aws

Guía para desarrolladores

Amazon Kendra

Amazon Kendra: Guía para desarrolladores

Table of Contents

..... xiii
¿Qué es Amazon Kendra? 1
Consulta de Amazon Kendra 1
Ventajas de Amazon Kendra 2
Ediciones de Amazon Kendra 2
Precios de Amazon Kendra 4
¿Es la primera vez que usa Amazon Kendra? 4
Cómo funciona Amazon Kendra 6
Índice 7
Uso de Amazon Kendra campos de documentos comunes o reservados 7
Buscar en índices 9
Documentos 9
Tipos o formatos de documentos 9
Atributos o campos del documento 12
Origen de datos 15
Consultas 17
Etiquetas 18
Etiquetado de recursos 18
Restricciones de las etiquetas 19
Configuración de Amazon Kendra 20
Inscríbase en AWS 20
Regiones y puntos de conexión 21
Configuración del AWS CLI 21
Configuración de los SDK AWS 22
IAM roles de acceso para Amazon Kendra 23
IAM roles para índices 23
IAM funciones para la BatchPutDocument API 26
IAM funciones para las fuentes de datos 29
Función de nube privada virtual (VPC) IAM 119
IAM funciones para las preguntas frecuentes (FAQ) 121
IAM funciones para sugerencias de consultas 123
IAM funciones para el mapeo principal de usuarios y grupos 124
IAM funciones para AWS IAM Identity Center 127
IAM roles para Amazon Kendra experiencias 128

IAM funciones para el enriquecimiento de documentos personalizados 131
Implementación de Amazon Kendra 135
Información general 136
Requisitos previos 136
Configurar el ejemplo 137
Página de búsqueda principal 138
Componente de búsqueda 138
Componente de resultados 138
Componente de facetas 138
Componente de paginación 139
Implementación de una aplicación de búsqueda sin código 139
Cómo funciona la búsqueda de Experience Builder 139
Diseñe y ajuste su experiencia de búsqueda 140
Proporcionar acceso a la página de búsqueda 141
Configurar una experiencia de búsqueda 142
Ajuste de la capacidad 147
Visualización de la capacidad 148
Agregar y eliminar capacidad 148
Amazon Kendra Capacidad de clasificación inteligente 149
Capacidad de sugerencias de consulta 149
Amazon Kendra capacidad de experiencia 149
Capacidad de experiencia de búsqueda 149
Ráfaga de consultas adaptativas 150
Introducción 151
Requisitos previos 151
Inscríbase en un Cuenta de AWS 151
Creación de un usuario con acceso administrativo 152
Amazon Kendra recursos: SDK AWS CLI, consola 154
Cómo empezar a utilizar la Amazon Kendra consola 160
Introducción (AWS CLI) 161
Introducción (SDK para Python [Boto3]) 162
Introducción (SDK para Java) 166
Introducción a S3 (consola) 170
Introducción a MySQL (consola) 171
Introducción a un origen de identidad de IAM Identity Center (consola) 174
Cambiar el origen de identidad de IAM Identity Center 177

Creación de un índice ... 178
Adición de documentos directamente a un índice mediante la carga por lotes ... 183
Añadir documentos con la BatchPutDocument API ... 184
Adición de documentos desde un bucket de S3 ... 186
Adición de preguntas frecuentes a un índice ... 189
Crear campos de índice para un archivo de preguntas frecuentes ... 190
Archivo CSV básico ... 191
Archivo CSV personalizado ... 191
Archivo JSON ... 193
Uso del archivo de preguntas frecuentes ... 195
Archivos de preguntas frecuentes en idiomas distintos del inglés ... 197
Creación de campos de documento personalizados ... 197
Actualización de campos de documentos personalizados ... 198
Control de acceso de usuarios a los documentos con tokens ... 201
Uso de OpenID ... 202
Uso de un JSON Web Token (JWT) con un secreto compartido ... 205
Uso de un JSON Web Token (JWT) con una clave pública ... 208
Uso de JSON ... 212
Creación de un conector de origen de datos ... 215
Establecimiento de un programa de actualizaciones ... 216
Configuración del idioma ... 216
Conectores de origen de datos ... 217
Esquemas de plantillas de origen de datos ... 218
Adobe Experience Manager ... 606
Alfresco ... 616
Aurora (MySQL) ... 625
Aurora (PostgreSQL) ... 634
Amazon FSx (Windows) ... 642
Amazon FSx (NetApp DISPONIBLE) ... 651
Amazon RDS/Aurora ... 660
Amazon RDS (Microsoft SQL Server) ... 669
Amazon RDS (MySQL) ... 678
Amazon RDS (Oracle) ... 686
Amazon RDS (PostgreSQL) ... 695
Amazon S3 ... 704
Amazon Kendra Rastreador web ... 721

Amazon WorkDocs 743
Box (Cuadro) 749
Confluence 757
Conector de orígenes de datos personalizados 777
Dropbox 786
Drupal 795
GitHub 806
Gmail 817
Google Drive 826
IBM DB2 845
Jira 854
Microsoft Exchange 861
Microsoft OneDrive 870
Microsoft SharePoint 886
Microsoft SQL Server 921
Microsoft Teams 930
Microsoft Yammer 942
MySQL 950
Oracle Database 958
PostgreSQL 967
Quip 975
Salesforce 982
ServiceNow 1001
Slack 1021
Zendesk 1032
Asignación de campos de origen de datos 1040
Uso de campos de documentos comunes o Amazon Kendra reservados 7
Adición de documentos en idiomas distintos del inglés 1046
Configuración Amazon Kendra para usar un Amazon VPC 1049
Configurando Amazon VPC 1049
Conectarse a Amazon VPC 1052
Conexión a una base de datos 1053
Solución de problemas de conexión de VPC 1056
Eliminar un índice, un origen de datos o documentos cargados por lotes 1059
Eliminación de un índice 1059
Eliminación de un origen de datos 1060

Eliminar documentos cargados por lotes ... 1062
Enriquecimiento de sus documentos durante la ingesta ... 1064
Cómo funciona Custom Document Enrichment ... 1064
Operaciones básicas para cambiar los metadatos ... 1065
Funciones de Lambda: extraer y cambiar metadatos o contenido ... 1074
Contratos de datos para funciones Lambda ... 1083
Formato del documento estructurado ... 1085
Ejemplo de una función Lambda que se adhiere a los contratos de datos ... 1085
Búsqueda en un índice ... 1089
Consulta de un índice ... 1089
Requisitos previos ... 1090
Buscar en un índice (consola) ... 1091
Buscar en un índice (SDK) ... 1091
Buscar en un índice (Postman) ... 1094
Búsqueda con una sintaxis de consulta avanzada ... 1095
Buscar en otros idiomas ... 1100
Recuperación de pasajes ... 1104
Navegar por un índice ... 1107
Destacar resultados de búsqueda ... 1110
Búsqueda tabular de HTML ... 1114
Sugerencias de consulta ... 1118
Sugerencias de consulta mediante el historial de consultas ... 1119
Sugerencias de consulta mediante los campos del documento ... 1125
Bloquear determinadas consultas o contenidos de los campos del documento para que no usen en las sugerencias ... 1130
Corrector ortográfico de las consultas ... 1136
Uso del corrector ortográfico de las consultas con los límites predeterminados ... 1137
Filtrado y búsqueda por facetas ... 1137
Facetas ... 1138
Utilizar atributos del documento para filtrar los resultados de búsqueda ... 1142
Filtrar los atributos de cada documento en los resultados de búsqueda ... 1144
Filtrar por contexto de usuario ... 1144
Filtrado por token de usuario ... 1145
Filtrado por ID de usuario y grupo ... 1146
Filtrado por atributo de usuario ... 1148
Filtrado por contexto de usuario para los documentos añadidos directamente a un índice . 1149

Filtrado por contexto de usuario para las preguntas más frecuentes 1150
Filtrado por contexto de usuario para los orígenes de datos 1150
Respuestas a las consultas y tipos de respuestas 1169
Respuestas a las consultas 1169
Tipos de respuestas 1173
Ajustar y ordenar las respuestas 1177
Ajustar las respuestas 1178
Ordenar las respuestas 1179
Contraer o expandir los resultados de la consulta 1181
Contraer los resultados 1184
Elegir un documento principal mediante la ordenación 1184
No hay una estrategia clave para el documento 1185
Expandir los resultados 1185
Interacciones con otras Amazon Kendra funciones 1185
Ajuste de la relevancia de las búsquedas 1187
Ajuste de la relevancia a nivel de índice 1188
Ajuste de la relevancia a nivel de consulta 1189
Obtener información con análisis de búsqueda 1191
Métricas de búsqueda 1191
Tasa de clics 1192
Tasa de clics cero 1192
Tasa de resultados de búsqueda cero 1193
Tasa de respuesta instantánea 1193
Consultas principales 1193
Consultas principales con cero clics 1194
Consultas principales con cero resultados de búsqueda 1194
La mayoría de las veces has hecho clic en los documentos 1194
Consultas totales 1195
Documentos totales 1195
Ejemplo de recuperación de datos métricos 1195
Desde métricas hasta información procesable 1197
Visualización y generación de informes sobre los análisis de búsqueda 1198
Gráfico de consultas totales 1198
Gráfico de tasa de clics 1198
Gráfico de tasa de clics cero 1199
Gráfico de tasa de resultados de búsqueda cero 1199

Gráfico de tasa de respuesta instantánea 1199
Envío de valoraciones para el aprendizaje incremental 1200
Usar la Amazon Kendra JavaScript biblioteca para enviar comentarios 1202
Paso 1: Inserta una etiqueta de script en tu aplicación de búsqueda Amazon Kendra 1202
Paso 2: Añadir el token de valoración a los resultados de búsqueda 1205
Paso 3: Probar la cadena de valoración 1205
Uso de la Amazon Kendra API para enviar comentarios 1206
Adición de sinónimos personalizados a un índice 1209
Crear un archivo de tesauro 1211
Adición de un tesauro a un índice 1214
Actualización de un tesauro 1218
Eliminar un tesauro 1222
Aspectos destacados en los resultados de búsqueda 1224
Creación de una solución de búsqueda inteligente 1225
Requisitos previos 1226
Paso 1: Añadir documentos 1227
Descarga del conjunto de datos de muestra 1228
Creación de un bucket de Amazon S3 1230
Crear carpetas de datos y metadatos en su bucket de S3 1232
Cargar los datos de entrada. 1235
Paso 2: Detectar entidades 1237
Ejecución de un trabajo de análisis de entidades de Amazon Comprehend 1237
Paso 3: Formatear los metadatos 1246
Descargar y extraer el resultado de Amazon Comprehend 1247
Cargar la salida en el bucket de S3 1251
Conversión de la salida al formato de metadatos de Amazon Kendra 1253
Limpieza del bucket de Amazon S3 1257
Paso 4: Creación de un índice e ingesta de los metadatos 1259
Creación de un índice de Amazon Kendra 1260
Actualización del rol de IAM para el acceso a Amazon S3 1268
Creación de campos de índice de búsqueda personalizados de Amazon Kendra 1271
Agregar el bucket de Amazon S3 como origen de datos para el índice 1276
Sincronización del índice de Amazon Kendra 1281
Paso 5: Consultar el índice 1283
Consulta del índice de Amazon Kendra 1284
Filtrar los resultados de búsqueda 1290

Paso 6: Limpieza 1294
Limpieza de los archivos 1294
..... 1295
Monitoreo y registro 1297
Monitorización de índices 1297
Supervisión de llamadas a la API de Amazon Kendra con CloudTrail 1301
Información de Amazon Kendra en CloudTrail 1301
Ejemplo: Entradas del archivo de registro de Amazon Kendra 1302
Monitorización de las llamadas a la API de Amazon Kendra Intelligent Ranking con CloudTrail 1303
Información de Amazon Kendra Intelligent Ranking en CloudTrail 1304
Ejemplo: Entradas del archivo de registro de Amazon Kendra Intelligent Ranking 1305
Monitorización de Amazon Kendra con CloudWatch 1306
Visualización de métricas de Amazon Kendra 1307
Creación de una alarma 1307
Métricas de CloudWatch para trabajos de sincronización de índices 1308
Métricas para Origen de datos de Amazon Kendra 1310
Métricas de los documentos indexados 1312
Monitorización de Amazon Kendra con registros de CloudWatch 1313
Flujos de registro de Origen de datos 1314
Flujo de registro de documentos 1315
Seguridad 1317
Protección de datos 1318
Cifrado en reposo 1319
Cifrado en tránsito 1319
Administración de claves 1319
Puntos de conexión de VPC (AWS PrivateLink) 1320
Consideraciones sobre los puntos de enlace de VPC de Amazon Kendra y Amazon Kendra Intelligent Ranking 1320
Creación de un punto final de VPC de interfaz para Amazon Kendra y Amazon Kendra Intelligent Ranking 1320
Creación de una política de puntos de conexión de VPC para Amazon Kendra y Amazon Kendra Intelligent Ranking 1321
Administración de identidades y accesos 1323
Público 1323
Autenticación con identidades 1324

Administración de acceso mediante políticas ... 1327
Cómo funciona Amazon Kendra con IAM ... 1330
Ejemplos de políticas basadas en identidades ... 1335
AWS políticas gestionadas ... 1341
Resolución de problemas ... 1346
Prácticas recomendadas de seguridad ... 1348
Aplicación del principio de privilegios mínimos ... 1348
Permisos de control de acceso basado en roles (RBAC) ... 1349
Registro y monitoreo en Amazon Kendra ... 1349
Validación de conformidad ... 1349
Resiliencia ... 1351
Seguridad de la infraestructura ... 1351
Configuración y análisis de vulnerabilidades ... 1352
Cuotas ... 1353
Regiones de admitidas ... 1353
Cuotas ... 1353
Cuotas indexadas ... 1353
Cuotas del conector de fuente de datos ... 1354
Preguntas frecuentes sobre cuotas ... 1355
Cuotas del tesauro ... 1356
Amazon Kendra cuotas de experiencia ... 1356
Cuotas de consultas y resultados de búsqueda ... 1356
Cuotas de sugerencias de consultas ... 1358
Cuotas de documentos ... 1360
Cuotas de resultados de búsqueda destacados ... 1361
Recupera las cuotas de resultados de búsqueda ... 1362
Solución de problemas ... 1364
Solución de problemas con los orígenes de datos ... 1364
No se han indexado mis documentos ... 1364
Ha fallado mi trabajo de sincronización ... 1365
Mi trabajo de sincronización está incompleto ... 1365
Mi trabajo de sincronización se ha realizado correctamente, pero no hay documentos indexados ... 1366
Tengo problemas con el formato de los archivos al sincronizar mi origen de datos ... 1367
Quiero generar un informe del historial de sincronización de mis documentos ... 1367
¿Cuánto tiempo lleva sincronizar un origen de datos? ... 1368

¿Cuánto cuesta sincronizar un origen de datos? 1368
Recibo un error Amazon EC2 de autorización 1368
No puedo usar los enlaces del índice de búsqueda para abrir mis Amazon S3 objetos 1368
Aparece un mensaje de error AccessDenied al usar un archivo de certificado SSL 1369
Aparece un error de autorización al utilizar una fuente de SharePoint datos 1369
Mi índice no rastrea los documentos de mi origen de datos de Confluence 1369
Solución de problemas con los resultados de búsqueda de documentos 1370
Los resultados de búsqueda no son relevantes para mi consulta de búsqueda 1370
¿Por qué solo veo 100 resultados? 1371
¿Por qué faltan los documentos que espero ver? 1371
¿Por qué veo documentos que tienen una política de ACL? 1371
Solución de problemas generales 1371
Intelligent Ranking Amazon Kendra 1373
Clasificación inteligente para autogestionarse OpenSearch 1373
Cómo funciona el complemento de búsqueda inteligente 1373
Configuración del complemento de búsqueda inteligente 1374
Interactuar con el complemento de búsqueda inteligente 1380
Comparar OpenSearch los resultados con Amazon Kendra los resultados 1386
Clasificación semántica de los resultados de un servicio de búsqueda 1387
Historial de documentos 1397
Referencia de la API 1413
Glosario de AWS 1414
..... mcdxv

¿Qué es Amazon Kendra?

Amazon Kendra es un servicio de búsqueda inteligente que utiliza el procesamiento del lenguaje natural y algoritmos avanzados de machine learning para proporcionar respuestas específicas a las preguntas de búsqueda a partir de sus datos.

A diferencia de la búsqueda tradicional basada en palabras clave, Amazon Kendra utiliza sus capacidades de comprensión semántica y contextual para decidir si un documento es relevante para una consulta de búsqueda. Proporciona respuestas específicas a las preguntas, lo que brinda a los usuarios una experiencia similar a la de interactuar con un experto humano.

> **Note**
>
> También puede utilizar las capacidades de búsqueda semántica de Amazon Kendra para volver a clasificar los resultados de otro servicio de búsqueda. Consulte Clasificación inteligente de Amazon Kendra para obtener más información.

Con Amazon Kendra, puede crear una experiencia de búsqueda unificada conectando varios repositorios de datos a un índice e incorporando y rastreando documentos. Puede usar los metadatos de sus documentos para crear una experiencia de búsqueda personalizada y con numerosas características para sus usuarios, ayudándoles a encontrar de manera eficiente las respuestas correctas a sus consultas.

¿Qué es Amazon Kendra?

Consulta de Amazon Kendra

Puede realizar los siguientes tipos de consultas en Amazon Kendra:

Preguntas triviales: preguntas sencillas sobre quién, qué, cuándo o dónde. Por ejemplo: ¿Dónde está el centro de servicios más cercano a Seattle? Las preguntas triviales tienen respuestas basadas en hechos que se pueden responder con una sola palabra o frase. La respuesta se obtiene a partir de una sección de preguntas frecuentes o de sus documentos indexados.

Preguntas descriptivas: preguntas en las que la respuesta puede ser una oración, un fragmento o un documento completo. Por ejemplo: ¿Cómo conecto el Echo Plus a la red? O bien: ¿Cómo puedo obtener los beneficios fiscales para las familias con ingresos bajos?

Preguntas sobre palabras clave y lenguaje natural: preguntas que incluyen contenido conversacional complejo cuyo significado puede no estar claro. Por ejemplo: discurso de apertura. Cuando Amazon Kendra encuentra una palabra como “apertura”, que tiene varios significados contextuales, deduce correctamente el significado de la consulta de búsqueda y proporciona la información relevante.

Ventajas de Amazon Kendra

Amazon Kendra es altamente escalable, capaz de satisfacer las demandas de rendimiento, está estrechamente integrado con otros servicios de AWS, como Amazon S3 y Amazon Lex, y ofrece seguridad de nivel empresarial. Alguno de los beneficios de usar Amazon Kendra son:

Simplicidad: Amazon Kendra proporciona una consola y una API para gestionar los documentos en los que desee buscar. Puede utilizar una API de búsqueda sencilla para integrar Amazon Kendra en las aplicaciones de sus clientes, como sitios web o aplicaciones móviles.

Conectividad: Amazon Kendra puede conectarse a repositorios de datos u orígenes de datos de terceros, como Microsoft SharePoint. Puede indexar y buscar fácilmente en sus documentos utilizando su origen de datos.

Precisión: a diferencia de los servicios de búsqueda tradicionales que utilizan búsquedas por palabras clave, Amazon Kendra intenta entender el contexto de la pregunta y proporciona la palabra, el fragmento o el documento más relevante para la consulta. Amazon Kendra utiliza el machine learning para mejorar los resultados de búsqueda a lo largo del tiempo.

Seguridad: Amazon Kendra ofrece una experiencia de búsqueda empresarial altamente segura. Los resultados de la búsqueda reflejan el modelo de seguridad de su organización y se pueden filtrar en función del acceso de los usuarios o grupos a los documentos. Los clientes son responsables de autenticar y autorizar el acceso de los usuarios.

Ediciones de Amazon Kendra

Amazon Kendra tiene dos versiones: Developer Edition y Enterprise Edition. En la siguiente tabla se describen sus características y las diferencias entre las dos.

Amazon Kendra Developer Edition

Amazon Kendra Developer Edition proporciona todas las características de Amazon Kendra a un precio menor.

Caso de uso ideal

- Probar cómo Amazon Kendra indexa sus documentos
- Probar las características
- Desarrollar aplicaciones que usen Amazon Kendra

Características

- Incluye un nivel gratuito con 750 horas de uso
- Hasta 5 índices con hasta 5 orígenes de datos cada uno
- 10 000 documentos o 3 GB de texto extraído
- Aproximadamente 4000 consultas al día o 0,05 consultas por segundo
- Se ejecuta en 1 zona de disponibilidad (AZ); consulte Zonas de disponibilidad (centros de datos en las regiones de AWS)

Limitaciones

- No admite aplicaciones de producción
- No ofrece garantías de latencia o disponibilidad

Amazon Kendra Enterprise Edition

Amazon Kendra Enterprise Edition proporciona todas las características de Amazon Kendra y está diseñada para los contextos de producción.

Caso de uso ideal

- Indexar todos los documentos de su empresa
- Implementar su aplicación en un entorno de producción

Características

- Hasta 5 índices con hasta 50 orígenes de datos cada uno
- 100 000 documentos o 30 GB de texto extraído
- Aproximadamente 8000 consultas al día o 0,1 consultas por segundo
- Se ejecuta en 3 zonas de disponibilidad (AZ); consulte Zonas de disponibilidad (centros de datos en las regiones de AWS)

> **Note**
>
> Puede aumentar esta cuota mediante la consola Service Quotas.

Limitaciones

- Ninguno

> **Note**
>
> Para obtener una lista con las regiones, los puntos de conexión y las cuotas de servicio que admite Amazon Kendra, consulte Puntos de conexión y cuotas de Amazon Kendra.

Precios de Amazon Kendra

Puede empezar de forma gratuita con la Amazon Kendra Developer Edition, que ofrece un uso de hasta 750 horas durante los primeros 30 días.

Cuando finalice la versión de prueba, se le cobrará por todos los índices de Amazon Kendra aprovisionados, incluso si están vacíos y no se ejecuta ninguna consulta. Una vez que finalice la versión de prueba, se cobrarán cargos adicionales por el escaneo y la sincronización de documentos mediante los orígenes de datos de Amazon Kendra.

Para obtener una lista completa de los cargos y precios, consulte Precios de Amazon Kendra.

¿Es la primera vez que usa Amazon Kendra?

Si es la primera vez que utiliza Amazon Kendra, le recomendamos que lea las siguientes secciones en orden:

1	2	3	4	5	6
Cómo funciona Amazon Kendra	Introducción	Creación de un índice	Adición de documentos directamente a un índice mediante la carga por lotes	Creación de un conector de origen de datos	Búsqueda en un índice
Presenta los componentes de Amazon Kendra y describe	Explica cómo configurar una cuenta y probar la API de búsqueda	Explica cómo usar Amazon Kendra para crear un índice de	Explica cómo añadir documentos directamente a un índice	Explica cómo añadir documentos de un repositorio	Explica cómo usar la API de búsqueda de Amazon Kendra para

1	2	3	4	5	6
Cómo funciona Amazon Kendra	Introducción	Creación de un índice	Adición de documentos directamente a un índice mediante la carga por lotes	Creación de un conector de origen de datos	Búsqueda en un índice
cómo utilizarlos para crear una solución de búsqueda.	de Amazon Kendra.	búsqueda y añadir orígenes de datos para sincronizar los documentos.	de Amazon Kendra.	de datos a un índice de Amazon Kendra.	buscar en un índice.

Cómo funciona Amazon Kendra

Amazon Kendra proporciona funciones de búsqueda a su aplicación. Indexa sus documentos directamente o desde un repositorio de documentos de terceros y proporciona información relevante a sus usuarios de forma inteligente. Se puede utilizar Amazon Kendra para crear un índice actualizable de documentos de diversos tipos. Para obtener una lista de los tipos de documentos compatibles, Amazon Kendra consulte Tipos de documentos.

Amazon Kendra se integra con otros servicios. Por ejemplo, puedes potenciar los bots de Amazon Lex chat con la función de Amazon Kendra búsqueda para proporcionar respuestas útiles a las preguntas de los usuarios. Puedes usar un Amazon Simple Storage Service depósito como fuente de datos para conectarte Amazon Kendra a tus documentos e indexarlos. Además, puede configurar políticas de acceso o permisos a los recursos que utilizan AWS Identity and Access Management.

Amazon Kendra tiene los siguientes componentes:

- Un índice que contiene sus documentos y permite buscarlos.
- Un origen de datos que almacena los documentos y al que se conecta Amazon Kendra . Puede sincronizar automáticamente una fuente de datos con un Amazon Kendra índice para que el índice permanezca actualizado con el repositorio de fuentes.
- Una API de adición de documentos que agrega documentos directamente a un índice.

Puede usarlo Amazon Kendra a través de la consola o la API. Puede crear, actualizar y eliminar índices. Al eliminar un índice, se eliminan todos sus conectores de fuentes de datos y se elimina permanentemente toda la información del documento. Amazon Kendra

Temas

- Índice
- Documentos
- Origen de datos
- Consultas
- Etiquetas

Índice

Un índice contiene el contenido de los documentos y está estructurado de forma que permite realizar búsquedas en ellos. La forma en que agrega los documentos al índice depende de cómo los almacena.

- Si almacena los documentos en algún tipo de repositorio, como un Amazon S3 depósito o un SharePoint sitio de Microsoft, utiliza un conector de fuente de datos para indexar los documentos del repositorio.
- Si no guardas los documentos en un repositorio, utilizas la BatchPutDocumentAPI para indexarlos directamente.
- En el caso de las preguntas y respuestas de las preguntas frecuentes, que deben almacenarse en un bucket de Amazon Kendra (Amazon S3), las carga desde ese bucket

Puede crear índices con la Amazon Kendra consola AWS CLI, el o un AWS SDK. Para obtener información sobre los tipos de documentos que se pueden indexar, consulte Tipos de documentos.

Uso de Amazon Kendra campos de documentos comunes o reservados

Con la UpdateIndex API, puede crear campos reservados o comunes utilizando `DocumentMetadataConfigurationUpdates` y especificando el nombre del campo de índice Amazon Kendra reservado para asignarlos al atributo o nombre de campo del documento equivalente. También puede crear campos personalizados. Si utiliza un conector de fuente de datos, la mayoría incluye asignaciones de campos que asignan los campos del documento de la fuente de datos a campos de indexación. Amazon Kendra Si utiliza la consola, los campos se actualizan seleccionando el origen de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de asignación de campos para configurar el origen de datos.

Puede configurar el objeto `Search` para establecer un campo como visualizable, facetable, buscable y ordenable. Puede configurar el objeto `Relevance` para establecer el orden de clasificación, duración de potenciación o período de tiempo de un campo para aplicarlos a los valores de potenciación, actualización, valor de importancia y valores de importancia asignados a valores de campo específicos. Si utiliza la consola, puede configurar los ajustes de búsqueda de un campo seleccionando la opción de faceta en el menú de navegación. Para configurar el ajuste de relevancia, seleccione la opción de buscar en su índice en el menú de navegación, introduzca una consulta y utilice las opciones del panel lateral para ajustar la relevancia de la búsqueda. No puede cambiar el tipo de campo una vez que este se ha creado.

Amazon Kendra tiene los siguientes campos de documento reservados o comunes que puede usar:

- `_authors`: una lista de uno o más autores responsables del contenido del documento.
- `_category`: una categoría que coloca un documento en un grupo específico.
- `_created_at`: la fecha y hora en formato ISO 8601 de creación del documento. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en el horario de Europa Central.
- `_data_source_id`: el identificador del origen de datos que contiene el documento.
- `_document_body`: el contenido del documento.
- `_document_id`: un identificador único del documento.
- `_document_title`: el título del documento.
- `_excerpt_page_number`: el número de página de un archivo PDF en el que aparece el extracto del documento. Si el índice se creó antes del 8 de septiembre de 2020, debe volver a indexar los documentos antes de poder utilizar este atributo.
- `_faq_id`: si se trata de un documento tipo pregunta-respuesta (preguntas frecuentes), un identificador único para las preguntas frecuentes.
- `_file_type`: el tipo de archivo del documento, como pdf o doc.
- `_last_updated_at`: la fecha y hora en formato ISO 8601 de última actualización del documento. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en el horario de Europa Central.
- `_source_uri`: el URI en el que está disponible el documento. Por ejemplo, el URI del documento en el sitio web de una empresa.
- `_version`: un identificador de la versión específica de un documento.
- `_view_count`: el número de veces que se ha visto el documento.
- `_language_code` (cadena): el código de un idioma que se aplica al documento. Este valor se define por defecto en inglés si no especifica un idioma. Para obtener más información acerca de los idiomas admitidos, incluidos sus códigos, consulte Adición de documentos en idiomas distintos del inglés.

En el caso de campos personalizados, estos campos se crean mediante `DocumentMetadataConfigurationUpdates` con la API `UpdateIndex`, del mismo modo que cuando se crea un campo reservado o común. Debe establecer el tipo de datos adecuado para el campo personalizado. Si utiliza la consola, los campos se actualizan seleccionando el origen de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de asignación

de campos para configurar el origen de datos. Algunos orígenes de datos no admiten la adición de campos nuevos o campos personalizados. No puede cambiar el tipo de campo una vez que este se ha creado.

Los siguientes son los tipos que puede configurar para los campos personalizados:

- Date
- Número
- Cadena
- Lista de cadenas

Si ha añadido documentos al índice mediante la BatchPutDocumentAPI, `Attributes` muestra los campos/atributos de los documentos y crea campos con el `DocumentAttribute` objeto.

En el caso de los documentos indexados a partir de una fuente de Amazon S3 datos, los campos se crean mediante un archivo de metadatos JSON que incluye la información de los campos.

Si utiliza una base de datos compatible como origen de datos, puede configurar los campos mediante la opción de asignación de campos.

Buscar en índices

Después de crear un índice, puede empezar a buscar en los documentos. Para obtener más información, consulte Buscar en índices.

Documentos

En esta sección se explica cómo Amazon Kendra indexa los numerosos formatos de documentos que admite y los diferentes campos o atributos de los documentos.

Temas

- Tipos o formatos de documentos
- Atributos o campos del documento

Tipos o formatos de documentos

Amazon Kendra admite los tipos o formatos de documentos más populares, como PDF, HTML PowerPoint, Word y más. Un índice puede contener varios formatos de documento.

Amazon Kendra extrae el contenido de los documentos para poder buscarlos. Los documentos se analizan de forma que se optimice la búsqueda en el texto extraído y en cualquier contenido tabular (tablas HTML) de los documentos. Esto significa estructurar los documentos en campos o atributos que se utilizan para la búsqueda. Los metadatos del documento, como la fecha de la última modificación, pueden ser campos útiles para la búsqueda.

Los documentos se pueden organizar en filas y columnas. Por ejemplo, cada documento es una fila y cada campo/atributo del documento, como el título y el contenido del cuerpo, es una columna. Por ejemplo, si utiliza una base de datos como origen de datos, los datos deben estructurarse u organizarse en filas y columnas.

Puede añadir documentos al índice de las siguientes maneras:

- API de BatchPutDocument
- Conector de origen de datos

Si desea añadir un archivo de preguntas frecuentes, utilice la CreateFaqAPI para añadir el archivo almacenado en un Amazon S3 depósito. Puede elegir entre un formato CSV básico, un formato CSV que incluya campos/atributos personalizados en un encabezado y un formato JSON que incluya campos personalizados. El formato predeterminado es CSV básico.

A continuación, se proporciona información sobre cada formato de documento compatible y cómo Amazon Kendra trata cada formato al indexar documentos.

Formato del documento	Tratado como	Cómo se trata el documento	Estructura original
Formato de documento portátil (PDF)	HTML	Se convierte a HTML y, a continuación, se extrae el contenido.	No estructurado
HyperText Lenguaje de marcado (HTML)	HTML	Las etiquetas HTML se filtran para extraer el contenido. El contenido debe estar entre las etiquetas HTML principales de inicio y cierre	Semiestructurado

Formato del documento	Tratado como	Cómo se trata el documento	Estructura original
		(<HTML>content</HTML>).	
Lenguaje de marcado extensible (XML)	XML	Las etiquetas XML se filtran para extraer el contenido.	Semiestructurado
Transformación del lenguaje de hojas de estilo extensible (XSLT)	XSLT	Las etiquetas se filtran para extraer el contenido.	Semiestructurado
MarkDown (MD)	Texto no cifrado	El contenido se extrae con MarkDown la sintaxis incluida.	Semiestructurado
Valores separados por comas (CSV)	CSV	Contenido extraído de cada celda, con un único archivo tratado como resultado de un único documento.	Estructurado para archivos de preguntas frecuentes; de lo contrario, semiestructurado
Microsoft Excel (XLS y XLSX)	XLS y XLSX	Contenido extraído de cada celda, con un único archivo tratado como resultado de un único documento.	Semiestructurado
JavaScript Notación de objetos (JSON)	Texto no cifrado	El contenido se extrae con la sintaxis JSON incluida.	Semiestructurado
Formato de texto enriquecido (RTF)	RTF	La sintaxis RTF se filtra para extraer el contenido.	Semiestructurado

Formato del documento	Tratado como	Cómo se trata el documento	Estructura original
Microsoft PowerPoint (PPT)	PPT	Solo se extrae el contenido de texto de las PowerPoin t diapositivas para su búsqueda. Las imágenes y otros contenidos no se extraen.	No estructurado
Microsoft Word (DOCX)	DOCX	Solo se extrae el contenido de texto de las páginas de Word para su búsqueda. Las imágenes y otros contenidos no se extraen.	No estructurado
Texto sin formato (TXT)	TXT	Se extrae todo el texto del documento de texto.	No estructurado

Atributos o campos del documento

Un documento tiene atributos o campos asociados. Los campos de un documento son las propiedades de un documento o lo que contiene su estructura. Por ejemplo, cada uno de sus documentos puede contener el título, el cuerpo del texto y el autor. También puede añadir campos personalizados para sus documentos específicos. Por ejemplo, si el índice busca documentos fiscales, puede especificar un campo personalizado para el tipo de documento fiscal, como el W-2, el 1099, etc.

Antes de poder usar un campo de documento en una consulta, debe asignarse a un campo de índice. Por ejemplo, el campo de título se puede asignar al campo `_document_title`. Para obtener más información, consulte Asignación de campos. Para agregar un campo nuevo, debe crear un

campo de índice al que asignarlo. Los campos de índice se crean mediante la consola o mediante la UpdateIndexAPI.

Puede usar los campos del documento para filtrar las respuestas y crear resultados de búsqueda facetados. Por ejemplo, puede filtrar una respuesta para que muestre solo una versión específica de un documento, o puede filtrar las búsquedas para que solo se muestren documentos fiscales del tipo 1099 que coincidan con el término de búsqueda. Para obtener más información, consulte Filtrado y búsqueda de facetas.

También puede utilizar los campos del documento para ajustar manualmente la respuesta a la consulta. Por ejemplo, puede optar por aumentar la importancia del campo de título para aumentar el peso que se le Amazon Kendra asigna al campo a la hora de determinar qué documentos devolver en la respuesta. Para más información, consulte Ajuste de la relevancia de la búsqueda.

Si agrega un documento directamente a un índice, debe especificar los campos en el parámetro de entrada del documento en la BatchPutDocumentAPI. Los valores de los campos personalizados se especifican en una matriz de DocumentAttributeobjetos. Si utiliza un origen de datos, el método que utilice para agregar los campos del documento depende del origen de datos. Para obtener más información, consulte Asignación de campos de origen de datos.

Uso de campos de documentos comunes o Amazon Kendra reservados

Con la UpdateIndex API, puede crear campos reservados o comunes utilizando `DocumentMetadataConfigurationUpdates` y especificando el nombre del campo de índice Amazon Kendra reservado para asignarlos al atributo o nombre de campo del documento equivalente. También puede crear campos personalizados. Si utiliza un conector de fuente de datos, la mayoría incluye asignaciones de campos que asignan los campos del documento de la fuente de datos a campos de indexación. Amazon Kendra Si utiliza la consola, los campos se actualizan seleccionando el origen de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de asignación de campos para configurar el origen de datos.

Puede configurar el objeto `Search` para establecer un campo como visualizable, facetable, buscable y ordenable. Puede configurar el objeto `Relevance` para establecer el orden de clasificación, duración de potenciación o período de tiempo de un campo para aplicarlos a los valores de potenciación, actualización, valor de importancia y valores de importancia asignados a valores de campo específicos. Si utiliza la consola, puede configurar los ajustes de búsqueda de un campo seleccionando la opción de faceta en el menú de navegación. Para configurar el ajuste de relevancia, seleccione la opción de buscar en su índice en el menú de navegación, introduzca una consulta y

utilice las opciones del panel lateral para ajustar la relevancia de la búsqueda. No puede cambiar el tipo de campo una vez que este se ha creado.

Amazon Kendra tiene los siguientes campos de documento reservados o comunes que puede usar:

- `_authors`: una lista de uno o más autores responsables del contenido del documento.
- `_category`: una categoría que coloca un documento en un grupo específico.
- `_created_at`: la fecha y hora en formato ISO 8601 de creación del documento. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en el horario de Europa Central.
- `_data_source_id`: el identificador del origen de datos que contiene el documento.
- `_document_body`: el contenido del documento.
- `_document_id`: un identificador único del documento.
- `_document_title`: el título del documento.
- `_excerpt_page_number`: el número de página de un archivo PDF en el que aparece el extracto del documento. Si el índice se creó antes del 8 de septiembre de 2020, debe volver a indexar los documentos antes de poder utilizar este atributo.
- `_faq_id`: si se trata de un documento tipo pregunta-respuesta (preguntas frecuentes), un identificador único para las preguntas frecuentes.
- `_file_type`: el tipo de archivo del documento, como pdf o doc.
- `_last_updated_at`: la fecha y hora en formato ISO 8601 de última actualización del documento. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en el horario de Europa Central.
- `_source_uri`: el URI en el que está disponible el documento. Por ejemplo, el URI del documento en el sitio web de una empresa.
- `_version`: un identificador de la versión específica de un documento.
- `_view_count`: el número de veces que se ha visto el documento.
- `_language_code` (cadena): el código de un idioma que se aplica al documento. Este valor se define por defecto en inglés si no especifica un idioma. Para obtener más información acerca de los idiomas admitidos, incluidos sus códigos, consulte Adición de documentos en idiomas distintos del inglés.

En el caso de campos personalizados, estos campos se crean mediante `DocumentMetadataConfigurationUpdates` con la API `UpdateIndex`, del mismo modo que

cuando se crea un campo reservado o común. Debe establecer el tipo de datos adecuado para el campo personalizado. Si utiliza la consola, los campos se actualizan seleccionando el origen de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de asignación de campos para configurar el origen de datos. Algunos orígenes de datos no admiten la adición de campos nuevos o campos personalizados. No puede cambiar el tipo de campo una vez que este se ha creado.

Los siguientes son los tipos que puede configurar para los campos personalizados:

- Date
- Número
- Cadena
- Lista de cadenas

Si ha añadido documentos al índice mediante la BatchPutDocumentAPI, `Attributes` muestra los campos/atributos de los documentos y crea campos con el `DocumentAttribute` objeto.

En el caso de los documentos indexados a partir de una fuente de Amazon S3 datos, los campos se crean mediante un archivo de metadatos JSON que incluye la información de los campos.

Si utiliza una base de datos compatible como origen de datos, puede configurar los campos mediante la opción de asignación de campos.

Origen de datos

Una fuente de datos es un repositorio o ubicación de datos que Amazon Kendra se conecta a los documentos o el contenido y los indexa. Por ejemplo, puede configurarlo Amazon Kendra para conectarse a Microsoft SharePoint para rastrear e indexar los documentos almacenados en esta fuente. También puede indexar las páginas web proporcionando las direcciones URL para Amazon Kendra rastrearlas. Puede sincronizar automáticamente una fuente de datos con un Amazon Kendra índice para que los documentos agregados, actualizados o eliminados en la fuente de datos también se agreguen, actualicen o eliminen en el índice.

Los orígenes de datos admitidos son:

- Adobe Experience Manager
- Alfresco
- Aurora (MySQL)

- Aurora (PostgreSQL)
- Amazon FSx (Windows)
- Amazon FSx (NetApp ONTAP)
- Orígenes de datos de bases de datos
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon S3 cubos
- Amazon Kendra Rastreador web
- Amazon WorkDocs
- Box
- Confluence
- Orígenes de datos personalizados
- Dropbox
- Drupal
- GitHub
- Gmail
- Unidades de Workspace de Google
- IBM DB2
- Jira
- Microsoft Exchange
- Microsoft OneDrive
- Microsoft SharePoint
- Microsoft Teams
- Microsoft SQL Server
- Microsoft Yammer
- MySQL
- Oracle Database
- PostgreSQL

- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

Para obtener una lista de los tipos o formatos de documentos compatibles, Amazon Kendra consulte Tipos de documentos. Primero debe crear un índice, antes de crear un conector de origen de datos para indexar los documentos a partir del origen de datos.

> **Note**
>
> Para crear un índice de documentos, no es necesario utilizar un origen de datos. Puede agregar documentos directamente a un índice mediante la carga por lotes. Para obtener más información, consulte Añadir documentos directamente a un índice.

Para ver un tutorial sobre el uso de la Amazon Kendra consola, la AWS CLI o los SDK, consulte Primeros pasos.

Consultas

Para obtener respuestas, los usuarios consultan un índice. Los usuarios pueden utilizar el lenguaje natural en sus consultas. La respuesta contiene información, como el título, un extracto del texto y la ubicación de los documentos en el índice que proporcionan la mejor respuesta.

Amazon Kendra utiliza toda la información que se proporciona sobre los documentos, no solo el contenido de los mismos, para determinar si un documento es relevante para la consulta. Por ejemplo, si el índice contiene información sobre cuándo se actualizaron los documentos por última vez, puede asignar una mayor relevancia a los documentos que se actualizaron más recientemente. Amazon Kendra

Una consulta también puede contener criterios sobre cómo filtrar la respuesta de modo que solo se Amazon Kendra devuelvan los documentos que cumplan los criterios de filtro. Por ejemplo, si ha creado un campo de índice denominado departamento, puede filtrar la respuesta para que solo se devuelvan los documentos con el campo de departamento establecido como legal. Para obtener más información, consulte Filtrado de búsquedas.

Puede influir en los resultados de una consulta ajustando la relevancia de los campos individuales del índice. El ajuste cambia la importancia de un campo en los resultados. Por ejemplo, si eleva la importancia de los documentos de la categoría nuevos, es más probable que los documentos de esta categoría se incluyan en la respuesta. Para más información, consulte Ajuste de la relevancia de la búsqueda.

Para más información acerca del uso de consultas, consulte Buscar en un índice.

Etiquetas

Administre sus índices, orígenes de datos y preguntas frecuentes mediante la asignación de etiquetas. Puede usar etiquetas para clasificar Amazon Kendra los recursos de varias maneras. Por ejemplo, según su finalidad, propietario o aplicación, o cualquier combinación. Cada etiqueta consta de una clave y un valor, ambos definidos por el usuario.

Las etiquetas ayudan a:

- Identifique y organice sus AWS recursos. Muchos AWS servicios admiten el etiquetado, por lo que puede asignar la misma etiqueta a los recursos de diferentes servicios para indicar que los recursos están relacionados. Por ejemplo, puedes etiquetar un índice y el Amazon Lex bot que lo usa con la misma etiqueta.
- Asignar costos. Las etiquetas se activan en el AWS Billing and Cost Management panel de control. AWS usa etiquetas para categorizar sus costos y entregarle un informe mensual de asignación de costos. Para obtener más información, consulte Cost Allocation and Tagging en About AWS Billing and Cost Management.
- Controle el acceso a los recursos. Puede utilizar etiquetas en políticas de AWS Identity and Access Management (IAM) que controlan el acceso a los recursos de Amazon Kendra . Puede adjuntar estas políticas a un IAM rol o usuario para activar el control de acceso basado en etiquetas. Para obtener más información, consulte Autorización basada en etiquetas.

Puede crear y administrar etiquetas mediante la AWS Management Console, la AWS Command Line Interface (AWS CLI) o la Amazon Kendra API.

Etiquetado de recursos

Si utilizas la Amazon Kendra consola, puedes etiquetar los recursos al crearlos o añadirlos más adelante. También puede utilizar la consola para actualizar o quitar etiquetas.

Si usas AWS Command Line Interface (AWS CLI) o la Amazon Kendra API, usa las siguientes operaciones para administrar las etiquetas de tus recursos:

- CreateDataSource—Aplique etiquetas al crear una fuente de datos.
- CreateFaq—Aplica etiquetas al crear una sección de preguntas frecuentes.
- CreateIndex—Aplique etiquetas al crear un índice.
- ListTagsForResource—Ver las etiquetas asociadas a un recurso.
- TagResource—Añadir y modificar las etiquetas de un recurso.
- UntagResource—Eliminar etiquetas de un recurso.

Restricciones de las etiquetas

Las siguientes restricciones se aplican a las etiquetas de los Amazon Kendra recursos:

- Número máximo de etiquetas: 50
- Longitud máxima de la clave: 128 caracteres
- Longitud máxima del valor: 256 caracteres
- Caracteres válidos para claves y valores: a-z, A-Z, espacio y los siguientes caracteres: _ . : / = + - y @
- Las claves y los valores distinguen entre mayúsculas y minúsculas
- No utilice `aws:` como prefijo para claves, ya que está reservado para AWS .

Configuración de Amazon Kendra

Antes de utilizar Amazon Kendra, debe tener una cuenta de Amazon Web Services (AWS). Una vez que tenga una AWS cuenta, podrá acceder a Amazon Kendra a través de la consola Amazon Kendra, AWS Command Line Interface el AWS CLI () o los SDK. AWS

Esta guía incluye ejemplos AWS CLI de Java y Python.

Temas

- Inscríbase en AWS
- Regiones y puntos de conexión
- Configurar el AWS CLI
- Configuración de los SDK AWS

Inscríbase en AWS

Cuando te registras en Amazon Web Services (AWS), tu cuenta se registra automáticamente en todos los servicios de Amazon AWS Kendra, incluido Amazon. Solo se le cobrará por los servicios que utilice.

Si ya tienes una AWS cuenta, pasa a la siguiente tarea. Si no dispone de una cuenta de AWS , utilice el siguiente procedimiento para crear una.

Para registrarte en AWS

1. Abre https://aws.amazon.com y, a continuación, selecciona Crear una AWS cuenta.
2. Siga las instrucciones que aparecen en pantalla para completar la creación de la cuenta. Anote su número de cuenta de AWS de 12 dígitos. Parte del procedimiento de inscripción consiste en recibir una llamada telefónica e introducir un número PIN con el teclado del teléfono.
3. Cree un usuario administrador AWS Identity and Access Management (IAM). Para obtener instrucciones, consulte Creación del primer grupo y usuario administrador de IAM en la Guía del usuario de AWS Identity and Access Management .

Regiones y puntos de conexión

Un punto de enlace es una URL que es el punto de entrada de un servicio web. Cada punto final está asociado a una AWS región específica. Si utiliza una combinación de la consola de Amazon Kendra y los SDK de Amazon Kendra AWS CLI, preste atención a sus regiones predeterminadas, ya que todos los componentes de Amazon Kendra de una campaña determinada (índice, consulta, etc.) deben crearse en la misma región. Para conocer las regiones y puntos de conexión que admite Amazon Kendra, consulte Regiones y puntos de conexión.

Configurar el AWS CLI

La interfaz de línea de AWS comandos (AWS CLI) es una herramienta de desarrollador unificada para administrar AWS servicios, incluido Amazon Kendra. Recomendamos que la instale.

1. Para instalarla AWS CLI, siga las instrucciones de la AWS Guía del usuario de la interfaz de línea de AWS comandos.
2. Para configurar AWS CLI y configurar un perfil al que llamar AWS CLI, siga las instrucciones de la Guía del usuario sobre la AWS CLI configuración de la interfaz de línea de AWS comandos.
3. Para confirmar que el AWS CLI perfil está configurado correctamente, ejecute el siguiente comando:

   ```
   aws configure --profile default
   ```

   Si el perfil está configurado correctamente, la salida debería ser similar a la siguiente:

   ```
   AWS Access Key ID [****************52FQ]:
   AWS Secret Access Key [****************xgyZ]:
   Default region name [us-west-2]:
   Default output format [json]:
   ```

4. Para comprobar que AWS CLI está configurado para su uso con Amazon Kendra, ejecute los siguientes comandos:

   ```
   aws kendra help
   ```

   Si AWS CLI está configurado correctamente, verá una lista de los AWS CLI comandos compatibles con Amazon Kendra, Amazon Kendra runtime y Amazon Kendra events.

Configuración de los SDK AWS

Descarga e instala los AWS SDK que quieras usar. En esta guía, se ofrecen ejemplos para Python. Para obtener información sobre otros AWS SDK, consulte Herramientas para Amazon Web Services.

El paquete para el SDK de Python se llama Boto3.

Antes de ejecutar los siguientes comandos de Python, primero debe descargar e instalar Python 3.6 o versiones posteriores para su sistema operativo. El soporte para Python 3.5 y versiones anteriores está obsoleto. Si no tiene pip incluido en su directorio de cadenas de Python, puede descargar get-pip.py y almacenarlo en su directorio de cadenas. También puede configurar su directorio de Python como una ruta o variable de entorno mediante un programa de terminal.

```
# Install the latest Boto3 release via pip
pip install boto3

# You can install a specific version of Boto3 for compatibility reasons
# Install Boto3 version 1.0 specifically
pip install boto3==1.0.0

# Make sure Boto3 is no older than version 1.15.0
pip install boto3>=1.15.0

# Avoid versions of Boto3 newer than version 1.15.3
pip install boto3<=1.15.3
```

Para usar Boto3, debe configurar las credenciales de autenticación de su AWS cuenta mediante la consola de IAM.

IAM roles de acceso para Amazon Kendra

Al crear un índice, una fuente de datos o una sección de preguntas frecuentes, Amazon Kendra necesita acceder a los AWS recursos necesarios para crear el Amazon Kendra recurso. Debe crear una política AWS Identity and Access Management (IAM) antes de crear el Amazon Kendra recurso. Al llamar a la operación, debe proporcionar el nombre de recurso de Amazon (ARN) del rol con la política adjunta. Por ejemplo, si llamas a la BatchPutDocumentAPI para añadir documentos desde un Amazon S3 depósito, Amazon Kendra asignas un rol con una política que tenga acceso al depósito.

Puedes crear un IAM rol nuevo en la Amazon Kendra consola o elegir uno IAM existente para usarlo. La consola muestra los roles que tienen la cadena “kendra” o “Kendra” en el nombre del rol.

En los temas siguientes se proporcionan detalles sobre las políticas necesarias. Si crea IAM roles mediante la Amazon Kendra consola, estas políticas se crean automáticamente.

Temas

- IAM roles para índices
- IAM funciones para la BatchPutDocument API
- IAM funciones para las fuentes de datos
- Función de nube privada virtual (VPC) IAM
- IAM funciones para las preguntas frecuentes (FAQ)
- IAM funciones para sugerencias de consultas
- IAM funciones para el mapeo principal de usuarios y grupos
- IAM funciones para AWS IAM Identity Center
- IAM roles para Amazon Kendra experiencias
- IAM funciones para el enriquecimiento de documentos personalizados

IAM roles para índices

Al crear un índice, debe proporcionar un IAM rol con permiso para escribir en un Amazon CloudWatch. También debe proporcionar una política de confianza que Amazon Kendra permita asumir el rol. Las siguientes son las políticas que se deben proporcionar.

IAM funciones para los índices

Una política de roles para permitir Amazon Kendra el acceso a un CloudWatch registro.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "cloudwatch:PutMetricData",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "cloudwatch:namespace": "AWS/Kendra"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": "logs:DescribeLogGroups",
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": "logs:CreateLogGroup",
            "Resource": "arn:aws:logs:your-region:your-account-id:log-group:/aws/
kendra/*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "logs:DescribeLogStreams",
                "logs:CreateLogStream",
                "logs:PutLogEvents"
            ],
            "Resource": "arn:aws:logs:your-region:your-account-id:log-group:/aws/
kendra/*:log-stream:*"
        }
    ]
}
```

Una política de roles para Amazon Kendra permitir el acceso AWS Secrets Manager. Si utiliza el contexto de usuario Secrets Manager como ubicación clave, puede utilizar la siguiente política.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Action":"cloudwatch:PutMetricData",
         "Resource":"*",
         "Condition":{
            "StringEquals":{
               "cloudwatch:namespace":"AWS/Kendra"
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":"logs:DescribeLogGroups",
         "Resource":"*"
      },
      {
         "Effect":"Allow",
         "Action":"logs:CreateLogGroup",
         "Resource":"arn:aws:logs:your-region:your-account-id:log-group:/aws/kendra/*"
      },
      {
         "Effect":"Allow",
         "Action":[
            "logs:DescribeLogStreams",
            "logs:CreateLogStream",
            "logs:PutLogEvents"
         ],
         "Resource":"arn:aws:logs:your-region:your-account-id:log-group:/aws/kendra/
*:log-stream:*"
      },
      {
         "Effect":"Allow",
         "Action":[
            "secretsmanager:GetSecretValue"
         ],
         "Resource":[
            "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
         ]
      },
      {
```

```
         "Effect":"Allow",
         "Action":[
            "kms:Decrypt"
         ],
         "Resource":[
            "arn:aws:kms:your-region:your-account-id:key/key-id"
         ],
         "Condition":{
            "StringLike":{
               "kms:ViaService":[
                  "secretsmanager.your-region.amazonaws.com"
               ]
            }
         }
      }
   ]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para la BatchPutDocument API

> **Warning**
>
> Amazon Kendra no utiliza una política de bucket que conceda permisos a un Amazon Kendra director para interactuar con un bucket de S3. En su lugar, usa roles de IAM . Asegúrese de Amazon Kendra no incluirlo como miembro de confianza en su política de bucket para evitar cualquier problema de seguridad de los datos si se conceden permisos accidentalmente

a directores arbitrarios. Sin embargo, puedes añadir una política de grupos para utilizarlos Amazon S3 en distintas cuentas. Para obtener más información, consulte Políticas de uso de Amazon S3 en varias cuentas. Para obtener más información sobre los roles de IAM para orígenes de datos de S3, consulte Roles de IAM.

Cuando utilizas la BatchPutDocumentAPI para indexar documentos en un Amazon S3 depósito, debes proporcionar Amazon Kendra un IAM rol con acceso al depósito. También debes proporcionar una política de confianza que te Amazon Kendra permita asumir el rol. Si los documentos del depósito están cifrados, debe dar permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar los documentos.

IAM funciones para la API BatchPutDocument

Una política de roles obligatoria para permitir Amazon Kendra el acceso a un Amazon S3 bucket.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
```

```
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Se recomienda incluir `aws:sourceAccount` y `aws:sourceArn` en la política de confianza. Esto limita los permisos y comprueba de forma segura si `aws:sourceAccount` los mismos que se indican en la política de IAM roles de la `sts:AssumeRole` acción. `aws:sourceArn` Esto evita que entidades no autorizadas accedan a tus IAM funciones y a sus permisos. Para obtener más información, consulte la AWS Identity and Access Management guía sobre el problema de los diputados confusos.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "kendra.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "your-account-id"
                },
                "StringLike": {
                    "aws:SourceArn": "arn:aws:kendra:your-region:your-account-id:index/
*"
                }
            }
        }
    ]
}
```

Una política de funciones opcional que permite Amazon Kendra utilizar una clave maestra AWS KMS del cliente (CMK) para descifrar los documentos de un Amazon S3 depósito.

```
{
    "Version": "2012-10-17",
    "Statement": [
```

```
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

IAM funciones para las fuentes de datos

Al utilizar la CreateDataSourceAPI, debe asignar Amazon Kendra un IAM rol que tenga permiso para acceder a los recursos. Los permisos específicos necesarios dependen del origen de datos.

IAM funciones para las fuentes de datos de Adobe Experience Manager

Cuando se utiliza Adobe Experience Manager, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su Adobe Experience Manager.
- Permiso para llamar a las API públicas necesarias para el conector de Adobe Experience Manager.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos de Adobe Experience Manager a Amazon Kendra través Amazon VPC de. Si está utilizando un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
```

```
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
```

```
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Alfresco

Cuando se utiliza Alfresco, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su Alfresco.
- Permiso para llamar a las API públicas necesarias para el conector de Alfresco.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos de Alfresco a través de. Amazon Kendra Amazon VPC Si utiliza un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
```

```
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para fuentes de datos Aurora (MySQL)

Cuando usas Aurora (MySQL), proporcionas un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su Aurora (MySQL).
- Permiso para llamar a las API públicas requeridas para el conector Aurora (MySQL).
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos Aurora (MySQL) a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
```

```
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
```

```
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para fuentes de datos Aurora (PostgreSQL)

Cuando utiliza Aurora (PostgreSQL), proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su Aurora (PostgreSQL).
- Permiso para llamar a las API públicas necesarias para el conector de Aurora (PostgreSQL).
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos Aurora (PostgreSQL) a través de. Amazon Kendra Amazon VPC Si está utilizando un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
```

```
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
```

```
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes Amazon FSx de datos

Cuando los usa Amazon FSx, proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su sistema de Amazon FSx archivos.
- Permiso de acceso Amazon Virtual Private Cloud (VPC) al lugar donde reside su sistema de Amazon FSx archivos.
- Permiso para obtener el nombre de dominio de Active Directory para su sistema de Amazon FSx archivos.
- Permiso para llamar a las API públicas necesarias para el conector de Amazon FSx .
- Permiso para llamar a las API `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.

```
{
    "Version": "2012-10-17",
        "Statement": [
        {
          "Effect": "Allow",
          "Action": [
            "secretsmanager:GetSecretValue"
          ],
          "Resource": [
            "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:
{{secret-id}}"
          ]
        },
        {
          "Effect": "Allow",
          "Action": [
```

```
            "kms:Decrypt"
          ],
          "Resource": [
            "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/{{key-id}}"
          ],
          "Condition": {
            "StringLike": {
              "kms:ViaService": [
                "secretsmanager.{{your-region}}.amazonaws.com"
              ]
            }
          }
        },
        {
          "Effect": "Allow",
          "Action":[
            "ec2:CreateNetworkInterface",
            "ec2:DeleteNetworkInterface"
          ],
          "Resource": [
                "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-interface/*",
                "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-ids]]"
          ]
        },
        {
          "Effect": "Allow",
          "Action": [
            "ec2:DescribeSubnets",
            "ec2:DescribeNetworkInterfaces"
          ],
          "Resource": "*"
        },
        {
          "Effect": "Allow",
          "Action": [
            "ec2:CreateNetworkInterfacePermission"
          ],
          "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-
interface/*",
          "Condition": {
            "StringEquals": {
              "ec2:AuthorizedService": "kendra.*.amazonaws.com"
            },
            "ArnEquals": {
```

```
            "ec2:Subnet": [
              "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-ids]]"
            ]
          }
        }
      },
      {
        "Sid": "AllowsKendraToGetDomainNameOfActiveDirectory",
        "Effect": "Allow",
        "Action": "ds:DescribeDirectories",
        "Resource": "*"
      },
      {
        "Sid": "AllowsKendraToCallRequiredFsxAPIs",
        "Effect": "Allow",
        "Action": [
            "fsx:DescribeFileSystems"
        ],
        "Resource": "*"
      },
      {
        "Sid": "iamPassRole",
        "Effect": "Allow",
        "Action": "iam:PassRole",
        "Resource": "*",
        "Condition": {
          "StringEquals": {
            "iam:PassedToService": [
              "kendra.*.amazonaws.com"
            ]
          }
        }
      },
      {
        "Effect": "Allow",
        "Action": [
          "kendra:BatchPutDocument",
          "kendra:BatchDeleteDocument"
        ],
        "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/
{{index-id}}"
      }
      ]
```

```
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de bases de datos

Cuando utiliza una base de datos como fuente de datos, Amazon Kendra proporciona un rol que tiene los permisos necesarios para conectarse a. Entre ellos se incluyen:

- Permiso para acceder al AWS Secrets Manager secreto que contiene el nombre de usuario y la contraseña del sitio. Para obtener más información sobre el contenido del secreto, consulte orígenes de datos.
- Permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el nombre de usuario y la contraseña secretos almacenados en. Secrets Manager
- Permiso para utilizar las operaciones `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.
- Permiso para acceder al Amazon S3 depósito que contiene el certificado SSL utilizado para comunicarse con el sitio.

> **Note**
>
> Puede conectar las fuentes de datos de la base de datos a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:GetSecretValue"
            ],
            "Resource": [
                "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": [
                "arn:aws:kendra:your-region:your-account-id:index/index-id"
            "Condition": {
                "StringLike": {
                    "kms:ViaService": [
                        "kendra.your-region.amazonaws.com"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
```

```
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Hay dos políticas opcionales que puede utilizar con un origen de datos.

Si has cifrado el Amazon S3 depósito que contiene el certificado SSL utilizado para comunicarse con él, proporciona una política para dar Amazon Kendra acceso a la clave.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

Si utiliza una VPC, proporcione una política que dé Amazon Kendra acceso a los recursos necesarios. Consulte Roles de IAM para los orígenes de datos y VPC para ver la política requerida.

Una política de confianza que permita Amazon Kendra asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
```

```
}
```

IAM roles para fuentes de datos Amazon RDS (Microsoft SQL Server)

Cuando utiliza un conector de fuente de datos Amazon RDS (Microsoft SQL Server), proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su instancia de fuente de datos Amazon RDS (Microsoft SQL Server).
- Permiso para llamar a las API públicas necesarias para el conector de fuente de datos Amazon RDS (Microsoft SQL Server).
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos Amazon RDS (Microsoft SQL Server) a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
```

```
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
```

```
      }
   ]
}
```

IAM roles para fuentes de datos Amazon RDS (MySQL)

Cuando utiliza un conector de fuente de datos Amazon RDS (MySQL), proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su instancia de fuente de datos Amazon RDS (MySQL).
- Permiso para llamar a las API públicas requeridas para el conector de fuente de datos Amazon RDS (MySQL).
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos Amazon RDS (MySQL) a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
```

```
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
```

```
            "Action":"sts:AssumeRole"
        }
    ]
}
```

IAM funciones para las fuentes de datos Amazon RDS (Oracle)

Cuando utiliza un conector de fuentes de datos de Amazon RDS Oracle, proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su instancia de fuente de datos Amazon RDS (Oracle).
- Permiso para llamar a las API públicas necesarias para el conector de la fuente de datos Amazon RDS (Oracle).
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos de Amazon RDS Oracle a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
```

```
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
```

```
        },
        "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para fuentes de datos Amazon RDS (PostgreSQL)

Cuando utiliza un conector de fuente de datos Amazon RDS (PostgreSQL), proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su instancia de fuente de datos Amazon RDS (PostgreSQL).
- Permiso para llamar a las API públicas necesarias para el conector de origen de datos de Amazon RDS (PostgreSQL).
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos Amazon RDS (PostgreSQL) a través de. Amazon Kendra Amazon VPC Si está utilizando un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
```

```
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
```

```
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes Amazon S3 de datos

> **Warning**
>
> Amazon Kendra no utiliza una política de bucket que conceda permisos a un Amazon Kendra director para interactuar con un bucket de S3. En su lugar, usa IAM roles. Asegúrate de Amazon Kendra no incluirlo como miembro de confianza en tu política de grupos para evitar problemas de seguridad de los datos al conceder permisos accidentalmente a directores arbitrarios. Sin embargo, puede añadir una política de bucket para utilizar un bucket de Amazon S3 en distintas cuentas. Para obtener más información, consulte Políticas para usar Amazon S3 en varias cuentas (más abajo).

Cuando utiliza un Amazon S3 depósito como fuente de datos, proporciona un rol que tiene permiso para acceder al depósito y utilizar las `BatchDeleteDocument` operaciones `BatchPutDocument` y. Si los documentos del Amazon S3 depósito están cifrados, debe dar permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar los documentos.

Las siguientes políticas de rol deben permitir Amazon Kendra asumir un rol. Desplácese más abajo para ver una política de confianza para asumir un rol.

Una política de roles obligatoria para Amazon Kendra permitir el uso de un Amazon S3 bucket como fuente de datos.

```
{
    "Version": "2012-10-17",
    "Statement": [
         {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
```

```
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name"
            ],
            "Effect": "Allow"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": [
                "arn:aws:kendra:your-region:your-account-id:index/index-id"
            ]
        }
    ]
}
```

Una política de funciones opcional que permite Amazon Kendra utilizar una clave maestra AWS KMS del cliente (CMK) para descifrar los documentos de un Amazon S3 depósito.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

Una política de roles opcional que permite acceder Amazon Kendra a un Amazon S3 depósito mientras se usa un Amazon VPC depósito y sin activar AWS KMS ni compartir AWS KMS permisos.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::{{bucket-name}}/*"
      ],
      "Effect": "Allow"
    },
    {
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::{{bucket-name}}"
      ],
      "Effect": "Allow"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": [
        "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-ids]]",
        "arn:aws:ec2:{{your-region}}:{{your-account-id}}:security-group/[[security-
group]]"
      ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "ec2:CreateNetworkInterface"
        ],
        "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-interface/
*",
        "Condition": {
```

```
            "StringLike": {
                "aws:RequestTag/AWS_KENDRA": "kendra_{{your-account-id}}_{{index-
id}}_{data-source-id}}_*"
            }
        }
    },
    {
        "Effect": "Allow",
        "Action": [
            "ec2:CreateTags"
        ],
        "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-interface/
*",
        "Condition": {
            "StringEquals": {
                "ec2:CreateAction": "CreateNetworkInterface"
            }
        }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeSubnets"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeNetworkInterfaces"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterfacePermission"
      ],
      "Resource": "arn:aws:ec2:{{your-region}}:{{your-accoount-id}}:network-interface/
*",
      "Condition": {
        "StringEquals": {
          "ec2:AuthorizedService": "kendra.amazonaws.com"
        },
```

```
        "ArnEquals": {
          "ec2:Subnet": [
            "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-ids]]"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
      ],
      "Resource": [
        "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}",
        "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
        "kendra:BatchDeleteDocument"
      ],
      "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}"
    }
  ]
}
```

Una política de roles opcional que permite acceder Amazon Kendra a un Amazon S3 bucket mientras se usa un Amazon VPC bucket y con AWS KMS los permisos activados.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetObject"
      ],
```

```
      "Resource": [
        "arn:aws:s3:::{{bucket-name}}/*"
      ],
      "Effect": "Allow"
    },
    {
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::{{bucket-name}}"
      ],
      "Effect": "Allow"
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt"
      ],
      "Resource": [
        "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/{{key-id}}"
      ],
      "Condition": {
        "StringLike": {
          "kms:ViaService": [
            "s3.{{your-region}}.amazonaws.com"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": [
        "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-ids]]",
        "arn:aws:ec2:{{your-region}}:{{your-account-id}}:security-group/[[security-
group]]"
      ]
    },
    {
        "Effect": "Allow",
        "Action": [
```

```
            "ec2:CreateNetworkInterface"
        ],
        "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-interface/
*",
        "Condition": {
            "StringLike": {
                "aws:RequestTag/AWS_KENDRA": "kendra_{{your-account-id}}_{{index-
id}}_{data-source-id}}_*"
            }
        }
    },
    {
        "Effect": "Allow",
        "Action": [
            "ec2:CreateTags"
        ],
        "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-interface/
*",
        "Condition": {
            "StringEquals": {
                "ec2:CreateAction": "CreateNetworkInterface"
            }
        }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeSubnets"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeNetworkInterfaces"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterfacePermission"
      ],
```

```
      "Resource": "arn:aws:ec2:{{your-region}}:{{your-account-id}}:network-interface/
*",
      "Condition": {
        "StringEquals": {
          "ec2:AuthorizedService": "kendra.amazonaws.com"
        },
        "ArnEquals": {
          "ec2:Subnet": [
            "arn:aws:ec2:{{your-region}}:{{your-account-id}}:subnet/[[subnet-ids]]"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
      ],
      "Resource": [
        "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}",
        "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
        "kendra:BatchDeleteDocument"
      ],
      "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}"
    }
  ]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
```

```
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Políticas para usar Amazon S3 en varias cuentas

Si tu Amazon S3 grupo está en una cuenta diferente a la que usas para tu Amazon Kendra índice, puedes crear políticas para usarlo en todas las cuentas.

Una política de rol para usar el Amazon S3 depósito como fuente de datos cuando el depósito se encuentre en una cuenta diferente a la de su Amazon Kendra índice. Tenga en cuenta que `s3:PutObject` y `s3:PutObjectAcl` son opcionales y que puede utilizarlos si desea incluir un archivo de configuración para la lista de control de acceso.

```
{
    "Version": "2012-10-17",
    "Statement": [
         {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::$bucket-in-other-account/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::$bucket-in-other-account/*"
            ],
            "Effect": "Allow"
        },
        {
```

```
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": [
                "arn:aws:kendra:$your-region:$your-account-id:index/$index-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject",
                "s3:PutObject",
                "s3:PutObjectAcl"
            ],
            "Resource": "arn:aws:s3:::$bucket-in-other-account/*"
        }
    ]
}
```

Una política de compartimentos para permitir que el rol de fuente de Amazon S3 datos acceda al Amazon S3 agrupamiento en todas las cuentas. Tenga en cuenta que `s3:PutObject` y `s3:PutObjectAcl` son opcionales y que puede utilizarlos si desea incluir un archivo de configuración para la lista de control de acceso.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "$kendra-s3-connector-role-arn"
            },
            "Action": [
                "s3:GetObject",
                "s3:PutObject",
                "s3:PutObjectAcl"
            ],
            "Resource": [
                "arn:aws:s3:::$bucket-in-other-account/*"
            ]
        },
```

```
        {
            "Effect": "Allow",
            "Principal": {
                "AWS": "$kendra-s3-connector-role-arn"
            },
            "Action": "s3:ListBucket",
            "Resource": "arn:aws:s3:::$bucket-in-other-account"
        }
    ]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Amazon Kendra Web Crawler

Cuando utiliza Amazon Kendra Web Crawler, proporciona un rol con las siguientes políticas:

- Permiso para acceder al AWS Secrets Manager secreto que contiene las credenciales para conectarse a sitios web o a un servidor proxy web respaldado por una autenticación básica. Para obtener más información sobre el contenido del secreto, consulte Utilizar un origen de datos de rastreador web.
- Permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el nombre de usuario y la contraseña secreta guardados en. Secrets Manager
- Permiso para utilizar las operaciones `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.
- Si utilizas un Amazon S3 depósito para almacenar tu lista de direcciones URL iniciales o mapas de sitio, incluye el permiso para acceder al depósito. Amazon S3

> **Note**
>
> Puede conectar una fuente de datos de Amazon Kendra Web Crawler a través de. Amazon Kendra Amazon VPC Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
```

```
}
```

Si guardas las URL iniciales o los mapas de sitio en un Amazon S3 depósito, debes añadir este permiso al rol.

```
,
{"Effect": "Allow",
     "Action": [
         "s3:GetObject"
      ],
      "Resource": [
         "arn:aws:s3:::bucket-name/*"
      ]
}
```

Una política de confianza que permita Amazon Kendra asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes Amazon WorkDocs de datos

Cuando los usa Amazon WorkDocs, proporciona un rol con las siguientes políticas

- Permiso para verificar el ID de directorio (ID de organización) que corresponde al repositorio del sitio de Amazon WorkDocs .
- Permiso para obtener el nombre de dominio de Active Directory que contiene el directorio del sitio de Amazon WorkDocs .
- Permiso para llamar a las API públicas necesarias para el conector de Amazon WorkDocs .

- Permiso para llamar a las API `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowsKendraToGetDomainNameOfActiveDirectory",
      "Effect": "Allow",
      "Action": "ds:DescribeDirectories",
      "Resource": "*"
    },
    {
      "Sid": "AllowsKendraToCallRequiredWorkDocsAPIs",
      "Effect": "Allow",
      "Action": [
        "workdocs:GetDocumentPath",
        "workdocs:GetGroup",
        "workdocs:GetDocument",
        "workdocs:DownloadDocumentVersions",
        "workdocs:DescribeUsers",
        "workdocs:DescribeFolderContents",
        "workdocs:DescribeActivities",
        "workdocs:DescribeComments",
        "workdocs:GetFolder",
        "workdocs:DescribeResourcePermissions",
        "workdocs:GetFolderPath",
        "workdocs:DescribeInstances"
      ],
      "Resource": "*"
    },
    {
      "Sid": "iamPassRole",
      "Effect": "Allow",
      "Action": "iam:PassRole",
      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "iam:PassedToService": [
            "kendra.amazonaws.com"
          ]
        }
```

```
      }
    },
    {
      "Sid": "AllowsKendraToCallBatchPutDeleteAPIs",
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
        "kendra:BatchDeleteDocument"
      ],
      "Resource": [
        "arn:aws:kendra:your-region:account-id:index/$index-id"
      ]
    }
  ]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Box

Cuando se utiliza Box, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a tu AWS Secrets Manager secreto para autenticar tu Slack.
- Permiso para llamar a las API públicas necesarias para el conector de Box.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puedes conectar una fuente de datos de Box a Amazon Kendra través de. Amazon VPC Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
```

```
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-d}}:index/{{index-
id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Confluence

IAM funciones para Confluence Connector v1.0

Cuando se utiliza un servidor de Confluence como origen de datos, se proporciona un rol con las siguientes políticas:

- Permiso para acceder al AWS Secrets Manager secreto que contiene las credenciales necesarias para conectarse a Confluence. Para obtener más información acerca del contenido del secreto, consulte orígenes de datos de Confluence.
- Permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el nombre de usuario y la contraseña secreta guardada por. Secrets Manager

- Permiso para utilizar las operaciones `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.

> **Note**
>
> Puedes conectar una fuente de datos de Confluence a través de. Amazon Kendra Amazon VPC Si utilizas un Amazon VPC, necesitas añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
```

```
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Si utiliza una VPC, proporcione una política que dé Amazon Kendra acceso a los recursos necesarios. Consulte Roles de IAM para los orígenes de datos y VPC para ver la política requerida.

Una política de confianza que permita Amazon Kendra asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para Confluence Connector v2.0

Para un origen de datos de Confluence Connector v2.0, se proporciona un rol con las siguientes políticas.

- Permiso para acceder al AWS Secrets Manager secreto que contiene las credenciales de autenticación de Confluence. Para obtener más información acerca del contenido del secreto, consulte orígenes de datos de Confluence.
- Permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el nombre de usuario y la contraseña secreta que guarda. AWS Secrets Manager
- Permiso para utilizar las operaciones `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.

También debe adjuntar una política de confianza que permita Amazon Kendra asumir el rol.

> **Note**
>
> Puedes conectar una fuente de datos de Confluence a Amazon Kendra través Amazon VPC de. Si utilizas un Amazon VPC, necesitas añadir permisos adicionales.

Una política de roles que te permita conectarte Amazon Kendra a Confluence.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "secretsmanager:GetSecretValue"
      ],
      "Resource": [
        "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt"
      ],
      "Resource": [
        "arn:aws:kms:your-region:your-account-id:key/key-id"
      ],
      "Condition": {
        "StringLike": {
          "kms:ViaService": [
            "secretsmanager.your-region.amazonaws.com"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
```

```
      ],
      "Resource": [
        "arn:aws:kendra:your-region:your-account-id:index/index-id",
        "arn:aws:kendra:your-region:your-account-id:index/index-id/data-source/*"
      ]
    }
    {
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
        "kendra:BatchDeleteDocument"
      ],
      "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
    }
  ]
}
```

Una política de confianza que permite Amazon Kendra asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Dropbox

Cuando se utiliza Dropbox, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a tu AWS Secrets Manager secreto para autenticar tu Dropbox.
- Permiso para llamar a las API públicas necesarias para el conector de Dropbox.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puedes conectar una fuente de datos de Dropbox a Amazon Kendra través Amazon VPC de. Si utilizas un Amazon VPC, tendrás que añadir permisos adicionales.

```
{
"Version": "2012-10-17",
  "Statement": [
  {"Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {"Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {"StringLike": {"kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {"Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"]
  },
  {"Effect": "Allow",
```

```
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Drupal

Cuando se utiliza Drupal, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a tu AWS Secrets Manager secreto para autenticar tu Drupal.
- Permiso para llamar a las API públicas necesarias para el conector de Drupal.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puedes conectar una fuente de datos de Drupal a través de. Amazon Kendra Amazon VPC Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
```

```
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes GitHub de datos

Cuando los usa GitHub, proporciona un rol con las siguientes políticas.

- Permiso para acceder a tu AWS Secrets Manager secreto para autenticar tu GitHub.
- Permiso para llamar a las API públicas necesarias para el GitHub conector.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente GitHub de datos a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
```

```
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
```

```
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Gmail

Cuando se utiliza Gmail, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a tu AWS Secrets Manager secreto para autenticar tu cuenta de Gmail.
- Permiso para llamar a las API públicas necesarias para el conector de Gmail.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puedes conectar una fuente de datos de Gmail a Amazon Kendra través Amazon VPC de. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
"Version": "2012-10-17",
  "Statement": [
  {"Effect": "Allow",
```

```
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {"Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {"StringLike": {"kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {"Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"]
  },
  {"Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
```

```
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Google Drive

Cuando utilizas una fuente de datos de Google Workspace Drive, Amazon Kendra asignas un rol que tiene los permisos necesarios para conectarse al sitio. Entre ellos se incluyen:

- Permiso para obtener y descifrar el AWS Secrets Manager secreto que contiene el correo electrónico de la cuenta de cliente, el correo de la cuenta de administrador y la clave privada necesarios para conectarse al sitio de Google Drive. Para obtener más información acerca del contenido de este secreto, consulte orígenes de datos de Google Drive.
- Permiso para usar las BatchDeleteDocumentAPI BatchPutDocumenty.

> **Note**
>
> Puedes conectar una fuente de datos de Google Drive a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

La siguiente IAM política proporciona los permisos necesarios:

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
```

```
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
```

```
}
```

IAM funciones para las fuentes de datos de IBM DB2

Cuando se utiliza un conector de origen de datos de IBM DB2, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su instancia de fuente de datos de IBM DB2.
- Permiso para llamar a las API públicas necesarias para el conector de origen de datos de IBM DB2.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos de IBM DB2 a través de. Amazon Kendra Amazon VPC Si utiliza un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
```

```
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
```

```
    ]
}
```

IAM funciones para las fuentes de datos de Jira

Cuando se utiliza Jira, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a tu AWS Secrets Manager secreto para autenticar tu Jira.
- Permiso para llamar a las API públicas necesarias para el conector de Jira.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puedes conectar una fuente de datos de Jira a través de. Amazon Kendra Amazon VPC Si utilizas un Amazon VPC, tendrás que añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
```

```
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
```

```
}
```

IAM roles para fuentes de datos de Microsoft Exchange

Cuando utiliza una fuente de datos de Microsoft Exchange, Amazon Kendra proporciona un rol que tiene los permisos necesarios para conectarse al sitio. Entre ellos se incluyen:

- Permiso para obtener y descifrar el AWS Secrets Manager secreto que contiene el identificador de aplicación y la clave secreta necesarios para conectarse al sitio de Microsoft Exchange. Para obtener más información acerca del contenido del secreto, consulte orígenes de datos de Microsoft Exchange.
- Permiso para usar las BatchDeleteDocumentAPI BatchPutDocumenty.

> **Note**
>
> Puede conectar una fuente de datos de Microsoft Exchange a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

La siguiente IAM política proporciona los permisos necesarios:

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
```

```
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Si va a almacenar la lista de usuarios para indexarlos en un Amazon S3 bucket, también debe conceder permiso para utilizar la `GetObject` operación S3. La siguiente política de IAM proporciona los permisos necesarios:

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Action": [
      "s3:GetObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name/*"
    ],
    "Effect": "Allow"
```

```
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/[[key-ids]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com",
          "s3.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Una política de confianza que Amazon Kendra permita asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para fuentes de OneDrive datos de Microsoft

Cuando utiliza una fuente de OneDrive datos de Microsoft, Amazon Kendra proporciona un rol que tiene los permisos necesarios para conectarse al sitio. Entre ellos se incluyen:

- Permiso para obtener y descifrar el AWS Secrets Manager secreto que contiene el identificador de la aplicación y la clave secreta necesarios para conectarse al OneDrive sitio. Para obtener más información sobre el contenido del secreto, consulte Fuentes de OneDrive datos de Microsoft.
- Permiso para usar las BatchDeleteDocumentAPI BatchPutDocumenty.

> **Note**
>
> Puede conectar una fuente de OneDrive datos de Microsoft a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

La siguiente IAM política proporciona los permisos necesarios:

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
```

```
            "secretsmanager.your-region.amazonaws.com"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
        "kendra:BatchDeleteDocument"
      ],
      "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
    }]
}
```

Si va a almacenar la lista de usuarios para indexarlos en un Amazon S3 bucket, también debe conceder permiso para utilizar la `GetObject` operación S3. La siguiente política de IAM proporciona los permisos necesarios:

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Action": [
      "s3:GetObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name/*"
    ],
    "Effect": "Allow"
  },
  {
    "Effect": "Allow",
    "Action": [
```

```
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/[[key-ids]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com",
          "s3.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Una política de confianza que Amazon Kendra permita asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para fuentes de SharePoint datos de Microsoft

IAM funciones para SharePoint Connector v1.0

Para una fuente de datos de Microsoft SharePoint Connector v1.0, debe proporcionar un rol con las siguientes políticas.

- Permiso para acceder al AWS Secrets Manager secreto que contiene el nombre de usuario y la contraseña del SharePoint sitio. Para obtener más información sobre el contenido del secreto, consulte Fuentes de SharePoint datos de Microsoft.
- Permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el nombre de usuario y la contraseña secretos almacenados en. AWS Secrets Manager
- Permiso para utilizar las operaciones `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.
- Permiso para acceder al Amazon S3 depósito que contiene el certificado SSL utilizado para comunicarse con el SharePoint sitio.

También debes adjuntar una política de confianza que Amazon Kendra permita asumir el rol.

> **Note**
>
> Puede conectar una fuente de SharePoint datos de Microsoft a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:GetSecretValue"
            ],
            "Resource": [
                "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
            ]
        },
        {
            "Effect": "Allow",
```

```
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": [
                "arn:aws:kendra:your-region:your-account-id:index/index-id"
            ],
            "Condition": {
                "StringLike": {
                    "kms:ViaService": [
                        "kendra.your-region.amazonaws.com"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Si has cifrado el Amazon S3 depósito que contiene el certificado SSL utilizado para comunicarte con el SharePoint sitio, proporciona una política para dar Amazon Kendra acceso a la clave.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
```

```
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

Una política de confianza que Amazon Kendra permita asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para SharePoint Connector v2.0

Para una fuente de datos de Microsoft SharePoint Connector v2.0, debe proporcionar un rol con las siguientes políticas.

- Permiso para acceder al AWS Secrets Manager secreto que contiene las credenciales de autenticación del SharePoint sitio. Para obtener más información sobre el contenido del secreto, consulte Fuentes de SharePoint datos de Microsoft.
- Permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el nombre de usuario y la contraseña secretos almacenados en. AWS Secrets Manager
- Permiso para utilizar las operaciones `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.
- Permiso para acceder al Amazon S3 depósito que contiene el certificado SSL utilizado para comunicarse con el SharePoint sitio.

También debes adjuntar una política de confianza que Amazon Kendra permita asumir el rol.

> **Note**
>
> Puede conectar una fuente de SharePoint datos de Microsoft a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "secretsmanager:GetSecretValue"
      ],
      "Resource": [
        "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt"
      ],
      "Resource": [
        "arn:aws:kms:your-region:your-account-id:key/key-id"
      ],
      "Condition": {
        "StringLike": {
          "kms:ViaService": [
            "secretsmanager.your-region.amazonaws.com"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
```

```
        "kendra:DescribePrincipalMapping"
      ],
      "Resource": [
        "arn:aws:kendra:your-region:your-account-id:index/index-id",
        "arn:aws:kendra:your-region:your-account-id:index/index-id/data-source/*"
      ]
    },
    {
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name/key-name"
      ],
      "Effect": "Allow"
    },
    {
      "Effect": "Allow",
      "Action": [
        "kendra:BatchPutDocument",
        "kendra:BatchDeleteDocument"
      ],
      "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": [
        "arn:aws:ec2:your-region:your-account-id:subnet/subnet-ids",
        "arn:aws:ec2:your-region:your-account-id:security-group/security-group"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": "arn:aws:ec2:region:account_id:network-interface/*",
      "Condition": {
        "StringLike": {
          "aws:RequestTag/AWS_KENDRA": "kendra_your-account-id_index-id_*"
        }
```

```
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:your-region:your-account-id:network-interface/*",
      "Condition": {
        "StringEquals": {
          "ec2:CreateAction": "CreateNetworkInterface"
        }
      }
    },

{
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterfacePermission"
      ],
      "Resource": "arn:aws:ec2:your-region:your-account-id:network-interface/*",
      "Condition": {
        "StringLike": {
          "aws:ResourceTag/AWS_KENDRA": "kendra_your-account-id_index-id_*"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeNetworkInterfaces",
        "ec2:DescribeAvailabilityZones",
        "ec2:DescribeNetworkInterfaceAttribute",
        "ec2:DescribeVpcs",
        "ec2:DescribeRegions",
        "ec2:DescribeNetworkInterfacePermissions",
        "ec2:DescribeSubnets"
      ],
      "Resource": "*"
    }
  ]
}
```

Si has cifrado el Amazon S3 depósito que contiene el certificado SSL utilizado para comunicarte con el SharePoint sitio, proporciona una política para dar Amazon Kendra acceso a la clave.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:youraccount-id:key/key-id"
            ]
        }
    ]
}
```

Una política de confianza que Amazon Kendra permita asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para fuentes de datos de Microsoft SQL Server

Cuando se utiliza Microsoft SQL Server, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su instancia de Microsoft SQL Server.
- Permiso para llamar a las API públicas necesarias para el conector de Microsoft SQL Server.

- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos de Microsoft SQL Server a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
```

```
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
 "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para fuentes de datos de Microsoft Teams

Cuando utiliza una fuente de datos de Microsoft Teams, Amazon Kendra proporciona un rol que tiene los permisos necesarios para conectarse al sitio. Entre ellos se incluyen:

- Permiso para obtener y descifrar el AWS Secrets Manager secreto que contiene el ID de cliente y el secreto de cliente necesarios para conectarse a Microsoft Teams. Para obtener más información acerca del contenido del secreto, consulte orígenes de datos de Microsoft Teams.

> **Note**
>
> Puede conectar una fuente de datos de Microsoft Teams a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

La siguiente IAM política proporciona los permisos necesarios:

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:client-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
```

```
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para fuentes de datos de Microsoft Yammer

Cuando utiliza una fuente de datos de Microsoft Yammer, Amazon Kendra proporciona un rol que tiene los permisos necesarios para conectarse al sitio. Entre ellos se incluyen:

- Permiso para obtener y descifrar el AWS Secrets Manager secreto que contiene el identificador de aplicación y la clave secreta necesarios para conectarse al sitio de Microsoft Yammer. Para obtener más información acerca del contenido del secreto, consulte orígenes de datos de Microsoft Yammer.
- Permiso para usar las API BatchPutDocumenty. BatchDeleteDocument

> **Note**
>
> Puede conectar una fuente de datos de Microsoft Yammer a Amazon Kendra través Amazon VPC de. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

La siguiente IAM política proporciona los permisos necesarios:

```
{
```

```
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Si va a almacenar la lista de usuarios para indexarlos en un Amazon S3 bucket, también debe conceder permiso para utilizar la `GetObject` operación S3. La siguiente política de IAM proporciona los permisos necesarios:

```
{
  "Version": "2012-10-17",
```

```
"Statement": [
{
  "Effect": "Allow",
  "Action": [
    "secretsmanager:GetSecretValue"
  ],
  "Resource": [
    "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
  ]
},
{
  "Action": [
    "s3:GetObject"
  ],
  "Resource": [
    "arn:aws:s3:::bucket-name/*"
  ],
  "Effect": "Allow"
},
{
  "Effect": "Allow",
  "Action": [
    "kms:Decrypt"
  ],
  "Resource": [
    "arn:aws:kms:your-region:your-account-id:key/[[key-ids]]"
  ],
  "Condition": {
    "StringLike": {
      "kms:ViaService": [
        "secretsmanager.your-region.amazonaws.com",
        "s3.your-region.amazonaws.com"
      ]
    }
  }
},
{
  "Effect": "Allow",
  "Action": [
    "kendra:BatchPutDocument",
    "kendra:BatchDeleteDocument"
  ],
  "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
}]
```

```
}
```

Una política de confianza que Amazon Kendra permita asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para fuentes de datos de MySQL

Cuando se utiliza un conector de origen de datos de MySQL, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su instancia de fuente de datos de My SQL.
- Permiso para llamar a las API públicas necesarias para el conector de origen de datos de MySQL.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos MySQL a Amazon Kendra través de Amazon VPC. Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
```

```
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
"arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
```

```
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Oracle

Cuando se utiliza un conector de origen de datos de Oracle, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su instancia de fuente de datos de Oracle.
- Permiso para llamar a las API públicas necesarias para el conector de origen de datos de Oracle.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos de Oracle a Amazon Kendra través Amazon VPC de. Si está utilizando un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
```

```
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
"arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
  }]
```

```
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para fuentes de datos de PostgreSQL

Cuando se utiliza un conector de origen de datos de PostgreSQL, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su instancia de fuente de datos de PostgreSQL.
- Permiso para llamar a las API públicas necesarias para el conector de origen de datos de PostgreSQL.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos de PostgreSQL a través de. Amazon Kendra Amazon VPC Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
```

```
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{region}}:{{account_id}}:secret:[[secret_id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{region}}:{{account_id}}:key/[[key_id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.*.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}",
"arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{region}}:{{account_id}}:index/{{index_id}}"
```

```
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Quip

Cuando se utiliza Quip, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su Quip.
- Permiso para llamar a las API públicas necesarias para el conector de Quip.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos de Quip a través de. Amazon Kendra Amazon VPC Si utiliza un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
```

```
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{yoour-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{your-
index-id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{your-index-
id}}/data-source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
```

```
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Salesforce

Cuando se utiliza Salesforce como origen de datos, se proporciona un rol con las siguientes políticas:

- Permiso para acceder al AWS Secrets Manager secreto que contiene el nombre de usuario y la contraseña del sitio de Salesforce. Para obtener más información acerca del contenido del secreto, consulte orígenes de datos de Salesforce.
- Permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el nombre de usuario y la contraseña secretos almacenados en. Secrets Manager
- Permiso para utilizar las operaciones `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.

> **Note**
>
> Puede conectar una fuente de datos de Salesforce a través de. Amazon Kendra Amazon VPC Si está utilizando un Amazon VPC, debe añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
```

```
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:account-id:index/index-id"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
```

```
        },
        "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes ServiceNow de datos

Cuando utiliza un ServiceNow como fuente de datos, proporciona un rol con las siguientes políticas:

- Permiso para acceder al Secrets Manager secreto que contiene el nombre de usuario y la contraseña del ServiceNow sitio. Para obtener más información sobre el contenido del secreto, consulte ServiceNow orígenes de datos.
- Permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el nombre de usuario y la contraseña secretos almacenados en. Secrets Manager
- Permiso para utilizar las operaciones `BatchPutDocument` y `BatchDeleteDocument` para actualizar el índice.

> **Note**
>
> Puede conectar una fuente de ServiceNow datos a Amazon Kendra través de. Amazon VPC Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:your-region:your-account-id:secret:secret-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
```

```
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.your-region.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:your-region:your-account-id:index/index-id"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Slack

Cuando se utiliza Slack, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a tu AWS Secrets Manager secreto para autenticar tu Slack.

- Permiso para llamar a las API públicas necesarias para el conector de Slack.
- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puedes conectar una fuente de datos de Slack a través de. Amazon Kendra Amazon VPC Si utilizas un Amazon VPC, debes añadir permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
```

```
    "Action": [
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las fuentes de datos de Zendesk

Cuando se utiliza Zendesk, se proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar su Zendesk Suite.
- Permiso para llamar a las API públicas necesarias para el conector de Zendesk.

- Permiso para llamar a las API `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` y `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Puede conectar una fuente de datos de Zendesk a través de. Amazon Kendra Amazon VPC Si está utilizando un Amazon VPC, debe agregar permisos adicionales.

```
{
  "Version": "2012-10-17",
  "Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue"
    ],
    "Resource": [
      "arn:aws:secretsmanager:{{your-region}}:{{your-account-id}}:secret:[[secret-id]]"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt"
    ],
    "Resource": [
      "arn:aws:kms:{{your-region}}:{{your-account-id}}:key/[[key-id]]"
    ],
    "Condition": {
      "StringLike": {
        "kms:ViaService": [
          "secretsmanager.{{your-region}}.amazonaws.com"
        ]
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
```

```
        "kendra:PutPrincipalMapping",
        "kendra:DeletePrincipalMapping",
        "kendra:ListGroupsOlderThanOrderingId",
        "kendra:DescribePrincipalMapping"
    ],
    "Resource": ["arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-
id}}", "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}/data-
source/*"]
  },
  {
    "Effect": "Allow",
    "Action": [
      "kendra:BatchPutDocument",
      "kendra:BatchDeleteDocument"
    ],
    "Resource": "arn:aws:kendra:{{your-region}}:{{your-account-id}}:index/{{index-id}}"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Función de nube privada virtual (VPC) IAM

Si utiliza una nube privada virtual (VPC) para conectarse a la fuente de datos, debe proporcionar los siguientes permisos adicionales.

Función de VPC IAM

```
{
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": [
        "arn:aws:ec2:{{region}}:{{account_id}}:subnet/[[subnet_ids]]",
        "arn:aws:ec2:{{region}}:{{account_id}}:security-group/[[security_group]]"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface"
      ],
      "Resource": "arn:aws:ec2:{{region}}:{{account_id}}:network-interface/*",
      "Condition": {
        "StringLike": {
          "aws:RequestTag/AWS_KENDRA": "kendra_{{account_id}}_{{index_id}}_*"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:{{region}}:{{account_id}}:network-interface/*",
      "Condition": {
        "StringEquals": {
          "ec2:CreateAction": "CreateNetworkInterface"
        }
      }
    },

{
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterfacePermission"
      ],
      "Resource": "arn:aws:ec2:{{region}}:{{account_id}}:network-interface/*",
```

```
      "Condition": {
        "StringLike": {
          "aws:ResourceTag/AWS_KENDRA": "kendra_{{account_id}}_{{index_id}}_*"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeNetworkInterfaces",
        "ec2:DescribeAvailabilityZones",
        "ec2:DescribeNetworkInterfaceAttribute",
        "ec2:DescribeVpcs",
        "ec2:DescribeRegions",
        "ec2:DescribeNetworkInterfacePermissions",
        "ec2:DescribeSubnets"
      ],
      "Resource": "*"
    }
}
```

Una política de confianza que permite Amazon Kendra asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para las preguntas frecuentes (FAQ)

Cuando utilizas la CreateFaqAPI para cargar preguntas y respuestas en un índice, debes proporcionar Amazon Kendra un IAM rol con acceso al Amazon S3 depósito que contiene los

archivos de origen. Si los archivos de origen están cifrados, debe dar permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar los archivos.

IAM funciones para las preguntas frecuentes

Una política de roles obligatoria para permitir Amazon Kendra el acceso a un Amazon S3 bucket.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Una política de funciones opcional que permite Amazon Kendra utilizar una clave maestra AWS KMS del cliente (CMK) para descifrar los archivos de un Amazon S3 depósito.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ],
            "Condition": {
                "StringLike": {
                    "kms:ViaService": [
                        "kendra.your-region.amazonaws.com"
                    ]
                }
            }
```

```
        }
    ]
}
```

Una política de confianza que permite Amazon Kendra asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para sugerencias de consultas

Cuando utilizas un Amazon S3 archivo como lista de sugerencias de consultas bloqueadas, proporcionas un rol que tenga permiso para acceder al Amazon S3 archivo y al Amazon S3 bucket. Si el archivo de texto de la lista de bloqueados (el Amazon S3 archivo) del Amazon S3 depósito está cifrado, debe dar permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar los documentos.

IAM funciones para sugerencias de consultas

Una política de funciones obligatoria Amazon Kendra para poder utilizar el Amazon S3 archivo como lista de sugerencias de consulta bloqueadas.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {"Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
```

```
            ]
        }
    ]
}
```

Una política de funciones opcional que permite Amazon Kendra utilizar una clave maestra AWS KMS del cliente (CMK) para descifrar los documentos de un Amazon S3 depósito.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {"Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

Una política de confianza que permite Amazon Kendra asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM funciones para el mapeo principal de usuarios y grupos

Al utilizar la PutPrincipalMappingAPI para asignar usuarios a sus grupos y filtrar los resultados de búsqueda por contexto de usuario, es necesario proporcionar una lista de los usuarios o subgrupos

que pertenecen a un grupo. Si tu lista contiene más de 1000 usuarios o subgrupos para un grupo, debes proporcionar un rol que tenga permiso para acceder al Amazon S3 archivo de la lista y al Amazon S3 bucket. Si el archivo de texto (el Amazon S3 archivo) de la lista del Amazon S3 depósito está cifrado, debes dar permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar los documentos.

IAM funciones para el mapeo principal

Una política de roles obligatoria Amazon Kendra para poder usar el Amazon S3 archivo como lista de usuarios y subgrupos que pertenecen a un grupo.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {"Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket-name/*"
            ]
        }
    ]
}
```

Una política de funciones opcional que permite Amazon Kendra utilizar una clave maestra AWS KMS del cliente (CMK) para descifrar los documentos de un Amazon S3 depósito.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {"Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:your-region:your-account-id:key/key-id"
            ]
        }
    ]
}
```

Una política de confianza que permite Amazon Kendra asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Se recomienda incluir `aws:sourceAccount` y `aws:sourceArn` en la política de confianza. Esto limita los permisos y comprueba de forma segura si `aws:sourceAccount` los mismos que se indican en la política de IAM roles de la `sts:AssumeRole` acción. `aws:sourceArn` Esto evita que entidades no autorizadas accedan a tus IAM funciones y a sus permisos. Para obtener más información, consulte la AWS Identity and Access Management guía sobre el problema de los diputados confusos.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "kendra.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "your-account-id"
                },
                "StringLike": {
                    "aws:SourceArn": "arn:aws:kendra:your-region:your-account-id:index-id/*"
                }
            }
```

```
        }
    ]
}
```

IAM funciones para AWS IAM Identity Center

Cuando usas el UserGroupResolutionConfigurationobjeto para obtener los niveles de acceso de grupos y usuarios de una fuente de AWS IAM Identity Center identidad, debes proporcionar un rol que tenga permiso de acceso IAM Identity Center.

IAM roles para AWS IAM Identity Center

Una política de roles obligatoria para Amazon Kendra permitir el acceso IAM Identity Center.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "sso-directory:SearchUsers",
                "sso-directory:ListGroupsForUser",
                "sso-directory:DescribeGroups",
                "sso:ListDirectoryAssociations"
            ],
            "Resource": [
                 "*"
            ]
        },
        {
          "Sid": "iamPassRole",
          "Effect": "Allow",
          "Action": "iam:PassRole",
          "Resource": "*",
          "Condition": {
            "StringEquals": {
              "iam:PassedToService": [
                "kendra.amazonaws.com"
              ]
            }
          }
        }
```

```
    ]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

IAM roles para Amazon Kendra experiencias

Cuando utilice UpdateExperiencelas API CreateExperienceo las API para crear o actualizar una aplicación de búsqueda, debe proporcionar un rol que tenga permiso para acceder a las operaciones necesarias y al Centro de identidad de IAM.

IAM roles para la experiencia Amazon Kendra de búsqueda

Una política de funciones obligatoria para poder acceder Amazon Kendra a `Query` las operaciones, `QuerySuggestions` `SubmitFeedback` las operaciones y el centro de identidad de IAM, que almacena la información de sus usuarios y grupos.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowsKendraSearchAppToCallKendraApi",
      "Effect": "Allow",
      "Action": [
        "kendra:GetQuerySuggestions",
        "kendra:Query",
        "kendra:DescribeIndex",
        "kendra:ListFaqs",
```

```
        "kendra:DescribeDataSource",
        "kendra:ListDataSources",
        "kendra:DescribeFaq",
        "kendra:SubmitFeedback"
      ],
      "Resource": [
        "arn:aws:kendra:your-region:your-account-id:index/index-id"
      ]
    },
    {
      "Sid": "AllowKendraSearchAppToDescribeDataSourcesAndFaq",
      "Effect": "Allow",
      "Action": [
        "kendra:DescribeDataSource",
        "kendra:DescribeFaq"
      ],
      "Resource": [
        "arn:aws:kendra:your-region:your-account-id:index/index-id/data-source/data-
source-id",
        "arn:aws:kendra:your-region:your-account-id:index/index-id/faq/faq-id"
      ]
    },
    {
      "Sid": "AllowKendraSearchAppToCallSSODescribeUsersAndGroups",
      "Effect": "Allow",
      "Action": [
        "sso-directory:ListGroupsForUser",
        "sso-directory:SearchGroups",
        "sso-directory:SearchUsers",
        "sso-directory:DescribeUser",
        "sso-directory:DescribeGroup",
        "sso-directory:DescribeGroups",
        "sso-directory:DescribeUsers",
        "sso:ListDirectoryAssociations"
      ],
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringLike": {
          "kms:ViaService": [
            "kendra.your-region.amazonaws.com"
          ]
        }
```

```
      }
    }
  ]
}
```

Una política de confianza que permite Amazon Kendra asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Se recomienda incluir `aws:sourceAccount` y `aws:sourceArn` en la política de confianza. Esto limita los permisos y comprueba de forma segura si `aws:sourceAccount` los mismos que se indican en la política de IAM roles de la `sts:AssumeRole` acción. `aws:sourceArn` Esto evita que entidades no autorizadas accedan a tus IAM funciones y a sus permisos. Para obtener más información, consulte la AWS Identity and Access Management guía sobre el problema de los diputados confusos.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "kendra.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "your-account-id"
```

```
                },
                "StringLike": {
                    "aws:SourceArn": "arn:aws:kendra:your-region:your-account-id:index-
id/*"
                }
            }
        }
    ]
}
```

IAM funciones para el enriquecimiento de documentos personalizados

Cuando utilice el CustomDocumentEnrichmentConfigurationobjeto para realizar modificaciones avanzadas en los metadatos y el contenido del documento, debe proporcionar un rol que tenga los permisos necesarios para ejecutar `PreExtractionHookConfiguration` y/o`PostExtractionHookConfiguration`. Se configura una función de Lambda para `PreExtractionHookConfiguration` o `PostExtractionHookConfiguration` para aplicar modificaciones avanzadas de los metadatos y el contenido del documento durante el proceso de ingesta. Si decide activar el cifrado del lado del servidor para su Amazon S3 depósito, debe conceder permiso para utilizar la clave maestra del AWS KMS cliente (CMK) a fin de cifrar y descifrar los objetos almacenados en el depósito. Amazon S3

IAM funciones para el enriquecimiento personalizado de documentos

Una política de roles obligatoria para Amazon Kendra permitir la ejecución `PreExtractionHookConfiguration` y `PostExtractionHookConfiguration` el cifrado de su Amazon S3 bucket.

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Action": [
      "s3:GetObject",
      "s3:PutObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name/*"
    ],
    "Effect": "Allow"
```

```
  },
  {
    "Action": [
      "s3:ListBucket"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name"
    ],
    "Effect": "Allow"
  },
  {
    "Effect": "Allow",
    "Action": [
      "kms:Decrypt",
      "kms:GenerateDataKey"
    ],
    "Resource": [
      "arn:aws:kms:your-region:your-account-id:key/key-id"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "lambda:InvokeFunction"
    ],
    "Resource": "arn:aws:lambda:your-region:your-account-id:function:lambda-function"
  }]
}
```

Una política de roles opcional Amazon Kendra que permite ejecutar `PreExtractionHookConfiguration` el Amazon S3 bucket `PostExtractionHookConfiguration` sin cifrar.

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Action": [
      "s3:GetObject",
      "s3:PutObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name/*"
    ],
```

```
    "Effect": "Allow"
  },
  {
    "Action": [
      "s3:ListBucket"
    ],
    "Resource": [
      "arn:aws:s3:::bucket-name"
    ],
    "Effect": "Allow"
  },
  {
    "Effect": "Allow",
    "Action": [
      "lambda:InvokeFunction"
    ],
    "Resource": "arn:aws:lambda:your-region:your-account-id:function:lambda-function"
  }]
}
```

Una política de confianza que Amazon Kendra permite asumir un rol.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Principal":{
            "Service":"kendra.amazonaws.com"
         },
         "Action":"sts:AssumeRole"
      }
   ]
}
```

Se recomienda incluir `aws:sourceAccount` y `aws:sourceArn` en la política de confianza. Esto limita los permisos y comprueba de forma segura si `aws:sourceAccount` los mismos que se indican en la política de IAM roles de la `sts:AssumeRole` acción. `aws:sourceArn` Esto evita que entidades no autorizadas accedan a tus IAM funciones y a sus permisos. Para obtener más información, consulte la AWS Identity and Access Management guía sobre el problema de los diputados confusos.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "kendra.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "your-account-id"
                },
                "StringLike": {
                    "aws:SourceArn": "arn:aws:kendra:your-region:your-account-id:index-
id/*"
                }
            }
        }
    ]
}
```

Implementación de Amazon Kendra

Cuando llegue el momento de implementar la búsqueda de Amazon Kendra en su sitio web, le proporcionaremos el código fuente que podrá usar con React para empezar con su aplicación. El código fuente se proporciona sin cargo alguno en virtud de una licencia del MIT modificada. Puede usarlo tal cual o cambiarlo según sus propias necesidades. La aplicación React proporcionada es un ejemplo para ayudarlo a comenzar. No es una aplicación lista para producción.

Para implementar una aplicación de búsqueda sin código y generar una URL de punto de conexión para su página de búsqueda con control de acceso, consulte Amazon Kendra Experience Builder.

El siguiente código de ejemplo añade la búsqueda de Amazon Kendra a una aplicación web de React existente:

- https://kendrasamples.s3.amazonaws.com/kendrasamples-react-app.zip: archivos de muestra que los desarrolladores pueden usar para crear una experiencia de búsqueda funcional en su aplicación web de React existente.

Los ejemplos se basan en la página de búsqueda de la consola de Amazon Kendra. Tienen las mismas características para buscar y mostrar los resultados de búsqueda. Puede usar el ejemplo completo o elegir solo una de las características para su propio uso.

Para ver los tres componentes de la página de búsqueda en la consola de Amazon Kendra, seleccione el icono de código (</>) en el menú de la derecha. Sitúe el puntero del ratón sobre cada sección para ver una breve descripción del componente y obtener la URL de la fuente del componente.

Temas

- Información general
- Requisitos previos
- Configurar el ejemplo
- Página de búsqueda principal
- Componente de búsqueda
- Componente de resultados
- Componente de facetas

- Componente de paginación
- Creación de una experiencia de búsqueda sin código

Información general

Para activar la búsqueda, agrega el código de ejemplo a una aplicación web de React existente. El código de ejemplo incluye un archivo Readme con los pasos para configurar un nuevo entorno de desarrollo de React. Los datos de ejemplo del código de ejemplo se pueden usar para demostrar una búsqueda. Los archivos y componentes de búsqueda del código de ejemplo se estructuran de la siguiente manera:

- Página de búsqueda principal (`Search.tsx`): es la página principal que contiene todos los componentes. Aquí es donde se integra la aplicación con la API de Amazon Kendra.
- Barra de búsqueda: es el componente en el que el usuario introduce un término de búsqueda y llama a la función de búsqueda.
- Resultados: este es el componente que muestra los resultados de Amazon Kendra. Tiene tres componentes: respuestas sugeridas, resultados de preguntas frecuentes y documentos recomendados.
- Facetas: este es el componente que muestra las facetas en los resultados de la búsqueda y permite elegir una faceta para restringir la búsqueda.
- Paginación: este es el componente desde el que se pagina la respuesta de Amazon Kendra.

Requisitos previos

Antes de comenzar, necesitará lo siguiente:

- Node.js y npm instalados. Se requiere la versión 19 o anterior de Node.js.
- Python 3 o Python 2 descargados e instalados.
- SDK for Java o AWS SDK for JavaScript para realizar llamadas a la API a Amazon Kendra.
- Una aplicación web de React existente. El código de ejemplo incluye un archivo Readme con los pasos sobre cómo configurar un nuevo entorno de desarrollo de React, incluido utilizar los marcos/bibliotecas necesarios. También puede seguir las instrucciones de inicio rápido de la documentación de React para crear una aplicación web de React.
- Las bibliotecas y dependencias necesarias configuradas en su entorno de desarrollo. El código de ejemplo incluye un archivo Readme que enumera las bibliotecas y las dependencias de paquetes

necesarias. Tenga en cuenta que `sass` es obligatorio, ya que `node-sass` está obsoleto. Si instaló anteriormente `node-sass`, desinstálelo e instale `sass`.

Configurar el ejemplo

Un procedimiento completo para añadir una búsqueda de Amazon Kendra a una aplicación de React se encuentra en el archivo Readme incluido en el código de ejemplo.

Para empezar a usar kendrasamples-react-app.zip

1. Asegúrese de haber completado los Requisitos previos, incluida la descarga e instalación de Node.js y npm.
2. Descargue kendrasamples-react-app.zip y descomprímalo.
3. Abra su terminal y vaya a `aws-kendra-example-react-app/src/services/`. Abra `local-dev-credentials.json` y proporcione sus credenciales. No añada este archivo a ningún repositorio público.
4. Vaya a `aws-kendra-example-react-app` e instale las dependencias en `package.json`. Ejecute `npm install`.
5. Inicie una versión de demostración de la aplicación en su servidor local. Ejecute `npm start`. Puede detener el servidor local ingresando en su teclado `Cmd/Ctrl + C`.
6. Para cambiar el puerto o el host (por ejemplo, la dirección IP), vaya a `package.json` y actualice el host y el puerto: `"start": "HOST=[host] PORT=[port] react-scripts start"`. Si utiliza Windows: `"start": "set HOST=[host] && set PORT=[port] && react-scripts start"`.
7. Si tiene un dominio de sitio web registrado, puede especificarlo en `package.json` después del nombre de su aplicación. Por ejemplo, `"homepage": "https://mywebsite.com"`. Debe volver a ejecutar `npm install` para actualizar las nuevas dependencias y, a continuación, ejecutar `npm start`.
8. Para compilar la aplicación, ejecute `npm build`. Cargue el contenido del directorio de compilación a su proveedor de alojamiento.

> **Warning**
>
> La aplicación React no está lista para producción. Es un ejemplo de cómo implementar una aplicación de búsqueda de Amazon Kendra.

Página de búsqueda principal

La página de búsqueda principal (`Search.tsx`) contiene todos los componentes de búsqueda de ejemplo. Incluye el componente de barra de búsqueda para la salida, los componentes de resultados para mostrar la respuesta de la API Query y un componente de paginación para paginar la respuesta.

Componente de búsqueda

El componente de búsqueda proporciona un cuadro de texto para introducir el texto de la consulta. La función `onSearch` es un enlace que llama a la función principal en `Search.tsx` para realizar la llamada a la API de Query de Amazon Kendra.

Componente de resultados

El componente de resultados muestra la respuesta de la API `Query`. Los resultados se muestran en tres áreas distintas.

- Respuestas sugeridas: estos son los principales resultados devueltos por la API `Query`. Contiene hasta tres respuestas sugeridas. En la respuesta, tienen el tipo de resultado `ANSWER`.
- Respuestas a preguntas frecuentes: son los resultados de las preguntas frecuentes que arroja la respuesta. Las preguntas frecuentes se añaden al índice por separado. En la respuesta, tienen el tipo `QUESTION_ANSWER`. Para obtener más información, consulte Preguntas y respuestas.
- Documentos recomendados: son documentos adicionales que Amazon Kendra devuelve en la respuesta. En la respuesta de la API `Query`, tienen el tipo `DOCUMENT`.

Los componentes de resultados comparten un conjunto de componentes para características como el resaltado, los títulos, los enlaces, etc. Los componentes compartidos deben estar presentes para que los componentes de resultados funcionen.

Componente de facetas

El componente de facetas muestra las facetas disponibles en los resultados de la búsqueda. Cada faceta clasifica la respuesta en función de una dimensión específica, como el autor. Puede restringir la búsqueda a una faceta específica seleccionando una de la lista.

Tras seleccionar una faceta, el componente llama a `Query` con un filtro de atributos que restringe la búsqueda a los documentos que coincidan con la faceta.

Componente de paginación

El componente de paginación le permite mostrar los resultados de búsqueda de la API `Query` en varias páginas. Llama a la API `Query` con los parámetros `PageSize` y `PageNumber` para obtener una página de resultados específica.

Creación de una experiencia de búsqueda sin código

Puede crear e implementar una aplicación de búsqueda de Amazon Kendra sin necesidad de ningún código de frontend. Amazon Kendra Experience Builder le ayuda a crear e implementar una aplicación de búsqueda completamente funcional con unos pocos clics para que pueda empezar a buscar de inmediato. Puede personalizar el diseño de la página de búsqueda y ajustar la búsqueda para adaptar la experiencia a las necesidades de sus usuarios. Amazon Kendra genera una URL de punto de conexión única y totalmente alojada de su página de búsqueda para empezar a buscar en sus documentos y preguntas frecuentes. Puede crear rápidamente una prueba de concepto de su experiencia de búsqueda y compartirla con otros usuarios.

Utiliza la plantilla de experiencia de búsqueda disponible en el generador para personalizar la búsqueda. Puede invitar a otras personas a colaborar en la creación de su experiencia de búsqueda o a evaluar los resultados de la búsqueda con fines de ajuste. Una vez que la experiencia de búsqueda esté lista para que los usuarios comiencen a buscar, solo tiene que compartir la URL del punto de conexión seguro.

Cómo funciona la búsqueda de Experience Builder

El proceso general de crear una experiencia de búsqueda es el siguiente:

1. Para crear su experiencia de búsqueda, debe asignarle un nombre, una descripción y elegir los orígenes de datos que desee utilizar en la experiencia de búsqueda.
2. Configura la lista de usuarios y grupos en AWS IAM Identity Center y, a continuación, les asigna derechos de acceso a su experiencia de búsqueda. Se incluye a usted mismo como propietario de la experiencia. Para obtener más información, consulte the section called “Proporcionar acceso a la página de búsqueda”.

3. Abra el Amazon Kendra Experience Builder para diseñar y ajustar su página de búsqueda. Puede compartir la URL del punto de conexión de su experiencia de búsqueda con otras personas a las que asigne derechos de acceso propios para editar o ver la búsqueda.

Llama a la API CreateExperience para crear y configurar su experiencia de búsqueda. Si usa la consola, selecciona su índice y, a continuación, selecciona Experiencias en el menú de navegación para configurar su experiencia.

Diseñe y ajuste su experiencia de búsqueda

Una vez que haya creado y configurado su experiencia de búsqueda, abra la experiencia de búsqueda mediante una URL de punto de conexión para empezar a personalizar la búsqueda como propietario con derechos de acceso de editor. Escribe la consulta en el cuadro de búsqueda y, a continuación, personaliza la búsqueda mediante las opciones de edición del panel lateral para ver cómo se aplican a su página. Cuando esté listo para publicar, seleccione Publicar. También puede cambiar entre Cambiar a la visualización en directo, para ver la última versión publicada de la página de búsqueda, y Cambiar al modo de compilación, para editar o personalizar la página de búsqueda.

Las siguientes son formas en las que puede personalizar su experiencia de búsqueda.

Filtro

Añada una búsqueda por facetas o filtre por atributos del documento. Esto incluye atributos personalizados. Puede añadir un filtro mediante sus propios campos de metadatos configurados. Por ejemplo, para buscar facetas por cada categoría de ciudad, utilice un atributo de documento personalizado `_category` que contenga todas las categorías de ciudades.

Respuesta sugerida

Agrega respuestas generadas por machine learning a las consultas de tus usuarios. Por ejemplo, “¿Qué tan difícil es este curso?”. Amazon Kendra puede recuperar el texto más relevante de todos los documentos que hacen referencia a la dificultad de un curso y sugerir la respuesta más relevante.

Preguntas frecuentes

Agregue un documento de preguntas frecuentes para proporcionar respuestas a las preguntas más frecuentes. Por ejemplo, “¿Cuántas horas se necesitan para completar este curso?”. Amazon Kendra puede utilizar el documento de preguntas frecuentes que contiene la respuesta a esta pregunta y dar la respuesta correcta.

Ordenar

Añada una clasificación de los resultados de la búsqueda para que los usuarios puedan organizarlos por relevancia, hora de creación, hora de la última actualización y otros criterios de clasificación.

Documentos

Configure cómo se muestran los documentos o los resultados de la búsqueda en la página de búsqueda. Puede configurar el número de resultados que se muestran en la página, incluir la paginación (por ejemplo, los números de página), activar un botón de comentarios de los usuarios y organizar la forma en que se muestran los campos de metadatos de los documentos en los resultados de la búsqueda.

Idioma

Seleccione un idioma para filtrar los resultados de la búsqueda o los documentos en el idioma seleccionado.

Cuadro de búsqueda

Configura el tamaño y el texto del marcador de posición del cuadro de búsqueda, además de permitir sugerencias de consultas.

Ajuste de relevancia

Añada potenciación a los campos de metadatos de los documentos para darles más peso cuando los usuarios busquen documentos. Puede añadir un peso que comience en 1 y aumente gradualmente hasta 10. Puede potenciar los tipos de campos de texto, fecha y numéricos. Por ejemplo, para dar a `_last_updated_at` y `_created_at` más peso o importancia que a otros campos, asigne a estos campos un peso de 1 a 10, según su importancia. Puede aplicar diferentes configuraciones de ajuste de relevancia para cada aplicación o experiencia de búsqueda.

Proporcionar acceso a la página de búsqueda

El acceso a su experiencia de búsqueda se realiza a través de IAM Identity Center. Al configurar su experiencia de búsqueda, permite que otras personas incluidas en el directorio de su Identity Center accedan a su página de búsqueda de Amazon Kendra. Estas reciben un correo electrónico en el que se les indica que inicien sesión con sus credenciales en IAM Identity Center para acceder a la página de búsqueda. Debe configurar IAM Identity Center a nivel de organización o a nivel de titular de la

cuenta en AWS Organizations. Para obtener más información sobre configurar IAM Identity Center, consulte Primeros pasos con IAM Identity Center.

Puede activar las identidades de los usuarios en IAM Identity Center con su experiencia de búsqueda y asignar permisos de acceso de Lector o Propietario mediante la API o la consola.

- Lector: se le permite realizar consultas, recibir sugerencias de respuestas relacionadas con su búsqueda y contribuir con sus comentarios a Amazon Kendra para seguir mejorando la búsqueda.
- Propietario: se le permite personalizar el diseño de la página de búsqueda, ajustar la búsqueda y utilizar la aplicación de búsqueda como Lector. Actualmente, no se admite la desactivación del acceso a los lectores en la consola.

Para asignar a otras personas acceso a su experiencia de búsqueda, primero debe activar las identidades de usuario en IAM Identity Center con su experiencia de Amazon Kendra mediante el objeto ExperienceConfiguration. Debe especificar el nombre del campo que contiene los identificadores de sus usuarios, como el nombre de usuario o la dirección de correo electrónico. A continuación, concede a su lista de usuarios acceso a su experiencia de búsqueda mediante la API AssociateEntitiesToExperience y define sus permisos como Lector o Propietario mediante la API AssociatePersonasToEntities. Especifica cada usuario o grupo mediante el objeto EntityConfiguration y si ese usuario o grupo es un Espectador o un Propietario mediante el objeto EntityPersonaConfiguration.

Para asignar a otras personas acceso a su experiencia de búsqueda mediante la consola, primero debe crear una experiencia y confirmar su identidad y que usted es el propietario. A continuación, puede asignar a otros usuarios o grupos como lectores o propietarios. En la consola, seleccione su índice y, a continuación, seleccione Experiencias en el menú de navegación. Tras crear la experiencia, puede seleccionarla de la lista. Vaya a Administración de acceso para asignar usuarios o grupos como lectores o propietarios.

Configurar una experiencia de búsqueda

A continuación se muestra un ejemplo de configuración o creación de una experiencia de búsqueda.

Console

Para crear una experiencia de búsqueda de Amazon Kendra

1. En el panel de navegación izquierdo, en Índices, seleccione Experiencias y, a continuación, seleccione Crear experiencia.

2. En la página Configurar experiencia, introduzca un nombre y una descripción para su experiencia, elija sus fuentes de contenido y elija el rol de IAM para su experiencia. Para obtener más información acerca de roles de IAM, consulte Roles de IAM para experiencias de Amazon Kendra.
3. En la página Confirmar su identidad desde un directorio de Identity Center, seleccione su ID de usuario, como su correo electrónico. Si no tiene un directorio de Identity Center, simplemente introduzca su nombre completo y su correo electrónico para crear un directorio de Identity Center. Esto lo incluye como usuario de la experiencia y le asigna automáticamente derechos de acceso de propietario.
4. En la página Revisar para abrir Experience Builder, revise los detalles de configuración y seleccione Crear experiencia y abrir Experience Builder para empezar a editar la página de búsqueda.

CLI

Para crear una experiencia de Amazon Kendra

```
aws kendra create-experience \
 --name experience-name \
 --description "experience description" \
 --index-id index-id \
 --role-arn arn:aws:iam::account-id:role/role-name \
 --configuration '{"ExperienceConfiguration":[{"ContentSourceConfiguration":
{"DataSourceIds":["data-source-1","data-source-2"]},
 "UserIdentityConfiguration":"identity attribute name"}]}'

aws kendra describe-experience \
 --endpoints experience-endpoint-URL(s)
```

Python

Para crear una experiencia de Amazon Kendra

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")
```

```
print("Create an experience.")

# Provide a name for the experience
name = "experience-name"
# Provide an optional description for the experience
description = "experience description"
# Provide the index ID for the experience
index_id = "index-id"
# Provide the IAM role ARN required for Amazon Kendra experiences
role_arn = "arn:aws:iam::${account-id}:role/${role-name}"
# Configure the experience
configuration = {"ExperienceConfiguration":
        [{
            "ContentSourceConfiguration":{"DataSourceIds":["data-source-1","data-
source-2"]},
            "UserIdentityConfiguration":"identity attribute name"
        }]
    }

try:
    experience_response = kendra.create_experience(
        Name = name,
        Description = description,
        IndexId = index_id,
        RoleArn = role_arn,
        Configuration = configuration
    )

    pprint.pprint(experience_response)

    experience_endpoints = experience_response["Endpoints"]

    print("Wait for Amazon Kendra to create the experience.")

    while True:
        # Get the details of the experience, such as the status
        experience_description = kendra.describe_experience(
            Endpoints = experience_endpoints
        )
        status = experience_description["Status"]
        print(" Creating experience. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break
```

```
except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

Para crear un Amazon Kendra

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateExperienceRequest;
import software.amazon.awssdk.services.kendra.model.CreateExperienceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeExperienceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeExperienceResponse;
import software.amazon.awssdk.services.kendra.model.ExperienceStatus;


public class CreateExperienceExample {

    public static void main(String[] args) throws InterruptedException {
        System.out.println("Create an experience");

        String experienceName = "experience-name";
        String experienceDescription = "experience description";
        String indexId = "index-id";
        String experienceRoleArn = "arn:aws:iam::account-id:role/role-name";

        KendraClient kendra = KendraClient.builder().build();

        CreateExperienceRequest createExperienceRequest = CreateExperienceRequest
            .builder()
            .name(experienceName)
            .description(experienceDescription)
            .roleArn(experienceRoleArn)
            .configuration(
                ExperienceConfiguration
                    .builder()
                    .contentSourceConfiguration(
                        ContentSourceConfiguration(
                            .builder()
```

```
                            .dataSourceIds("data-source-1","data-source-2")
                            .build()
                        )
                    )
                    .userIdentityConfiguration(
                        UserIdentityConfiguration(
                            .builder()
                            .identityAttributeName("identity-attribute-name")
                            .build()
                        )
                    ).build()
            ).build();

        CreateExperienceResponse createExperienceResponse =
 kendra.createExperience(createExperienceRequest);
        System.out.println(String.format("Experience response %s",
 createExperienceResponse));

        String experienceEndpoints = createExperienceResponse.endpoints();

        System.out.println(String.format("Wait for Kendra to create the
 experience.", experienceEndpoints));
        while (true) {
            DescribeExperienceRequest describeExperienceRequest =
 DescribeExperienceRequest.builder().endpoints(experienceEndpoints).build();
            DescribeExperienceResponse describeEpxerienceResponse =
 kendra.describeExperience(describeExperienceRequest);
            ExperienceStatus status = describeExperienceResponse.status();
            TimeUnit.SECONDS.sleep(60);
            if (status != ExperienceStatus.CREATING) {
                break;
            }
        }

        System.out.println("Experience creation is complete.");
    }
}
```

Ajuste de la capacidad

Amazon Kendra proporciona recursos para su índice en unidades de capacidad. Cada unidad de capacidad proporciona recursos adicionales para el índice. Hay unidades de capacidad independientes para el almacenamiento de documentos y para las consultas. Solo puede añadir unidades de capacidad a los índices de Amazon Kendra Enterprise Edition. No puede agregar capacidad a un índice de Developer Edition.

Una unidad de capacidad de almacenamiento de documentos proporciona el siguiente espacio de almacenamiento adicional para el índice.

- 100 000 documentos o 30 GB de almacenamiento.

Una unidad de capacidad de consulta proporciona las siguientes consultas adicionales para el índice.

- 0,1 consultas de por segundo o aproximadamente 8000 consultas al día.

Cada índice incluye una capacidad base equivalente a 1 unidad de capacidad (30 GB de almacenamiento y 0,1 consultas por segundo). Hay un coste adicional por cada unidad de capacidad adicional. Consulte Precios de Amazon Kendra para obtener más información.

Puede agregar hasta 100 unidades de capacidad adicional a sus recursos de almacenamiento y de consulta para un índice. Si necesita más unidades, simplemente póngase en contacto con Support.

Puede ajustar las unidades de capacidad hasta 5 veces al día para adaptarse a sus necesidades de uso. No puede reducir la capacidad de almacenamiento de documentos por debajo del número de documentos almacenados en el índice. Por ejemplo, si almacena 150 000 documentos, no puede reducir la capacidad de almacenamiento por debajo de 1 unidad adicional.

Puede ver los recursos que utiliza un índice en la consola seleccionando el nombre del índice para abrir la configuración del índice y otra información, o bien puede utilizar la DescribeIndexAPI.

Amazon Kendra también devuelve excepciones cuando se supera la capacidad de un índice. Se obtiene un valor `ServiceQuotaExceededException` cuando el tamaño total extraído de todos los documentos supera el límite de un índice. Se obtiene un valor `InvalidRequest` para cada documento cuando el número de documentos supere el límite de un índice. Se obtiene un valor `ThrottlingException` cuando el número de consultas por segundo supera el límite. Para obtener más información sobre los límites, consulte Cuotas para Amazon Kendra.

Las consultas acumuladas durarán hasta 24 horas.

Visualización de la capacidad

Para ver los recursos que utiliza el índice con la Amazon Kendra consola, seleccione el nombre del índice para acceder a los detalles. La consola también proporciona gráficos de uso para que pueda determinar la capacidad de almacenamiento y consulta que utiliza el índice. Puede utilizar esta información para planificar cuándo agregar capacidad adicional.

Para ver el almacenamiento de documentos y el uso de consultas (consola)

1. Inicie sesión AWS Management Console y abra la Amazon Kendra consola en https://console.aws.amazon.com/kendra/home.
2. En la lista de índices, seleccione el índice al que desea acceder.
3. Desplácese hasta la sección de configuración para ver la capacidad total actual de almacenamiento de documentos y consulta.

Para ver la capacidad mediante la Amazon Kendra API, utilice el `CapacityUnits` parámetro de la DescribeIndexAPI.

Agregar y eliminar capacidad

Si necesita capacidad adicional para su índice, puede agregarla mediante la consola o la Amazon Kendra API.

Para agregar o eliminar capacidad de almacenamiento o de consulta (consola)

1. Inicie sesión AWS Management Console y abra la Amazon Kendra consola en https://console.aws.amazon.com/kendra/home.
2. En la lista de índices, seleccione el índice al que desea acceder.
3. Seleccione Editar o seleccione Editar en el menú desplegable Acciones.
4. Seleccione Siguiente para ir a la página de detalles de aprovisionamiento.
5. Agregue o elimine unidades de capacidad de almacenamiento de documentos o de consulta.
6. Continúe seleccionando Siguiente para ir a la página de revisión y, a continuación, seleccione Actualizar para guardar los cambios.

Después de la actualización de la capacidad del índice, pueden pasar varios minutos hasta que los cambios surtan efecto.

Para añadir o eliminar capacidad mediante la Amazon Kendra API, utiliza el `CapacityUnits` parámetro de la UpdateIndexAPI.

Amazon Kendra Capacidad de clasificación inteligente

Una unidad de capacidad proporciona las siguientes solicitudes de repuntuación adicionales por segundo para un plan de ejecución de repuntuación. Un plan de ejecución de repuntuación es un recurso que se utiliza para aprovisionar la API Rescore.

- 0,01 solicitudes por segundo.

Cada plan de ejecución de repuntuación incluye una capacidad base igual a 1 unidad de capacidad (0,01 solicitudes por segundo). Hay un coste adicional por cada unidad de capacidad adicional. Consulte Precios de Amazon Kendra para obtener más información.

Puede agregar hasta 1000 unidades de capacidad adicional para un plan de ejecución de repuntuación. Si necesita más unidades, simplemente póngase en contacto con Support.

Capacidad de sugerencias de consulta

Cuando se utilizan sugerencias de consultas, hay una capacidad de consulta básica de 2,5 GetQuerySuggestionsllamadas por segundo. La capacidad de `GetQuerySuggestions` es cinco veces la capacidad de consulta aprovisionada para un índice o la capacidad base de 2,5 llamadas por segundo, la que sea mayor. Por ejemplo, la capacidad base de un índice es de 0,1 consultas por segundo, y la capacidad de `GetQuerySuggestions` tiene una base de 2,5 llamadas por segundo. Si se agregan otras 0,1 consultas por segundo al total de 0,2 consultas por segundo para un índice, la capacidad de `GetQuerySuggestions` es de 2,5 llamadas por segundo (mayor que cinco veces 0,2 consultas por segundo).

Amazon Kendra capacidad de experiencia

Capacidad de experiencia de búsqueda

Amazon Kendra comienza a acelerarse `QueryQuerySuggestions`, `SubmitFeedback` para su Amazon Kendra experiencia, a 15 solicitudes por segundo y 40 solicitudes por segundo a ráfagas de

consultas. En el caso de un índice con más de 150 unidades de capacidad de consulta, se seguirán aplicando estos límites.

Por ejemplo, las unidades de capacidad de consulta del índice son 150, por lo que la aplicación de experiencia de búsqueda puede gestionar 15 solicitudes por segundo. Sin embargo, si lo escalara a 200 unidades de capacidad de consulta, su aplicación de experiencia de búsqueda seguiría gestionando solo 15 solicitudes por segundo. Si limita el índice a 100 unidades de capacidad de consulta, su aplicación de experiencia de búsqueda gestionaría solo 10 solicitudes por segundo.

Ráfaga de consultas adaptativas

Amazon Kendra tiene una capacidad base aprovisionada de 1 unidad de capacidad de consulta. Puede utilizar hasta 8000 consultas al día con un rendimiento mínimo de 0,1 consultas por segundo (por unidad de capacidad de consulta). Las consultas acumuladas durarán hasta 24 horas y pueden adaptarse a ráfagas de tráfico. La cantidad de ráfaga permitida varía porque depende de la carga del clúster en un momento dado. Aprovisione suficientes unidades de capacidad de consulta para gestionar los picos de carga.

Un enfoque adaptativo para gestionar ráfagas de tráfico inesperadas que superen el rendimiento previsto es la fragmentación Amazon Kendra de consultas adaptativa integrada. La ráfaga de consultas adaptativas está disponible en la edición Enterprise de Amazon Kendra.

La fragmentación adaptativa de consultas es una función integrada que permite utilizar la capacidad de consulta no utilizada para gestionar el tráfico inesperado. Amazon Kendra acumula las consultas no utilizadas a una tasa de consultas aprovisionadas por segundo, cada segundo, hasta el número máximo de consultas que haya aprovisionado para su índice. Amazon Kendra Estas consultas acumuladas se utilizan para el tráfico inesperado que supere la capacidad asignada. El rendimiento óptimo de la ráfaga de consultas adaptativas puede variar en función de varios factores, como el tamaño total del índice, la complejidad de las consultas, la acumulación de consultas no utilizadas y la carga general del índice. Se recomienda realizar sus propias pruebas de carga para medir con precisión la capacidad de transmisión por ráfagas.

Introducción

En esta sección se muestra cómo crear una fuente de datos y cómo añadir los documentos a un Amazon Kendra índice. Se proporcionan instrucciones para la AWS consola, el AWS CLI, un programa Python que usa el AWS SDK for Python (Boto3), y un programa Java que usa el AWS SDK for Java.

Temas

- Requisitos previos
- Cómo empezar a utilizar la Amazon Kendra consola
- Introducción (AWS CLI)
- Introducción (AWS SDK for Python (Boto3))
- Introducción (AWS SDK for Java)
- Introducción a un origen de datos de Amazon S3 (consola)
- Introducción a un origen de datos de base de datos de MySQL (consola)
- Introducción a una fuente de AWS IAM Identity Center identidad (consola)

Requisitos previos

Los siguientes pasos son requisitos previos para los ejercicios de introducción. Estos pasos te muestran cómo configurar tu cuenta, crear un IAM rol que dé Amazon Kendra permiso para hacer llamadas en tu nombre e indexar documentos desde un Amazon S3 bucket. Se utiliza un bucket de S3 como ejemplo, pero puede utilizar otras fuentes de datos Amazon Kendra compatibles. Consulte Orígenes de datos.

Inscríbase en un Cuenta de AWS

Si no tiene uno Cuenta de AWS, complete los siguientes pasos para crearlo.

Para suscribirte a una Cuenta de AWS

1. Abra https://portal.aws.amazon.com/billing/signup.
2. Siga las instrucciones que se le indiquen.

   Parte del procedimiento de registro consiste en recibir una llamada telefónica e indicar un código de verificación en el teclado del teléfono.

Cuando te registras en una Cuenta de AWS, Usuario raíz de la cuenta de AWSse crea una. El usuario raíz tendrá acceso a todos los Servicios de AWS y recursos de esa cuenta. Como práctica recomendada de seguridad, asigne acceso administrativo a un usuario y utilice únicamente el usuario raíz para realizar tareas que requieren acceso de usuario raíz.

AWS te envía un correo electrónico de confirmación una vez finalizado el proceso de registro. Puede ver la actividad de la cuenta y administrar la cuenta en cualquier momento entrando en https://aws.amazon.com/ y seleccionando Mi cuenta.

Creación de un usuario con acceso administrativo

Después de crear un usuario administrativo Cuenta de AWS, asegúrelo Usuario raíz de la cuenta de AWS AWS IAM Identity Center, habilite y cree un usuario administrativo para no usar el usuario root en las tareas diarias.

Proteja su Usuario raíz de la cuenta de AWS

1. Inicie sesión AWS Management Consolecomo propietario de la cuenta seleccionando el usuario root e introduciendo su dirección de Cuenta de AWS correo electrónico. En la siguiente página, escriba su contraseña.

   Para obtener ayuda para iniciar sesión con el usuario raíz, consulte Signing in as the root user en la Guía del usuario de AWS Sign-In .

2. Active la autenticación multifactor (MFA) para el usuario raíz.

   Para obtener instrucciones, consulte Habilitar un dispositivo MFA virtual para el usuario Cuenta de AWS raíz (consola) en la Guía del usuario de IAM.

Creación de un usuario con acceso administrativo

1. Activar IAM Identity Center.

   Consulte las instrucciones en Activar AWS IAM Identity Center en la Guía del usuario de AWS IAM Identity Center .

2. En IAM Identity Center, conceda acceso administrativo a un usuario.

Para ver un tutorial sobre su uso Directorio de IAM Identity Center como fuente de identidad, consulte Configurar el acceso de los usuarios con la configuración predeterminada Directorio de IAM Identity Center en la Guía del AWS IAM Identity Center usuario.

Iniciar sesión como usuario con acceso de administrador

- Para iniciar sesión con el usuario de IAM Identity Center, utilice la URL de inicio de sesión que se envió a la dirección de correo electrónico cuando creó el usuario de IAM Identity Center.

  Para obtener ayuda para iniciar sesión con un usuario del Centro de identidades de IAM, consulte Iniciar sesión en el portal de AWS acceso en la Guía del AWS Sign-In usuario.

Concesión de acceso a usuarios adicionales

1. En IAM Identity Center, cree un conjunto de permisos que siga la práctica recomendada de aplicar permisos de privilegios mínimos.

   Para conocer las instrucciones, consulte Create a permission set en la Guía del usuario de AWS IAM Identity Center .
2. Asigne usuarios a un grupo y, a continuación, asigne el acceso de inicio de sesión único al grupo.

   Para conocer las instrucciones, consulte Add groups en la Guía del usuario de AWS IAM Identity Center .

- Si utiliza un depósito de S3 que contiene documentos para realizar pruebas Amazon Kendra, cree un depósito de S3 en la misma región que utiliza Amazon Kendra. Para obtener instrucciones, consulte Crear y configurar un bucket de S3 en la guía del usuario de Amazon Simple Storage Service.

  Carga de los documentos en el bucket de S3. Para ver las instrucciones, consulte Carga, descarga y administración de objetos en la Guía del usuario de Amazon Simple Storage Service.

  Si utiliza otro origen de datos, debe tener un sitio activo y credenciales para conectarse al origen de datos.

Si va a utilizar la consola para empezar, comience con Cómo empezar a utilizar la Amazon Kendra consola.

Amazon Kendra recursos: SDK AWS CLI, consola

Se requieren ciertos permisos si usa la CLI, el SDK o la consola.

Para usarlo Amazon Kendra para la CLI, el SDK o la consola, debe tener permisos que le Amazon Kendra permitan crear y administrar recursos en su nombre. Según su caso de uso, estos permisos incluyen el acceso a la propia Amazon Kendra API, AWS KMS keys si desea cifrar sus datos a través de un CMK personalizado o un directorio de Identity Center si desea integrarlos con una experiencia de búsqueda AWS IAM Identity Center o crearla. Para obtener una lista completa de los permisos para los distintos casos de uso, consulte Roles de IAM.

En primer lugar, debe adjuntar los siguientes permisos a su usuario de IAM.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Stmt1644430853544",
      "Action": [
        "kms:CreateGrant",
        "kms:DescribeKey"
      ],
      "Effect": "Allow",
      "Resource": "*"
    },
    {
      "Sid": "Stmt1644430878150",
      "Action": "kendra:*",
      "Effect": "Allow",
      "Resource": "*"
    },
    {
      "Sid": "Stmt1644430973706",
      "Action": [
        "sso:AssociateProfile",
        "sso:CreateManagedApplicationInstance",
        "sso:DeleteManagedApplicationInstance",
        "sso:DisassociateProfile",
        "sso:GetManagedApplicationInstance",
        "sso:GetProfile",
```

```
        "sso:ListDirectoryAssociations",
        "sso:ListProfileAssociations",
        "sso:ListProfiles"
      ],
      "Effect": "Allow",
      "Resource": "*"
    },
    {
      "Sid": "Stmt1644430999558",
      "Action": [
        "sso-directory:DescribeGroup",
        "sso-directory:DescribeGroups",
        "sso-directory:DescribeUser",
        "sso-directory:DescribeUsers"
      ],
      "Effect": "Allow",
      "Resource": "*"
    },
    {
      "Sid": "Stmt1644431025960",
      "Action": [
        "identitystore:DescribeGroup",
        "identitystore:DescribeUser",
        "identitystore:ListGroups",
        "identitystore:ListUsers"
      ],
      "Effect": "Allow",
      "Resource": "*"
    }
  ]
}
```

En segundo lugar, si usa la CLI o el SDK, también debe crear un IAM rol y una política para acceder Amazon CloudWatch Logs. Si está utilizando la consola, no tendrá que crear un rol de IAM ni una política para ello. Esto se crea como parte del procedimiento de la consola.

Para crear un IAM rol y una política para el AWS CLI SDK que le permitan acceder Amazon Kendra a su Amazon CloudWatch Logs.

1. Inicie sesión en la consola de IAM AWS Management Console y ábrala en https://console.aws.amazon.com/iam/.
2. En el menú de la izquierda, elija Políticas y, a continuación, seleccione Crear política.

3. Seleccione JSON y sustituya la política predeterminada con lo siguiente:

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "cloudwatch:PutMetricData"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "cloudwatch:namespace": "AWS/Kendra"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "logs:DescribeLogGroups"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "logs:CreateLogGroup"
            ],
            "Resource": [
                "arn:aws:logs:region:account ID:log-group:/aws/kendra/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "logs:DescribeLogStreams",
                "logs:CreateLogStream",
                "logs:PutLogEvents"
            ],
            "Resource": [
                "arn:aws:logs:region:account ID:log-group:/aws/kendra/*:log-
stream:*"
```

```
            ]
        }
    ]
}
```

4. Elija Revisar política.
5. Asigne a la política el nombre "KendraPolicyForGettingStartedIndex" y, a continuación, seleccione Crear política.
6. En el menú de la izquierda, seleccione Roles y, a continuación, Crear rol.
7. Selecciona Otra AWS cuenta y, a continuación, escribe tu ID de cuenta en ID de cuenta. Elija Siguiente: permisos.
8. Elija la política que creó anteriormente y, a continuación, seleccione Siguiente: etiquetas
9. No añada ninguna etiqueta. Elija Siguiente: revisar.
10. Asigne al rol el nombre "KendraRoleForGettingStartedIndex" y, a continuación, seleccione Crear rol.
11. Busque el rol que acaba de crear. Seleccione el nombre del rol para abrir el resumen. Seleccione Relaciones de confianza, y, a continuación, seleccione Editar relación de confianza.
12. Reemplace la relación de confianza existente con lo siguiente:

```
{
    "Version": "2012-10-17",
    "Statement": [
      {
        "Effect": "Allow",
        "Principal": {
          "Service": "kendra.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
      }
    ]
}
```

13. Elija Actualizar política de confianza.

En tercer lugar, si utilizas una Amazon S3 para almacenar tus documentos o utilizas S3 para realizar pruebas Amazon Kendra, también debes crear un IAM rol y una política para acceder a tu bucket. Si utiliza otro origen de datos, consulte los Roles de IAM para orígenes de datos.

Para crear un IAM rol y una política que le permitan acceder Amazon Kendra a su Amazon S3 bucket e indexarlo.

1. Inicie sesión en la consola de IAM AWS Management Console y ábrala en https://console.aws.amazon.com/iam/.
2. En el menú de la izquierda, elija Políticas y, a continuación, seleccione Crear política.
3. Seleccione JSON y sustituya la política predeterminada con lo siguiente:

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket name/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::bucket name"
            ],
            "Effect": "Allow"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:BatchPutDocument",
                "kendra:BatchDeleteDocument"
            ],
            "Resource": "arn:aws:kendra:region:account ID:index/*"
        }
    ]
}
```

4. Elija Revisar política.

5. Asigne el nombre "KendraPolicyForGettingStartedDataSource" a la política y, a continuación, seleccione Crear política.
6. En el menú de la izquierda, seleccione Roles y, a continuación, Crear rol.
7. Elige Otra AWS cuenta y, a continuación, escribe tu ID de cuenta en ID de cuenta. Elija Siguiente: permisos.
8. Elija la política que creó anteriormente y, a continuación, seleccione Siguiente: etiquetas
9. No añada ninguna etiqueta. Elija Siguiente: revisar.
10. Asigne el nombre KendraRoleForGettingStartedDataSource "» al rol y, a continuación, seleccione Crear rol.
11. Busque el rol que acaba de crear. Seleccione el nombre del rol para abrir el resumen. Seleccione Relaciones de confianza, y, a continuación, seleccione Editar relación de confianza.
12. Reemplace la relación de confianza existente con lo siguiente:

```
{
    "Version": "2012-10-17",
    "Statement": [
      {
        "Effect": "Allow",
        "Principal": {
          "Service": "kendra.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
      }
    ]
}
```

13. Elija Actualizar política de confianza.

En función de cómo desee utilizar la Amazon Kendra API, realice una de las siguientes acciones.

- Introducción (AWS CLI)
- Introducción (AWS SDK for Java)
- Introducción (AWS SDK for Python (Boto3))

Cómo empezar a utilizar la Amazon Kendra consola

Los siguientes procedimientos muestran cómo crear y probar un Amazon Kendra índice mediante la AWS consola. En los procedimientos, se crean un índice y un origen de datos para un índice. Por último, se prueba el índice realizando una solicitud de búsqueda.

Paso 1: Crear un índice (consola)

1. Inicie sesión en la consola AWS de administración y abra la Amazon Kendra consola en https://console.aws.amazon.com/kendra/.
2. Seleccione Create index (Crear índice) en la sección Indexes (Índices).
3. En la página Especificar detalles de índice, proporcione a su índice un nombre y una descripción.
4. En Rol de IAM , elija Crear un nuevo rol y, a continuación, asigne un nombre al rol. El IAM rol tendrá el prefijo "AmazonKendra-».
5. No cambie los demás valores predeterminados. Elija Siguiente.
6. En la página Configurar acceso del cliente, elija Siguiente paso.
7. En la página Detalles de aprovisionamiento, elija la Edición para desarrolladores.
8. Elija Crear para crear el índice.
9. Espere a que se cree el índice. Amazon Kendra aprovisiona el hardware para su índice. Esta operación puede llevar algún tiempo.

Paso 2: Añadir un origen de datos a un índice (consola)

1. Vea las fuentes de datos disponibles para conectarse Amazon Kendra a sus documentos e indexarlos.
2. En el panel de navegación, seleccione Orígenes de datos y, a continuación, seleccione Añadir origen de datos para el origen de datos elegido.
3. Siga los pasos para configurar el origen de datos.

Paso 3: Buscar en un índice (consola)

1. En el panel de navegación, elija la opción para buscar en el índice.
2. Introduzca un término de búsqueda que sea adecuado para su índice. Se muestran los resultados principales y los principales resultados de documentos.

Introducción (AWS CLI)

El siguiente procedimiento muestra cómo crear un Amazon Kendra índice mediante AWS CLI. El procedimiento crea un origen de datos y un índice y ejecuta una consulta en el índice.

Para crear un Amazon Kendra índice (CLI)

1. Realice los Requisitos previos.
2. Introduzca el siguiente comando para crear un índice.

```
aws kendra create-index \
 --name cli-getting-started-index \
 --description "Index for CLI getting started guide." \
 --role-arn arn:aws:iam::account id:role/KendraRoleForGettingStartedIndex
```

3. Espere Amazon Kendra a que se cree el índice. Compruebe el progreso usando el comando siguiente. Cuando el campo de estado sea `ACTIVE`, vaya al siguiente paso.

```
aws kendra describe-index \
 --id index id
```

4. En el símbolo del sistema, introduzca el siguiente comando para crear un origen de datos.

```
aws kendra create-data-source \
 --index-id index id \
 --name data source name \
 --role-arn arn:aws:iam::account id:role/KendraRoleForGettingStartedDataSource \
 --type S3 \
 --configuration '{"S3Configuration":{"BucketName":"S3 bucket name"}}'
```

Si se conecta al origen de datos mediante un esquema de plantilla, configure el esquema de plantilla.

```
aws kendra create-data-source \
 --index-id index id \
 --name data source name \
 --role-arn arn:aws:iam::account id:role/KendraRoleForGettingStartedDataSource \
 --type TEMPLATE \
 --configuration '{"TemplateConfiguration":{"Template":{JSON schema}}}'
```

5. La creación de la fuente de datos tardará Amazon Kendra un tiempo. Para verificar el progreso, ingrese el siguiente comando. Cuando el estado sea `ACTIVE`, vaya al siguiente paso.

```
aws kendra describe-data-source \
 --id data source ID \
 --index-id index ID
```

6. Ingrese el comando siguiente para sincronizar el origen de datos.

```
aws kendra start-data-source-sync-job \
 --id data source ID \
 --index-id index ID
```

7. Amazon Kendra indexará su fuente de datos. La cantidad de tiempo que tarde depende del número de documentos. Puede comprobar el estado del trabajo de sincronización mediante el siguiente comando. Cuando el estado sea `ACTIVE`, vaya al siguiente paso.

```
aws kendra describe-data-source \
 --id data source ID \
 --index-id index ID
```

8. Ingrese el comando siguiente para realizar una consulta.

```
aws kendra query \
 --index-id index ID \
 --query-text "search term"
```

   Los resultados de la búsqueda se muestran en formato JSON.

Introducción (AWS SDK for Python (Boto3))

El siguiente programa es un ejemplo de uso Amazon Kendra en un programa de Python. En el programa se realizan las siguientes tareas:

1. Crea un índice nuevo mediante la CreateIndexoperación.
2. Espera a que se complete la creación del índice. Utiliza la DescribeIndexoperación para supervisar el estado del índice.
3. Una vez que el índice está activo, crea una fuente de datos mediante la CreateDataSourceoperación.

4. Espera a que se complete la creación del origen de datos. Utiliza la DescribeDataSourceoperación para supervisar el estado de la fuente de datos.
5. Cuando la fuente de datos está activa, sincroniza el índice con el contenido de la fuente de datos mediante la StartDataSourceSyncJoboperación.

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create an index.")

# Provide a name for the index
index_name = "python-getting-started-index"
# Provide an optional decription for the index
description = "Getting started index"
# Provide the IAM role ARN required for indexes
index_role_arn = "arn:aws:iam::${accountId}:role/KendraRoleForGettingStartedIndex"

try:
    index_response = kendra.create_index(
        Description = description,
        Name = index_name,
        RoleArn = index_role_arn
    )

    pprint.pprint(index_response)

    index_id = index_response["Id"]

    print("Wait for Amazon Kendra to create the index.")

    while True:
        # Get the details of the index, such as the status
        index_description = kendra.describe_index(
            Id = index_id
        )
        # When status is not CREATING quit.
        status = index_description["Status"]
        print(" Creating index. Status: "+status)
```

```
        time.sleep(60)
        if status != "CREATING":
            break

    print("Create an S3 data source.")

    # Provide a name for the data source
    data_source_name = "python-getting-started-data-source"
    # Provide an optional description for the data source
    data_source_description = "Getting started data source."
    # Provide the IAM role ARN required for data sources
    data_source_role_arn = "arn:aws:iam::${accountId}:role/
KendraRoleForGettingStartedDataSource"
    # Provide the data source connection information
    S3_bucket_name = "S3-bucket-name"
    data_source_type = "S3"
    # Configure the data source
    configuration = {"S3Configuration":
        {
            "BucketName": S3_bucket_name
        }
    }

    """
    If you connect to your data source using a template schema,
    configure the template schema
    configuration = {"TemplateConfiguration":
        {
            "Template": {JSON schema}
        }
    }
    """

    data_source_response = kendra.create_data_source(
        Name = data_source_name,
        Description = data_source_name,
        RoleArn = data_source_role_arn,
        Type = data_source_type,
        Configuration = configuration,
        IndexId = index_id
    )

    pprint.pprint(data_source_response)
```

```
    data_source_id = data_source_response["Id"]

    print("Wait for Amazon Kendra to create the data source.")

    while True:
        # Get the details of the data source, such as the status
        data_source_description = kendra.describe_data_source(
            Id = data_source_id,
            IndexId = index_id
        )
        # If status is not CREATING, then quit
        status = data_source_description["Status"]
        print(" Creating data source. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

    print("Synchronize the data source.")

    sync_response = kendra.start_data_source_sync_job(
        Id = data_source_id,
        IndexId = index_id
    )

    pprint.pprint(sync_response)

    print("Wait for the data source to sync with the index.")

    while True:

        jobs = kendra.list_data_source_sync_jobs(
            Id = data_source_id,
            IndexId = index_id
        )

        # For this example, there should be one job
        status = jobs["History"][0]["Status"]

        print(" Syncing data source. Status: "+status)
        if status != "SYNCING":
            break
        time.sleep(60)

except  ClientError as e:
```

```
        print("%s" % e)

print("Program ends.")
```

Introducción (AWS SDK for Java)

El siguiente programa es un ejemplo de uso Amazon Kendra en un programa Java. En el programa se realizan las siguientes tareas:

1. Crea un índice nuevo mediante la CreateIndexoperación.
2. Espera a que se complete la creación del índice. Utiliza la DescribeIndexoperación para supervisar el estado del índice.
3. Una vez que el índice está activo, crea una fuente de datos mediante la CreateDataSourceoperación.
4. Espera a que se complete la creación del origen de datos. Utiliza la DescribeDataSourceoperación para supervisar el estado de la fuente de datos.
5. Cuando la fuente de datos está activa, sincroniza el índice con el contenido de la fuente de datos mediante la StartDataSourceSyncJoboperación.

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.DataSourceStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJob;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJobStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
```

```
import software.amazon.awssdk.services.kendra.model.S3DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;


public class CreateIndexAndDataSourceExample {

    public static void main(String[] args) throws InterruptedException {
        System.out.println("Create an index");

        String indexDescription = "Getting started index for Kendra";
        String indexName = "java-getting-started-index";
        String indexRoleArn = "arn:aws:iam::<your AWS account ID>:role/<name of an IAM
 role>";

        System.out.println(String.format("Creating an index named %s", indexName));
        KendraClient kendra = KendraClient.builder().build();

        CreateIndexRequest createIndexRequest = CreateIndexRequest
            .builder()
            .description(indexDescription)
            .name(indexName)
            .roleArn(indexRoleArn)
            .build();
        CreateIndexResponse createIndexResponse =
 kendra.createIndex(createIndexRequest);
        System.out.println(String.format("Index response %s", createIndexResponse));

        String indexId = createIndexResponse.id();

        System.out.println(String.format("Waiting until the index with index ID %s is
 created", indexId));
        while (true) {
            DescribeIndexRequest describeIndexRequest =
 DescribeIndexRequest.builder().id(indexId).build();
            DescribeIndexResponse describeIndexResponse =
 kendra.describeIndex(describeIndexRequest);
            IndexStatus status = describeIndexResponse.status();
            if (status != IndexStatus.CREATING) {
                break;
            }

            TimeUnit.SECONDS.sleep(60);
        }
```

```
        System.out.println("Creating an S3 data source");
        String dataSourceName = "java-getting-started-data-source";
        String dataSourceDescription = "Getting started data source";
        String s3BucketName = "an-aws-kendra-test-bucket";
        String dataSourceRoleArn = "arn:aws:iam::<your AWS account ID>:role/<name of an
IAM role>";

        CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
            .builder()
            .indexId(indexId)
            .name(dataSourceName)
            .description(dataSourceDescription)
            .roleArn(dataSourceRoleArn)
            .type(DataSourceType.S3)
            .configuration(
                DataSourceConfiguration
                    .builder()
                    .s3Configuration(
                        S3DataSourceConfiguration
                            .builder()
                            .bucketName(s3BucketName)
                            .build()
                    ).build()
            ).build();

        CreateDataSourceResponse createDataSourceResponse =
kendra.createDataSource(createDataSourceRequest);
        System.out.println(String.format("Response of creating data source: %s",
createDataSourceResponse));

        String dataSourceId = createDataSourceResponse.id();
        System.out.println(String.format("Waiting for Kendra to create the data source
%s", dataSourceId));
        DescribeDataSourceRequest describeDataSourceRequest = DescribeDataSourceRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            DescribeDataSourceResponse describeDataSourceResponse =
kendra.describeDataSource(describeDataSourceRequest);
```

```
            DataSourceStatus status = describeDataSourceResponse.status();
            System.out.println(String.format("Creating data source. Status: %s",
status));
            if (status != DataSourceStatus.CREATING) {
                break;
            }

            TimeUnit.SECONDS.sleep(60);
        }

        System.out.println(String.format("Synchronize the data source %s",
dataSourceId));
        StartDataSourceSyncJobRequest startDataSourceSyncJobRequest =
StartDataSourceSyncJobRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();
        StartDataSourceSyncJobResponse startDataSourceSyncJobResponse =
kendra.startDataSourceSyncJob(startDataSourceSyncJobRequest);
        System.out.println(String.format("Waiting for the data
source to sync with the index %s for execution ID %s", indexId,
startDataSourceSyncJobResponse.executionId()));

        // For this particular list, there should be just one job
        ListDataSourceSyncJobsRequest listDataSourceSyncJobsRequest =
ListDataSourceSyncJobsRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            ListDataSourceSyncJobsResponse listDataSourceSyncJobsResponse =
kendra.listDataSourceSyncJobs(listDataSourceSyncJobsRequest);
            DataSourceSyncJob job = listDataSourceSyncJobsResponse.history().get(0);
            System.out.println(String.format("Syncing data source. Status: %s",
job.status()));

            if (job.status() != DataSourceSyncJobStatus.SYNCING) {
                break;
            }

            TimeUnit.SECONDS.sleep(60);
```

```
        }

        System.out.println("Index setup is complete");
    }
}
```

Introducción a un origen de datos de Amazon S3 (consola)

Puede utilizar la consola de Amazon Kendra para empezar a utilizar un bucket de Amazon S3 como almacén de datos. Cuando usa la consola, especifica toda la información de conexión que necesita para indexar el contenido del bucket. Para obtener más información, consulte Amazon S3.

Utilice el siguiente procedimiento para crear un origen de datos de bucket de S3 básico con la configuración predeterminada. En el procedimiento, también se asume que ha creado un índice siguiendo las instrucciones del paso 1 de Cómo empezar a utilizar la Amazon Kendra consola.

Para crear un origen de datos de bucket de S3 con la consola de Amazon Kendra

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon Kendra en https://console.aws.amazon.com/kendra/home.
2. En la lista de índices, elija el índice al que quiera agregar el origen de datos.
3. Elija Agregar origen de datos.
4. En la lista de conectores de origen de datos, elija Amazon S3.
5. En la página Definir atributos, asigne un nombre al origen de datos y, si lo desea, una descripción. Deje el campo Etiquetas en blanco. Elija Siguiente para continuar.
6. En el campo Introduzca la ubicación del origen de datos, introduzca el nombre del bucket de S3 que contiene los documentos. Puede introducir el nombre directamente o puede buscarlo seleccionando Examinar. El bucket debe estar en la misma región que el índice.
7. En Rol de IAM, elija Crear un nuevo rol y, a continuación, escriba un nombre de rol. Para obtener más información, consulte Roles de IAM para orígenes de datos de Amazon S3.
8. En la sección Establecer un programa de ejecución de sincronización, elija Ejecutar bajo demanda.
9. Elija Siguiente para continuar.
10. En la página Revisar y crear, revise los detalles del origen de datos de S3. Si desea realizar cambios, pulse el botón Editar situado junto al elemento que desee cambiar. Cuando esté satisfecho con sus opciones, elija Crear para crear su origen de datos de S3.

Tras seleccionar Crear, Amazon Kendra comienza a crear el origen de datos. La creación del origen de datos puede tardar varios minutos. Cuando haya terminado, el estado del origen de datos cambia de Creando a Activo.

Tras crear el origen de datos, debe sincronizar el índice de Amazon Kendra con el origen de datos. Seleccione Sincronizar ahora para iniciar el proceso de sincronización. La sincronización del origen de datos puede tardar de varios minutos a varias horas, según la cantidad y el tamaño de los documentos.

Introducción a un origen de datos de base de datos de MySQL (consola)

Puede utilizar la consola de Amazon Kendra para empezar a utilizar una base de datos de MySQL como origen de datos. Cuando usa la consola, especifica la información de conexión que necesita para indexar el contenido de una base de datos de MySQL. Para obtener más información, consulte Uso de un origen de datos de base de datos.

Primero debe crear una base de datos de MySQL y, a continuación, crear un origen de datos para la base de datos.

Utilice el siguiente procedimiento para crear una base de datos de MySQL básica. En el procedimiento, también se asume que ya ha creado un índice después del paso 1 de Cómo empezar a utilizar la Amazon Kendra consola.

Para crear una base de datos de MySQL

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon RDS en https://console.aws.amazon.com/rds/.
2. En el panel de navegación, elija Grupos de subredes y, a continuación, elija Crear grupo de subred de base de datos.
3. Asigne un nombre al grupo y elija la nube privada virtual (VPC). Para obtener más información sobre la configuración de una VPC, consulte Configuración de Amazon Kendra para usar una VPC.
4. Agregue las subredes privadas de su VPC. Las subredes privadas son las que no están conectadas a la NAT. Seleccione Crear.
5. En el panel de navegación, seleccione Bases de datos y, a continuación, Crear base de datos.

6. Utilice los siguientes parámetros para crear la base de datos. No cambie los demás parámetros predeterminados.

   - Opciones de motor: MySQL
   - Plantillas: nivel gratuito
   - Configuración de credenciales: introduzca y confirme una contraseña
   - En Conectividad, elija Configuración de conectividad adicional. Realice los siguientes cambios.
     - Grupo de subredes: elija el grupo de subredes que ha creado en el paso 4.
     - Grupo de seguridad de VPC: elija el grupo que contiene las reglas de entrada y salida que creó en la VPC. Por ejemplo, **`DataSourceSecurityGroup`**. Para obtener más información sobre la configuración de una VPC, consulte Configuración de Amazon Kendra para usar una VPC.
   - En Configuración adicional, defina el Nombre de base de datos inicial en **`content`**.
7. Elija Crear base de datos.
8. En la lista de bases de datos, elija su nueva base de datos. Tome nota del punto de conexión de la base de datos.
9. Después de crear la base de datos, debe crear una tabla para guardar los documentos. La creación de una tabla queda fuera del ámbito de estas instrucciones. Cuando cree la tabla, tenga en cuenta lo siguiente:

   - Nombre de base de datos: **`content`**
   - Nombre de la tabla: **`documents`**
   - Columnas: **`ID`**, **`Title`**, **`Body`** y **`LastUpdate`**. Si lo desea, puede incluir columnas adicionales.

Ahora que ha creado su base de datos de MySQL, puede crear un origen de datos para la base de datos.

Para crear un origen de datos de MySQL

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon Kendra en https://console.aws.amazon.com/kendra/home.
2. Desde el panel de navegación, elija Índices y, a continuación, elija su índice.
3. Seleccione Añadir orígenes de datos y, a continuación, elija Amazon RDS.
4. Escriba un nombre y una descripción para el origen de datos y, a continuación, seleccione Siguiente.

5. Elija MySQL.
6. En Acceso de conexión, escriba la siguiente información:
   - Punto de conexión: el punto de conexión de la base de datos que creó anteriormente.
   - Puerto: el número de puerto de la base de datos. El puerto predeterminado para MySQL es el 3306.
   - Tipo de autenticación: elija Nuevo.
   - Nombre del nuevo contenedor secreto: un nombre para el contenedor Secrets Manager de las credenciales de la base de datos.
   - Nombre de usuario: el nombre de un usuario con acceso administrativo a la base de datos.
   - Contraseña: la contraseña del usuario y, a continuación, seleccione Guardar autenticación.
   - Nombre de base de datos: **`content`**.
   - Nombre de la tabla: **`documents`**.
   - Rol de IAM; elija Crear un nuevo rol y, a continuación, escriba un nombre para el rol.
7. En Configuración de columna, introduzca lo siguiente:
   - Nombre de la columna del ID del documento: **`ID`**
   - Nombre de la columna del título del documento: **`Title`**
   - Nombre de la columna de los datos del documento: **`Body`**
8. En Detección de cambios de columna, introduzca lo siguiente:
   - Columnas de detección de cambios: **`LastUpdate`**
9. En Configurar la VPC y el grupo de seguridad, proporcione lo siguiente:
   - En Nube privada virtual (VPC), seleccione su VPC.
   - En Subredes, elija las subredes que creó en su VPC.
   - En Grupos de seguridad de VPC, elija el grupo que contiene las reglas de entrada y salida que creó en su VPC para las bases de datos de MySQL. Por ejemplo, **`DataSourceSecurityGroup`**.
10. En Establecer un programa de ejecución de sincronización, elija Ejecutar bajo demanda y, a continuación, elija Siguiente.
11. En Asignación de campos de origen de datos, seleccione Siguiente.
12. Revise la configuración del origen de datos para asegurarse de que es correcta. Cuando esté seguro de que todo es correcto, elija Crear.

Introducción a una fuente de AWS IAM Identity Center identidad (consola)

Una fuente de AWS IAM Identity Center identidad contiene información sobre sus usuarios y grupos. Esto resulta útil para configurar el filtrado por contexto de usuario, en el que se Amazon Kendra filtran los resultados de búsqueda para distintos usuarios en función del acceso del usuario o de su grupo a los documentos.

Para crear un origen de identidad de IAM Identity Center, debe activar AIM Identity Center y crear una organización en AWS Organizations. Al activar AIM Identity Center y crear una organización por primera vez, automáticamente se establece de forma predeterminada en el directorio de Identity Center como origen de identidad. Puede cambiar a Active Directory (administrado por Amazon o autoadministrado) o a un proveedor de identidad externo como origen de identidad. Para ello, debe seguir las instrucciones correctas: consulte Cambio del origen de identidad de AIM Identity Center. Solo puede tener un origen de identidad por organización.

Para que a sus usuarios y grupos se les asignen diferentes niveles de acceso a los documentos, debe incluir a los usuarios y grupos en la lista de control de acceso cuando incorpore documentos a su índice. Esto permite a los usuarios y grupos buscar documentos Amazon Kendra de acuerdo con su nivel de acceso. Al realizar una consulta, el ID de usuario debe coincidir exactamente con el nombre de usuario de AIM Identity Center.

También debe conceder los permisos necesarios para utilizar IAM Identity Center con Amazon Kendra. Para obtener más información, consulte Roles de IAM para IAM Identity Center.

Para configurar un origen de identidad de IAM Identity Center

1. Abra la consola de IAM Identity Center.
2. Seleccione Activar el centro de identidad de IAM y, a continuación, seleccione Crear AWS organización.

   El directorio de Identity Center se crea de forma predeterminada y se le envía un correo electrónico para verificar la dirección de correo electrónico asociada a la organización.
3. Para añadir un grupo a su AWS organización, en el panel de navegación, elija Grupos.
4. En la página Grupos, elija Crear grupo e introduzca un nombre y una descripción del grupo en el cuadro de diálogo. Seleccione Crear.
5. Para agregar un usuario a sus Organizaciones, en el panel de navegación, elija Usuarios.

6. En la página Users (Usuarios), elija Add user (Añadir usuario). En User details (Detalles del usuario), especifique todos los campos obligatorios. En Password (Contraseña), elija Send an email to the user (Enviar un correo electrónico al usuario). Seleccione Siguiente.
7. Para añadir un usuario a un grupo, elija Grupos y seleccione un grupo.
8. En la página Detalles, bajo Miembros del grupo, elija Añadir usuario.
9. En la página Añadir usuarios a grupo, seleccione el usuario que desea añadir como miembro del grupo. Puede seleccionar varios usuarios para añadirlos a un grupo.
10. Para sincronizar la lista de usuarios y grupos con IAM Identity Center, cambie el origen de identidad a Active Directory o a un proveedor de identidad externo.

    El directorio de Identity Center es el origen de identidad predeterminada y requiere que añada manualmente sus usuarios y grupos utilizando este origen si no tiene su propia lista administrada por un proveedor. Para cambiar el origen de identidad, debe seguir las instrucciones correctas al respecto (consulte Cambio del origen de identidad de AIM Identity Center).

> **Note**
>
> Si utiliza Active Directory o un proveedor de identidad externo como origen de identidad, debe asignar las direcciones de correo electrónico de sus usuarios a los nombres de usuario de AIM Identity Center al especificar el protocolo del sistema de administración de identidades entre dominios (SCIM). Para obtener más información, consulte la Guía sobre IAM Identity Center en SCIM para habilitar IAM Identity Center.

Una vez que haya configurado el origen de identidad de IAM Identity Center, podrá activarlo en la consola al crear o editar el índice. Vaya a Control de acceso de usuarios en la configuración de su índice y edítela para poder obtener información sobre los grupos de usuarios de IAM Identity Center.

También puede activar el Centro de identidades de IAM mediante el UserGroupResolutionConfigurationobjeto. Debe proporcionar el `UserGroupResolutionMode` as `AWS_SSO` y crear un IAM rol que dé permiso para llamar `asso:ListDirectoryAssociations`,, `sso-directory:SearchUserssso-directory:ListGroupsForUser`,`sso-directory:DescribeGroups`.

> **Warning**
>
> Amazon Kendra actualmente no admite su uso `UserGroupResolutionConfiguration` con una cuenta de miembro de AWS la organización como fuente de identidad del IAM Identity Center. Debe crear el índice en la cuenta de administración de la organización para poder utilizar `UserGroupResolutionConfiguration`.

A continuación, se ofrece información general sobre cómo configurar un origen de datos con `UserGroupResolutionConfiguration` y un control de acceso de usuarios para filtrar los resultados de la búsqueda según el contexto del usuario. Esto supone que ya ha creado un índice y un IAM rol para los índices. Se crea un índice y se proporciona el IAM rol mediante la CreateIndexAPI.

Configurar un origen de datos con **`UserGroupResolutionConfiguration`** y filtrado de contexto de usuario

1. Cree un rol de IAM que dé permiso para acceder al origen de identidad de IAM Identity Center.
2. Para `UserGroupResolutionConfiguration`configurarlo, defina el modo `AWS_SSO` y llame UpdateIndexpara actualizar su índice y utilizar el IAM Identity Center.
3. Si desea utilizar el control de acceso de los usuarios basado en fichas para filtrar los resultados de la búsqueda según el contexto del usuario, configúrelo en el `USER_TOKEN` momento de UserContextPolicyla llamada. `UpdateIndex` De lo contrario, Amazon Kendra rastrea la lista de control de acceso de cada uno de sus documentos para la mayoría de los conectores de fuentes de datos. También puede filtrar los resultados de la búsqueda según el contexto del usuario en la API Query proporcionando información sobre los usuarios y los grupos en `UserContext`. También puede asignar usuarios a sus grupos de PutPrincipalMappingforma que solo necesite proporcionar el ID de usuario al realizar la consulta.
4. Cree un rol de IAM que conceda permiso para acceder a su origen de datos.
5. Configure su origen de datos. Debe proporcionar la información de conexión necesaria para conectarse a su origen de datos.
6. Cree una fuente de datos mediante la CreateDataSourceAPI. Proporcione el objeto `DataSourceConfiguration`, que incluye `TemplateConfiguration`, el ID de su índice, el rol de IAM del origen de datos y el tipo de origen de datos, y asigne un nombre al origen de datos. También puede actualizar el origen de datos.

Cambiar el origen de identidad de IAM Identity Center

> **Warning**
>
> Cambiar el origen de identidad en la Configuración de AIM Identity Center puede afectar a la conservación de la información de los usuarios y grupos. Para hacerlo de forma segura, se recomienda consultar las Consideraciones para cambiar el origen de identidad. Al cambiar el origen de identidad, se genera un nuevo ID del origen de identidad. Compruebe que está utilizando el ID correcto antes de configurar el modo como activado UserGroupResolutionConfiguration. `AWS_SSO`

Para cambiar el origen de identidad de IAM Identity Center

1. Abra la consola de IAM Identity Center.
2. Elija Configuración.
3. En la página Configuración, en Origen de identidad, seleccione Cambiar.
4. En la página Cambiar origen de identidad, seleccione el origen de identidad que prefiera y, a continuación, seleccione Siguiente.

Creación de un índice

Puede crear un índice mediante la consola o llamando a la CreateIndexAPI. Puedes usar AWS Command Line Interface (AWS CLI) o el SDK con la API. Después de crear el índice, puede agregar documentos directamente a él o desde un origen de datos.

Para crear un índice, debe proporcionar el nombre de recurso de Amazon (ARN) de una función AWS Identity and Access Management (IAM) para que los índices puedan acceder. CloudWatch Para obtener más información, consulte Roles de IAM para índices.

En las siguientes pestañas se proporciona un procedimiento para crear un índice mediante AWS Management Console, y ejemplos de código para utilizar los AWS CLI SDK de Python y Java.

Console

Para crear un índice

1. Inicie sesión en la consola AWS de administración y abra la Amazon Kendra consola en https://console.aws.amazon.com/kendra/.
2. Seleccione Create index (Crear índice) en la sección Indexes (Índices).
3. En Specify index details (Especificar detalles de índice), proporcione a su índice un nombre y una descripción.
4. En el IAM rol, proporcione un IAM rol. Para buscar un rol, elija entre los roles de su cuenta que contengan la palabra “kendra” o introduzca el nombre de otro rol. Para obtener más información sobre el rol y los permisos necesarios, consulte Roles de IAM para índices.
5. Elija Next (Siguiente).
6. En la página Configurar acceso del cliente, elija Siguiente paso. Puede actualizar el índice para utilizar tokens para el control de acceso después de crear un índice. Para más información, consulte Control de acceso a documentos.
7. En la página Detalles de aprovisionamiento, elija Crear.
8. El índice puede tardar un tiempo en crearse. Consulte la lista de índices para ver el progreso de la creación del índice. Cuando el estado del índice sea `ACTIVE`, el índice estará listo para utilizarse.

AWS CLI

Para crear un índice

1. Utilice el siguiente comando para crear un índice. `role-arn`Debe ser el nombre de recurso de Amazon (ARN) de un IAM rol que pueda ejecutar Amazon Kendra acciones. Para obtener más información, consulte Roles de IAM.

   El comando tiene formato para Linux y macOS. Si está usando Windows, reemplace el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

   ```
   aws kendra create-index \
    --name index name \
    --description "index description" \
    --role-arn arn:aws:iam::account ID:role/role name
   ```

2. El índice puede tardar un tiempo en crearse. Para comprobar el estado de su índice, utilice el ID de índice devuelto por `create-index` con el comando siguiente. Cuando el estado del índice sea `ACTIVE`, el índice estará listo para utilizarse.

   ```
   aws kendra describe-index \
    --index-id index ID
   ```

Python

Para crear un índice

- Proporcione valores para las siguientes variables en el siguiente ejemplo de código:
  - `description`: una descripción del índice que está creando. Es opcional.
  - `index_name`: el nombre del índice que está creando.
  - `role_arn`—El nombre de recurso de Amazon (ARN) de un rol que puede ejecutar Amazon Kendra API. Para obtener más información, consulte Roles de IAM.

  ```
  import boto3
  from botocore.exceptions import ClientError
  import pprint
  import time
  ```

```
kendra = boto3.client("kendra")

print("Create an index.")

# Provide a name for the index
index_name = "index-name"
# Provide an optional description for the index
description = "index description"
# Provide the IAM role ARN required for indexes
role_arn = "arn:aws:iam::${account id}:role/${role name}"

try:
    index_response = kendra.create_index(
        Name = index_name,
        Description = description,
        RoleArn = role_arn
    )

    pprint.pprint(index_response)

    index_id = index_response["Id"]

    print("Wait for Amazon Kendra to create the index.")

    while True:
        # Get the details of the index, such as the status
        index_description = kendra.describe_index(
            Id = index_id
        )
        # If status is not CREATING, then quit
        status = index_description["Status"]
        print(" Creating index. Status: "+status)
        if status != "CREATING":
            break
        time.sleep(60)

except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

Para crear un índice

- Proporcione valores para las siguientes variables en el siguiente ejemplo de código:
  - `description`: una descripción del índice que está creando. Es opcional.
  - `index_name`: el nombre del índice que está creando.
  - `role_arn`—El nombre de recurso de Amazon (ARN) de un rol que puede ejecutar Amazon Kendra API. Para obtener más información, consulte Roles de IAM.

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;


public class CreateIndexExample {

    public static void main(String[] args) throws InterruptedException {

        String indexDescription = "Getting started index for Kendra";
        String indexName = "java-getting-started-index";
        String indexRoleArn = "arn:aws:iam::<your AWS account ID>:role/
KendraRoleForGettingStartedIndex";

        System.out.println(String.format("Creating an index named %s",
 indexName));
        CreateIndexRequest createIndexRequest = CreateIndexRequest
            .builder()
            .description(indexDescription)
            .name(indexName)
            .roleArn(indexRoleArn)
            .build();
        KendraClient kendra = KendraClient.builder().build();
```

```
        CreateIndexResponse createIndexResponse =
 kendra.createIndex(createIndexRequest);
        System.out.println(String.format("Index response %s",
 createIndexResponse));

        String indexId = createIndexResponse.id();

        System.out.println(String.format("Waiting until the index with ID %s is
 created.", indexId));
        while (true) {
            DescribeIndexRequest describeIndexRequest =
 DescribeIndexRequest.builder().id(indexId).build();
            DescribeIndexResponse describeIndexResponse =
 kendra.describeIndex(describeIndexRequest);
            IndexStatus status = describeIndexResponse.status();
            if (status != IndexStatus.CREATING) {
                break;
            }

            TimeUnit.SECONDS.sleep(60);
        }

        System.out.println("Index creation is complete.");
    }
}
```

Después de crear el índice, tendrá que añadir documentos a él. Puede añadirlos directamente o crear un origen de datos que actualice el índice de forma periódica.

Temas

- Adición de documentos directamente a un índice mediante la carga por lotes
- Adición de preguntas frecuentes a un índice
- Creación de campos de documento personalizados
- Control de acceso de usuarios a los documentos con tokens

Adición de documentos directamente a un índice mediante la carga por lotes

Puede agregar documentos directamente a un índice mediante la API BatchPutDocument. No puede añadir documentos directamente con la consola. Si utiliza la consola, se conecta a un origen de datos para agregar documentos al índice. Los documentos se pueden agregar desde un bucket de S3 o suministrarse como datos binarios. Para obtener una lista de los tipos de documentos compatibles con, Amazon Kendra consulte Tipos de documentos.

La adición de documentos a un índice mediante `BatchPutDocument` es asincrónica. Después de llamar a la API `BatchPutDocument`, utilice la API BatchGetDocumentStatus para supervisar el progreso de la indexación de los documentos. Cuando se lama a la API `BatchGetDocumentStatus` con una lista de identificadores de documento, devuelve el estado del documento. Cuando el estado del documento sea `INDEXED` o `FAILED`, se habrá completado el procesamiento del documento. Cuando el estado sea `FAILED`, la API `BatchGetDocumentStatus` devolverá el motivo por el que el documento no se haya podido indexar.

Si desea modificar los campos de metadatos o atributos del contenido y el documento durante el proceso de ingesta de documentos, consulte Enriquecimiento de documentos personalizados de Amazon Kendra. Si desea utilizar un origen de datos personalizado, cada documento que envíe mediante la API `BatchPutDocument` requiere un ID de origen de datos y un ID de ejecución como atributos o campos. Para obtener más información, consulte Atributos obligatorios para orígenes de datos personalizados.

> **Note**
>
> Cada identificador de documento debe ser único por índice. No se puede crear un origen de datos para indexar los documentos con sus ID exclusivos y, a continuación, utilizar la API `BatchPutDocument` para indexar los mismos documentos o viceversa. No se puede crear un origen de datos para indexar los documentos con sus ID exclusivos y, a continuación, utilizar la API `BatchPutDocument` para indexar los mismos documentos o viceversa. El uso de las API `BatchPutDocument` y `BatchDeleteDocument` en combinación con un conector de origen de datos de Amazon Kendra para el mismo conjunto de documentos podría provocar inconsistencias en los datos. En su lugar, recomendamos utilizar el Amazon Kendra conector de origen de datos personalizado.

En los documentos de guía del desarrollador siguientes se muestra cómo añadir documentos directamente a un índice.

Temas

- Añadir documentos con la BatchPutDocument API
- Adición de documentos desde un bucket de S3

Añadir documentos con la BatchPutDocument API

En el siguiente ejemplo, se agrega un bloque de texto a un índice mediante una llamada BatchPutDocument. Puedes usar la `BatchPutDocument` API para añadir documentos directamente a tu índice. Para obtener una lista de los tipos de documentos compatibles, Amazon Kendra consulte Tipos de documentos.

Para ver un ejemplo de cómo crear un índice con los SDK AWS CLI y, consulta Crear un índice. Para configurar la CLI y los SDK, consulte Configurar Amazon Kendra.

> **Note**
>
> Los archivos añadidos al índice deben estar en un flujo de bytes codificado en UTF-8.

En los ejemplos siguientes, se añade al índice texto con codificación UTF-8.

CLI

En AWS Command Line Interface, usa el siguiente comando. El comando tiene formato para Linux y macOS. Si está usando Windows, reemplace el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

```
aws kendra batch-put-document \
   --index-id index-id \
   --documents '{"Id":"doc-id-1", "Blob":"Amazon.com is an online retailer.",
 "ContentType":"PLAIN_TEXT", "Title":"Information about Amazon.com"}'
```

Python

```
import boto3
```

```
kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"

# Provide the title and text
title = "Information about Amazon.com"
text = "Amazon.com is an online retailer."

document = {
    "Id": "1",
    "Blob": text,
    "ContentType": "PLAIN_TEXT",
    "Title": title
}

documents = [
    document
]

result = kendra.batch_put_document(
    IndexId = index_id,
    Documents = documents
)

print(result)
```

Java

```
package com.amazonaws.kendra;


import software.amazon.awssdk.core.SdkBytes;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentRequest;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentResponse;
import software.amazon.awssdk.services.kendra.model.ContentType;
import software.amazon.awssdk.services.kendra.model.Document;

public class AddDocumentsViaAPIExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();
```

```
        String indexId = "yourIndexId";

        Document testDoc = Document
            .builder()
            .title("The title of your document")
            .id("a_doc_id")
            .blob(SdkBytes.fromUtf8String("your text content"))
            .contentType(ContentType.PLAIN_TEXT)
            .build();

        BatchPutDocumentRequest batchPutDocumentRequest = BatchPutDocumentRequest
            .builder()
            .indexId(indexId)
            .documents(testDoc)
            .build();

        BatchPutDocumentResponse result =
 kendra.batchPutDocument(batchPutDocumentRequest);

        System.out.println(String.format("BatchPutDocument Result: %s", result));
    }
}
```

Adición de documentos desde un bucket de S3

Puedes añadir documentos directamente a tu índice desde un Amazon S3 bucket mediante la BatchPutDocumentAPI. Puede añadir hasta 10 documentos en la misma llamada. Cuando utilizas un bucket de S3, debes proporcionar un IAM rol con permiso para acceder al bucket que contiene tus documentos. Especifique el rol en el parámetro `RoleArn`.

El uso de la BatchPutDocumentAPI para añadir documentos desde un Amazon S3 depósito es una operación que se realiza una sola vez. Para mantener un índice sincronizado con el contenido de un depósito, cree una fuente de Amazon S3 datos. Para obtener más información, consulte Origen de datos de Amazon S3.

Para ver un ejemplo de cómo crear un índice con los SDK AWS CLI y, consulte Crear un índice. Para configurar la CLI y los SDK, consulte Configurar Amazon Kendra. Para obtener información sobre la creación de un bucket de S3, consulte la documentación de Amazon Simple Storage Service.

En el siguiente ejemplo, se agregan dos documentos de Microsoft Word al índice mediante la API `BatchPutDocument`.

Python

```
import boto3

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the IAM role ARN required to index documents in an S3 bucket
role_arn = "arn:aws:iam::${acccountID}:policy/${roleName}"

doc1_s3_file_data = {
    "Bucket": "bucket-name",
    "Key": "document1.docx"
}

doc1_document = {
    "S3Path": doc1_s3_file_data,
    "Title": "Document 1 title",
    "Id": "doc_1"
}

doc2_s3_file_data = {
    "Bucket": "bucket-name",
    "Key": "document2.docx"
}

doc2_document = {
    "S3Path": doc2_s3_file_data,
    "Title": "Document 2 title",
    "Id": "doc_2"
}

documents = [
    doc1_document,
    doc2_document
]

result = kendra.batch_put_document(
    Documents = documents,
    IndexId = index_id,
    RoleArn = role_arn
)
```

```
print(result)
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentRequest;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentResponse;
import software.amazon.awssdk.services.kendra.model.Document;
import software.amazon.awssdk.services.kendra.model.S3Path;

public class AddFilesFromS3Example {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indexId = "yourIndexId";
        String roleArn = "yourIndexRoleArn";

        Document pollyDoc = Document
            .builder()
            .s3Path(
                S3Path.builder()
                .bucket("an-aws-kendra-test-bucket")
                .key("What is Amazon Polly.docx")
                .build())
            .title("What is Amazon Polly")
            .id("polly_doc_1")
            .build();

        Document rekognitionDoc = Document
            .builder()
            .s3Path(
                S3Path.builder()
                .bucket("an-aws-kendra-test-bucket")
                .key("What is Amazon Rekognition.docx")
                .build())
            .title("What is Amazon rekognition")
            .id("rekognition_doc_1")
            .build();

        BatchPutDocumentRequest batchPutDocumentRequest = BatchPutDocumentRequest
            .builder()
```

```
            .indexId(indexId)
            .roleArn(roleArn)
            .documents(pollyDoc, rekognitionDoc)
            .build();

        BatchPutDocumentResponse result =
 kendra.batchPutDocument(batchPutDocumentRequest);

        System.out.println(String.format("BatchPutDocument result: %s", result));
    }
}
```

Adición de preguntas frecuentes a un índice

Puede agregar preguntas frecuentes directamente al índice mediante la consola o la API CreateFaq. La adición de preguntas frecuentes a un índice es asincrónica. Los datos de las preguntas frecuentes se colocan en un archivo que se almacena en un Amazon Simple Storage Service depósito. Puede utilizar archivos CSV o JSON como entrada para las preguntas frecuentes:

- CSV básico: un archivo CSV en el que cada fila contiene una pregunta, una respuesta y un URI de origen opcional.
- CSV personalizado: un archivo CSV que contiene preguntas, respuestas y encabezados para campos/atributos personalizados que puede utilizar para facetar, mostrar u ordenar las respuestas de preguntas frecuentes. También puede definir campos de control de acceso para limitar la respuesta de las preguntas frecuentes a determinados usuarios y grupos que pueden ver la respuesta de las preguntas frecuentes.
- JSON: un archivo JSON que contiene preguntas, respuestas y campos/atributos personalizados que puede utilizar para facetar, mostrar u ordenar las respuestas de preguntas frecuentes. También puede definir campos de control de acceso para limitar la respuesta de las preguntas frecuentes a determinados usuarios y grupos que pueden ver la respuesta de las preguntas frecuentes.

Por ejemplo, el siguiente es un archivo CSV básico que proporciona respuestas a preguntas sobre clínicas gratuitas en Spokane, Washington, EE. UU. y Mountain View, Missouri, EE. UU.

```
How many free clinics are in Spokane WA?, 13
How many free clinics are there in Mountain View Missouri?, 7
```

> **Note**
>
> El archivo de preguntas frecuentes debe estar con codificación UTF-8.

Temas

- Crear campos de índice para un archivo de preguntas frecuentes
- Archivo CSV básico
- Archivo CSV personalizado
- Archivo JSON
- Uso del archivo de preguntas frecuentes
- Archivos de preguntas frecuentes en idiomas distintos del inglés

Crear campos de índice para un archivo de preguntas frecuentes

Cuando utilizas un archivo CSV o JSON personalizado como entrada, puedes declarar campos personalizados para tus preguntas frecuentes. Por ejemplo, puede crear un campo personalizado que asigne cada pregunta frecuente a un departamento empresarial. Cuando se devuelve la pregunta frecuente en una respuesta, puede utilizar el departamento como faceta para restringir la búsqueda solo a “RR. HH.” o “Finanzas”, por ejemplo.

Un campo personalizado debe asignarse a un campo de índice. En la consola, utilice la página Facet definition (Definición de faceta) para crear un campo de índice. Al utilizar la API, primero debe crear un campo de índice mediante la API UpdateIndex.

El tipo de campo/atributo del archivo de preguntas frecuentes debe coincidir con el tipo del campo de índice asociado. Por ejemplo, el campo “Departamento” es un campo de tipo `STRING_LIST`. Por lo tanto, debe proporcionar en el archivo de preguntas frecuentes valores para el campo Departamento como una lista de cadenas. Puede comprobar el tipo de campos de índice utilizando la página Facet Definition (Definición de faceta) en la consola o mediante la API DescribeIndex.

Cuando se crea un campo de índice que se asigna a un atributo personalizado, se puede marcar como visualizable, facetable u ordenable. No se puede hacer que un atributo personalizado se pueda buscar.

Además de los atributos personalizados, también puede definir los campos reservados o comunes de Amazon Kendra en un archivo CSV o JSON personalizado. Para obtener más información, consulte Atributos o campos del documento.

Archivo CSV básico

Utilice un archivo CSV básico cuando desee utilizar una estructura sencilla para sus preguntas frecuentes. En un archivo CSV básico, cada fila tiene dos o tres campos: una pregunta, una respuesta y un URI de origen opcional que apunta a un documento con más información.

El contenido del archivo debe seguir el Formato común RFC 4180 y tipo MIME para archivos de valores separados por comas (CSV).

A continuación se muestra un archivo de preguntas frecuentes en formato CSV básico.

```
How many free clinics are in Spokane WA?, 13, https://s3.region.company.com/bucket-
name/directory/faq.csv
How many free clinics are there in Mountain View Missouri?, 7, https://
s3.region.company.com/bucket-name/directory/faq.csv
```

Archivo CSV personalizado

Utilice un archivo CSV personalizado cuando desee agregar campos/atributos personalizados a sus preguntas frecuentes. Para un archivo CSV personalizado, utilice una fila de encabezado del archivo CSV para definir los atributos adicionales.

El archivo CSV debe contener los dos campos obligatorios siguientes:

- `_question`: la pregunta frecuente
- `_answer`: la respuesta a la pregunta frecuente

El archivo puede contener tanto campos Amazon Kendra reservados como campos personalizados. A continuación se muestra un ejemplo de un archivo CSV personalizado.

```
_question,_answer,_last_updated_at,custom_string
How many free clinics are in Spokane WA?, 13, 2012-03-25T12:30:10+01:00, Note: Some
 free clinics require you to meet certain criteria in order to use their services
How many free clinics are there in Mountain View Missouri?, 7,
 2012-03-25T12:30:10+01:00, Note: Some free clinics require you to meet certain
 criteria in order to use their services
```

El contenido del archivo personalizado debe seguir el Formato común RFC 4180 y tipo MIME para archivos de valores separados por comas (CSV).

A continuación se enumeran los tipos de campos personalizados:

- Fecha: valores de fecha y hora con codificación ISO 8601.

  Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en la zona horaria de Europa Central.
- Largo: números, como `1234`.
- Cadena: valores de cadena. Si la cadena contiene comas, incluya todo el valor entre comillas dobles (") (por ejemplo, `"custom attribute, and more"`).
- Lista de cadenas: una lista de valores de cadenas. Enumere los valores de una lista separada por comas incluidos entre comillas (") (por ejemplo, `"item1, item2, item3"`). Si la lista contiene solo una entrada, puede omitir las comillas (por ejemplo, `item1`).

Un archivo CSV personalizado puede contener campos de control de acceso de usuarios. Puede utilizar estos campos para limitar el acceso a las preguntas frecuentes a determinados usuarios y grupos. Para filtrar el contexto de usuario, el usuario debe proporcionar en la consulta información de usuario y grupo. De lo contrario, se devuelven todas las preguntas frecuentes pertinentes. Para obtener más información, consulte Filtrado de contexto de usuario.

A continuación se enumeran los filtros de contexto de usuario para las preguntas frecuentes:

- `_acl_user_allow`: los usuarios de la lista de permitidos pueden ver las preguntas frecuentes en la respuesta a la consulta. Las preguntas frecuentes no se devuelven a otros usuarios.
- `_acl_user_deny`: los usuarios de la lista de denegados pueden ver las preguntas frecuentes en la respuesta a la consulta. Las preguntas frecuentes se devuelven a todos los demás usuarios cuando sea relevante para la consulta.
- `_acl_group_allow`: los usuarios que son miembros de un grupo permitido pueden ver las preguntas frecuentes en la respuesta a la consulta. Las preguntas frecuentes no se devuelven a los usuarios que sean miembros de otro grupo.
- `_acl_group_deny`: los usuarios que son miembros de un grupo denegado no pueden ver las preguntas frecuentes en la respuesta a la consulta. Las preguntas frecuentes se devuelven a todos los demás grupos cuando sea relevante para la consulta.

Proporcione los valores de las listas de permitir y denegar listas separadas por comas entre comillas (por ejemplo, `"user1,user2,user3"`). Puede incluir un usuario o un grupo en una lista de permisos o en una lista de denegación, pero no en ambas si el mismo usuario tiene un permiso individual pero forma parte de un grupo denegado. Si incluye un usuario o grupo en ambos, recibirá un error.

A continuación se muestra un ejemplo de un archivo CSV personalizado.

```
_question, _answer, _acl_user_allow, _acl_user_deny, _acl_group_allow, _acl_group_deny
How many free clinics are in Spokane WA?, 13, "userID6201,userID7552",
 "userID1001,userID2020", groupBasicPlusRate, groupPremiumRate
```

Archivo JSON

Puede utilizar un archivo JSON para proporcionar preguntas, respuestas y campos para su índice. Puede añadir cualquiera de los campos Amazon Kendra reservados o campos personalizados a las preguntas frecuentes.

A continuación se muestra el esquema para el archivo JSON.

```
{
    "SchemaVersion": 1,
    "FaqDocuments": [
        {
            "Question": string,
            "Answer": string,
            "Attributes": {
                string: object
                additional attributes
            },
            "AccessControlList": [
               {
                   "Name": string,
                   "Type": enum( "GROUP" | "USER" ),
                   "Access": enum( "ALLOW" | "DENY" )
               },
               additional user context
            ]
        },
        additional FAQ documents
    ]
```

```
}
```

En el siguiente ejemplo, el archivo JSON muestra dos documentos de preguntas frecuentes. Uno de ellos solo tiene la pregunta y la respuesta requeridas. El otro documento también incluye información adicional sobre el contexto del usuario y los campos o sobre el control de acceso.

```
{
    "SchemaVersion": 1,
    "FaqDocuments": [
        {
            "Question": "How many free clinics are in Spokane WA?",
            "Answer": "13"
        },
        {
            "Question": "How many free clinics are there in Mountain View Missouri?",
            "Answer": "7",
            "Attributes": {
                "_source_uri": "https://s3.region.company.com/bucket-name/directory/
faq.csv",
                "_category": "Charitable Clinics"
            },
            "AccessControlList": [
               {
                   "Name": "user@amazon.com",
                   "Type": "USER",
                   "Access": "ALLOW"
               },
               {
                   "Name": "Admin",
                   "Type": "GROUP",
                   "Access": "ALLOW"
               }
            ]
        }
    ]
}
```

A continuación se enumeran los tipos de campos personalizados:

- Fecha: valor de cadena JSON con valores de fecha y hora con codificación ISO 8601. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en la zona horaria de Europa Central.

- Largo: valor numérico de JSON, como `1234`.
- Cadena: valor de cadena de JSON (por ejemplo, `"custom attribute"`).
- Lista de cadenas: matriz de valores de cadenas de JSON (por ejemplo, `["item1,item2,item3"]`).

Un archivo JSON puede contener campos de control de acceso de usuarios. Puede utilizar estos campos para limitar el acceso a las preguntas frecuentes a determinados usuarios y grupos. Para filtrar el contexto de usuario, el usuario debe proporcionar en la consulta información de usuario y grupo. De lo contrario, se devuelven todas las preguntas frecuentes pertinentes. Para obtener más información, consulte Filtrado de contexto de usuario.

Puede incluir un usuario o un grupo en una lista de permisos o en una lista de denegación, pero no en ambas si el mismo usuario tiene un permiso individual pero forma parte de un grupo denegado. Si incluye un usuario o grupo en ambos, recibirá un error.

A continuación se muestra un ejemplo de cómo se incluye el control de acceso de usuarios en una pregunta frecuente de JSON.

```
"AccessControlList": [
                {
                    "Name": "group or user name",
                    "Type": "GROUP | USER",
                    "Access": "ALLOW | DENY"
                },
                additional user context
            ]
```

Uso del archivo de preguntas frecuentes

Después de almacenar el archivo de entrada de preguntas frecuentes en un bucket de S3, utilice la consola o la API `CreateFaq` para incluir las preguntas y respuestas en el índice. Si quiere actualizar una pregunta frecuente, elimínela y vuelva a crearla. Utilice la API `DeleteFaq` para eliminar una pregunta frecuente.

Debe proporcionar un IAM rol que tenga acceso al depósito de S3 que contiene sus archivos fuente. Puede especificar el rol en la consola o en el parámetro `RoleArn`. A continuación se muestra un ejemplo de adición de un archivo de preguntas frecuentes a un índice.

Python

```
import boto3

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the IAM role ARN required to index documents in an S3 bucket
role_arn = "arn:aws:iam::${accountId}:role/${roleName}"

# Provide the S3 bucket path information to the FAQ file
faq_path = {
    "Bucket": "bucket-name",
    "Key": "FreeClinicsUSA.csv"
}

response = kendra.create_faq(
    S3Path =  faq_path,
    Name = "FreeClinicsUSA",
    IndexId = index_id,
    RoleArn = role_arn
)

print(response)
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateFaqRequest;
import software.amazon.awssdk.services.kendra.model.CreateFaqResponse;
import software.amazon.awssdk.services.kendra.model.S3Path;

public class AddFaqExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indexId = "yourIndexId";
        String roleArn = "your role for accessing S3 files";

        CreateFaqRequest createFaqRequest = CreateFaqRequest
```

```
            .builder()
            .indexId(indexId)
            .name("FreeClinicsUSA")
            .roleArn(roleArn)
            .s3Path(
                S3Path
                    .builder()
                    .bucket("an-aws-kendra-test-bucket")
                    .key("FreeClinicsUSA.csv")
                    .build())
            .build();

        CreateFaqResponse response = kendra.createFaq(createFaqRequest);

        System.out.println(String.format("The result of creating FAQ: %s",
 response));

    }
}
```

Archivos de preguntas frecuentes en idiomas distintos del inglés

Puede indexar una pregunta frecuente en un idioma compatible. Amazon Kendra indexa las preguntas frecuentes en inglés de forma predeterminada si no especificas un idioma. El código de idioma se especifica al llamar a la CreateFaqoperación o se puede incluir el código de idioma de una pregunta frecuente en los metadatos de las preguntas frecuentes como un campo. Si una pregunta frecuente no contiene un código de idioma en sus metadatos especificado en un campo de metadatos, las preguntas frecuentes se indexan utilizando el código de idioma especificado al llamar a la operación `CreateFAQ`. Para indexar un documento de preguntas frecuentes en un idioma admitido en la consola, vaya a FAQs (Preguntas frecuentes) y seleccione Add FAQ (Añadir pregunta frecuente). Elija un idioma en el menú desplegable Language (Idioma).

Creación de campos de documento personalizados

Puede crear atributos o campos personalizados para sus documentos en su índice de Amazon Kendra. Por ejemplo, puede crear un campo o atributo personalizado denominado “Departamento” con los valores de “RR. HH.”, “Ventas” y “Fabricación”. Si asigna estos campos o atributos personalizados a su índice de Amazon Kendra, puede usarlos para filtrar los resultados de la búsqueda e incluir documentos por el atributo del departamento «Recursos humanos», por ejemplo.

Para poder utilizar un campo o atributo personalizado, primero debe crear el campo en el índice. Utilice la consola para editar las asignaciones de campos de la fuente de datos y añadir un campo personalizado o utilice la UpdateIndexAPI para crear el campo de índice. No puede cambiar el tipo de dato del campo una vez que este se ha creado.

Para la mayoría de orígenes de datos, asignará los campos del origen de datos externo a los campos correspondientes en Amazon Kendra. Para obtener más información, consulte Asignación de campos de origen de datos. Para los orígenes de datos de S3, puede crear atributos o campos personalizados mediante un archivo de metadatos JSON.

Puede crear hasta 500 campos o atributos personalizados.

También puede usar campos Amazon Kendra reservados o comunes. Para obtener más información, consulte Atributos o campos del documento.

Temas

- Actualización de campos de documentos personalizados

Actualización de campos de documentos personalizados

Con la API `UpdateIndex`, se agregan campos o atributos personalizados mediante el parámetro `DocumentMetadataConfigurationUpdates`.

En el siguiente ejemplo de JSON se utiliza `DocumentMetadataConfigurationUpdates` para agregar al índice un campo denominado “Department”.

```
"DocumentmetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE"
   }
]
```

En las siguientes secciones se incluyen ejemplos para añadir atributos o campos personalizados mediante BatchPutDocumenty para una fuente de datos de Amazon S3.

Temas

- Añadir atributos o campos personalizados con la BatchPutDocument API

- Adición de atributos o campos personalizados a un origen de datos de Amazon S3

Añadir atributos o campos personalizados con la BatchPutDocument API

Cuando utilizas la BatchPutDocumentAPI para añadir un documento a tu índice, especificas campos o atributos personalizados como parte de ella`Attributes`. Puede añadir varios campos o atributos al llamar a la API. Puede crear hasta 500 campos o atributos personalizados. El siguiente ejemplo es un campo o atributo personalizado que agrega “Departamento” a un documento.

```
"Attributes":
    {
        "Department": "HR",
        "_category": "Vacation policy"
    }
```

Adición de atributos o campos personalizados a un origen de datos de Amazon S3

Cuando se utilice un bucket de S3 como origen de datos para el índice, se agregan metadatos a los documentos con archivos de metadatos complementarios. Los archivos JSON de metadatos se colocan en una estructura de directorios paralela a los documentos. Para obtener más información, consulte Metadatos de documentos de S3.

Los campos o atributos personalizados se especifican en la estructura JSON de `Attributes`. Puede crear hasta 500 campos o atributos personalizados. Por ejemplo, en el siguiente ejemplo se utiliza `Attributes` para definir tres campos o atributos personalizados y un campo reservado.

```
"Attributes": {
        "brand": "Amazon Basics",
        "price": 1595,
        "_category": "sports",
        "subcategories": ["outdoors", "electronics"]
    }
```

En los siguientes pasos, se explica cómo añadir atributos personalizados a una fuente de datos de Amazon S3.

Temas

- Paso 1: Crear un índice de Amazon Kendra
- Paso 2: Actualice el índice para añadir campos de documentos personalizados

- Paso 3: Cree una fuente de datos de Amazon S3 y asigne los campos de la fuente de datos a atributos personalizados

Paso 1: Crear un índice de Amazon Kendra

Sigue los pasos Creación de un índice que se indican para crear tu índice de Amazon Kendra.

Paso 2: Actualice el índice para añadir campos de documentos personalizados

Después de crear un índice, se le añaden campos. El siguiente procedimiento muestra cómo agregar campos a un índice mediante la consola y la CLI.

Console

Para crear campos de índice

1. Asegúrese de haber creado un índice.
2. A continuación, en el menú de navegación de la izquierda, en Administración de datos, selecciona Definición de facetas.
3. En la guía de configuración de campos de índice, en Campos de índice, selecciona Añadir campo para añadir campos personalizados.
4. En el cuadro de diálogo Agregar campo de índice, haga lo siguiente:
   - Nombre de campo: agrega un nombre de campo.
   - Tipo de datos: seleccione el tipo de datos, ya sea cadena, lista de cadenas o fecha.
   - Tipos de uso: seleccione los tipos de uso, ya sean facetables, buscables, visualizables y ordenables.

     A continuación, selecciona Añadir.

   Repita el último paso para cualquier otro campo que desee mapear.

CLI

```
aws kendra update-index  \
--region $region \
--endpoint-url $endpoint \
--application-id $applicationId \
```

```
--index-id $indexId  \
--document-metadata-configuration-updates \
"[
    {
        "Name": "string",
        "Type": "STRING_VALUE"|"STRING_LIST_VALUE"|"LONG_VALUE"|"DATE_VALUE",
        "Relevance": {
            "Freshness": true|false,
            "Importance": integer,
            "Duration": "string",
            "RankOrder": "ASCENDING"|"DESCENDING",
            "ValueImportanceMap": {"string": integer
            ...}
    },
    "Search": {
        "Facetable": true|false,
        "Searchable": true|false,
        "Displayable": true|false,
        "Sortable": true|false
        }
    }
...
]"
```

Paso 3: Cree una fuente de datos de Amazon S3 y asigne los campos de la fuente de datos a atributos personalizados

Para crear una fuente de datos de Amazon S3 y asignarle campos, siga las instrucciones que se indican enAmazon S3.

Si utiliza la API, utilice el `fieldMappings` atributo que aparece debajo `configuration` cuando utilice la CreateDataSourceAPI.

Para obtener una descripción general de cómo se mapean los campos de las fuentes de datos, consulteAsignación de campos de origen de datos.

Control de acceso de usuarios a los documentos con tokens

Puede controlar qué usuarios o grupos pueden acceder a determinados documentos de su índice o ver determinados documentos en sus resultados de búsqueda. Esto se denomina filtrado por contexto de usuario. Es un tipo de búsqueda personalizada con la ventaja de controlar el acceso a

los documentos. Por ejemplo, no todos los equipos que buscan información en el portal corporativo deben acceder a los documentos de alto secreto de la empresa, ni estos documentos son relevantes para todos los usuarios. Solo los usuarios o grupos de equipos específicos que tengan acceso a documentos de alto secreto deberían ver estos documentos en sus resultados de búsqueda.

Amazon Kendra admite el control de acceso de usuarios basado en tokens con los siguientes tipos de tokens:

- OpenID
- JWT con un secreto compartido
- JWT con una clave pública
- JSON

Amazon Kendra ofrece una búsqueda empresarial altamente segura para sus aplicaciones de búsqueda. Los resultados de la búsqueda reflejan el modelo de seguridad de su organización. Los clientes son responsables de autenticar y autorizar a los usuarios a acceder a su aplicación de búsqueda. En el momento de la búsqueda, el servicio de Amazon Kendra filtra los resultados de la búsqueda en función del ID de usuario proporcionado por la aplicación de búsqueda del cliente y de las listas de control de acceso a los documentos (ACL) recopiladas por los conectores de Amazon Kendra durante el rastreo o la indexación. Los resultados de la búsqueda muestran direcciones URL que apuntan a los repositorios de documentos originales, además de breves fragmentos. El acceso al documento completo sigue siendo controlado por el repositorio original.

Temas

- Uso de OpenID
- Uso de un JSON Web Token (JWT) con un secreto compartido
- Uso de un JSON Web Token (JWT) con una clave pública
- Uso de JSON

Uso de OpenID

Para configurar un Amazon Kendra índice para usar un token de OpenID para el control de acceso, necesita la URL JWKS (JSON Web Key Set) del proveedor de OpenID. En la mayoría de los casos, la URL JWKS tiene el siguiente formato (si sigue la detección de OpenID) `https://domain-name/.well_known/jwks.json`.

Los siguientes ejemplos muestran cómo usar un token de OpenID para el control de acceso de los usuarios al crear un índice.

Console

1. Elija Crear índice para empezar a crear un índice nuevo.
2. En la página Especificar detalles de índice, proporcione a su índice un nombre y una descripción.
3. Para el rol de IAM , seleccione un rol o seleccione Crear un nuevo rol y especifique un nombre para crear un nuevo rol. La función de IAM tendrá el prefijo "-». AmazonKendra
4. No cambie los demás valores predeterminados. Elija Siguiente.
5. En la página Configurar el control de acceso de los usuarios, en Configuración de control de acceso, seleccione Sí para utilizar los tokens para el control de acceso.
6. En Configuración de token, seleccione OpenID como Tipo de token.
7. Especifique una URL de clave de firma. La URL debe apuntar a un conjunto de claves web JSON.
8. Opcional En Configuración avanzada:
    a. Especifique un Nombre de usuario para usarlo en la verificación de la ACL.
    b. Especifique uno o más Grupos para usarlos en la verificación de la ACL.
    c. Especifique el Emisor que validará el emisor del token.
    d. Especifique el o los ID de cliente. Debe especificar una expresión regular que coincida con el público del JWT.
9. En la página Detalles de aprovisionamiento, elija la Edición para desarrolladores.
10. Elija Crear para crear el índice.
11. Espere a que se cree el índice. Amazon Kendra aprovisiona el hardware para su índice. Esta operación puede llevar algún tiempo.

CLI

Para crear un índice AWS CLI con un archivo de entrada JSON, cree primero un archivo JSON con los parámetros que desee:

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
```

```
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",
    "UserTokenConfigurations": [
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "URL",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "URL": "https://example.com/.well-known/jwks.json"
            }
        }
    ],
    "UserContextPolicy": "USER_TOKEN"
}
```

Puede anular los nombres de campo de usuario y grupo predeterminados. El valor predeterminado de `UserNameAttributeField` es “user”. El valor predeterminado de `GroupAttributeField` es “groups”.

A continuación, llame a `create-index` usando el archivo de entrada. Por ejemplo, si el nombre de su archivo JSON es `create-index-openid.json`, puede usar lo siguiente:

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

Python

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account-id:role:/my-role',
    UserTokenConfigurations=[
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "URL",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "URL": "https://example.com/.well-known/jwks.json"
            }
        }
    ],
```

```
    UserContextPolicy='USER_TOKEN'
)
```

Uso de un JSON Web Token (JWT) con un secreto compartido

Los siguientes ejemplos muestran cómo usar el token web JSON (JWT) con un token secreto compartido para controlar el acceso de los usuarios al crear un índice.

Console

1. Elija Crear índice para empezar a crear un índice nuevo.
2. En la página Especificar detalles de índice, proporcione a su índice un nombre y una descripción.
3. Para el rol de IAM, seleccione un rol o seleccione Crear un nuevo rol y especifique un nombre para crear un nuevo rol. El IAM rol tendrá el prefijo "AmazonKendra-».
4. No cambie los demás valores predeterminados. Elija Siguiente.
5. En la página Configurar el control de acceso de los usuarios, en Configuración de control de acceso, seleccione Sí para utilizar los tokens para el control de acceso.
6. En Configuración de token, seleccione JWT con secreto compartido como Tipo de token.
7. En Parámetros para firmar el secreto compartido, elija el Tipo de secreto. Puede usar un secreto compartido existente de AWS Secrets Manager o crear uno nuevo.

   Para crear un secreto compartido nuevo, seleccione Nuevo y, a continuación, siga estos pasos:

   a. En Nuevo AWS Secrets Manager secreto, especifique un nombre secreto. El prefijo `AmazonKendra-` se agregará al guardar la clave pública.
   b. Especifique un ID de clave. El ID de clave es una sugerencia que indica la clave que se ha utilizado para proteger la firma web JSON del token.
   c. Elija el Algoritmo de firma del token. Este es el algoritmo criptográfico que se utiliza para proteger el token de ID. Para obtener más información sobre RSA, consulte Criptografía de RSA.
   d. Especifique un Secreto compartido introduciendo un secreto con codificación de URL en base64. También puede seleccionar Generar secreto para que se genere un secreto automáticamente. Debe asegurarse de que el secreto tenga codificación de URL en base64.

e. (Opcional) Especifique cuándo es válido el secreto compartido. Puede especificar la fecha y la hora desde o hasta las que un secreto es válido o ambas. El secreto será válido en el intervalo especificado.

f. Seleccione Guardar secreto para guardar el nuevo secreto.

8. (Opcional) En Configuración avanzada:

a. Especifique un Nombre de usuario para usarlo en la verificación de la ACL.

b. Especifique uno o más Grupos para usarlos en la verificación de la ACL.

c. Especifique el Emisor que validará el emisor del token.

d. Especifique el o los ID de reclamación. Debe especificar una expresión regular que coincida con el público del JWT.

9. En la página Detalles de aprovisionamiento, elija la Edición para desarrolladores.

10. Elija Crear para crear el índice.

11. Espere a que se cree el índice. Amazon Kendra aprovisiona el hardware para su índice. Esta operación puede llevar algún tiempo.

CLI

Puedes usar el token JWT con un secreto compartido en su interior. AWS Secrets Manager El secreto debe tener codificación de URL en base64. Necesita el Secrets Manager ARN y su Amazon Kendra rol debe tener acceso al `GetSecretValue` Secrets Manager recurso. Si va a cifrar el Secrets Manager recurso con AWS KMS, el rol también debe tener acceso a la acción de descifrado.

Para crear un índice con un archivo de AWS CLI entrada JSON, cree primero un archivo JSON con los parámetros que desee:

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",
    "UserTokenConfigurations": [
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "SECRET_MANAGER",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
```

```
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account
 id:secret:/my-user-context-secret
            }
        }
    ],
    "UserContextPolicy": "USER_TOKEN"
}
```

Puede anular los nombres de campo de usuario y grupo predeterminados. El valor predeterminado de `UserNameAttributeField` es “user”. El valor predeterminado de `GroupAttributeField` es “groups”.

A continuación, llame a `create-index` usando el archivo de entrada. Por ejemplo, si el nombre de su archivo JSON es `create-index-openid.json`, puede usar lo siguiente:

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

El secreto debe tener el siguiente formato AWS Secrets Manager:

```
{
  "keys": [
    {
      "kid": "key_id",
      "alg": "HS256|HS384|HS512",
      "kty": "OCT",
      "use": "sig", //this value can be sig only for now
      "k": "secret",
      "nbf":"ISO1806 date format"
      "exp":"ISO1806 date format"
    }
  ]
}
```

Para obtener más información acerca de JWT, consulte jwt.io.

Python

Puedes usar el token JWT con un secreto compartido en su interior. AWS Secrets Manager El secreto debe tener codificación de URL en base64. Necesita el Secrets Manager ARN y su Amazon Kendra rol debe tener acceso al `GetSecretValue` Secrets Manager recurso. Si va a

cifrar el Secrets Manager recurso con AWS KMS, el rol también debe tener acceso a la acción de descifrado.

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account-id:role:/my-role',
    UserTokenConfigurations=[
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "URL",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account
 id:secret:/my-user-context-secret"
            }
        }
    ],
    UserContextPolicy='USER_TOKEN'
)
```

Uso de un JSON Web Token (JWT) con una clave pública

Los siguientes ejemplos muestran cómo usar el token web JSON (JWT) con una clave pública para controlar el acceso de los usuarios al crear un índice. Para obtener más información acerca de JWT, consulte jwt.io.

Console

1. Elija Crear índice para empezar a crear un índice nuevo.
2. En la página Especificar detalles de índice, proporcione a su índice un nombre y una descripción.
3. Para el rol de IAM, seleccione un rol o seleccione Crear un nuevo rol y especifique un nombre para crear un nuevo rol. El IAM rol tendrá el prefijo "AmazonKendra-».
4. No cambie los demás valores predeterminados. Elija Siguiente.
5. En la página Configurar el control de acceso de los usuarios, en Configuración de control de acceso, seleccione Sí para utilizar los tokens para el control de acceso.

6. En Configuración de token, seleccione JWT con clave pública como Tipo de token.
7. En Parámetros para firmar la clave pública, elija el Tipo de secreto. Puede usar un secreto existente de AWS Secrets Manager o crear uno nuevo.

   Para crear un secreto nuevo, seleccione Nuevo y, a continuación, siga estos pasos:

   a. En Nuevo AWS Secrets Manager secreto, especifique un nombre secreto. El prefijo `AmazonKendra-` se agregará al guardar la clave pública.
   b. Especifique un ID de clave. El ID de clave es una sugerencia que indica la clave que se ha utilizado para proteger la firma web JSON del token.
   c. Elija el Algoritmo de firma del token. Este es el algoritmo criptográfico que se utiliza para proteger el token de ID. Para obtener más información sobre RSA, consulte Criptografía de RSA.
   d. En Atributos del certificado, especifique una Cadena de certificados opcional. La cadena de certificados se compone de una lista de certificados. Comienza con el certificado de un servidor y termina con el certificado raíz.
   e. Opcional: especifique la Huella digital. Debe ser un hash de un certificado, calculado sobre todos los datos del certificado y su firma.
   f. Especifique el Exponente. Este es el valor del exponente de la clave pública RSA. Se representa como un valor codificado en Base64urlUInt.
   g. Especifique el Módulo. Este es el valor del exponente de la clave pública RSA. Se representa como un valor codificado en Base64urlUInt.
   h. Seleccione Guardar clave para guardar la nueva clave.
8. Opcional En Configuración avanzada:

   a. Especifique un Nombre de usuario para usarlo en la verificación de la ACL.
   b. Especifique uno o más Grupos para usarlos en la verificación de la ACL.
   c. Especifique el Emisor que validará el emisor del token.
   d. Especifique el o los ID de cliente. Debe especificar una expresión regular que coincida con el público del JWT.
9. En la página Detalles de aprovisionamiento, elija la Edición para desarrolladores.
10. Elija Crear para crear el índice.
11. Espere a que se cree el índice. Amazon Kendra aprovisiona el hardware para su índice. Esta operación puede llevar algún tiempo.

CLI

Puede usar JWT con una clave pública dentro de un AWS Secrets Manager. Necesita el Secrets Manager ARN y su Amazon Kendra rol debe tener acceso al `GetSecretValue` Secrets Manager recurso. Si va a cifrar el Secrets Manager recurso con AWS KMS, el rol también debe tener acceso a la acción de descifrado.

Para crear un índice con un archivo de AWS CLI entrada JSON, cree primero un archivo JSON con los parámetros que desee:

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
    "RoleArn": "arn:aws:iam::account id:role:/my-role",
    "UserTokenConfigurationList": [
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "SECRET_MANAGER",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account
 id:secret:/my-user-context-secret
            }
        }
    ],    "UserContextPolicy": "USER_TOKEN"
}
```

Puede anular los nombres de campo de usuario y grupo predeterminados. El valor predeterminado de `UserNameAttributeField` es “user”. El valor predeterminado de `GroupAttributeField` es “groups”.

A continuación, llame a `create-index` usando el archivo de entrada. Por ejemplo, si el nombre de su archivo JSON es `create-index-openid.json`, puede usar lo siguiente:

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

El secreto debe tener el siguiente formato Secrets Manager:

```
{
```

```
  "keys": [
    {
      "alg": "RS256|RS384|RS512",
      "kty": "RSA", //this can be RSA only for now
      "use": "sig", //this value can be sig only for now
      "n": "modulus of standard pem",
      "e": "exponent of standard pem",
      "kid": "key_id",
      "x5t": "certificate thumprint for x.509 cert",
      "x5c": [
        "certificate chain"
      ]
    }
  ]
}
```

Para obtener más información acerca de JWT, consulte jwt.io.

Python

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account id:role:/my-role',
    UserTokenConfigurationList=[
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "URL",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account
 id:secret:/my-user-context-secret"
            }
        }
    ],
    UserContextPolicy='USER_TOKEN'
)
```

Uso de JSON

Los siguientes ejemplos muestran cómo usar JSON para el control de acceso de los usuarios al crear un índice.

> **Warning**
>
> El token JSON es una carga no validada. Solo debe usarse cuando las solicitudes a Amazon Kendra provienen de un servidor de confianza y nunca de un navegador.

Console

1. Elija Crear índice para empezar a crear un índice nuevo.
2. En la página Especificar detalles de índice, proporcione a su índice un nombre y una descripción.
3. Para el rol de IAM , seleccione un rol o seleccione Crear un nuevo rol y especifique un nombre para crear un nuevo rol. El IAM rol tendrá el prefijo "AmazonKendra-».
4. No cambie los demás valores predeterminados. Elija Siguiente.
5. En la página Configurar el control de acceso de los usuarios, en Configuración de control de acceso, seleccione Sí para utilizar los tokens para el control de acceso.
6. En Configuración de token, seleccione JSON como Tipo de token.
7. Especifique un Nombre de usuario para usarlo en la verificación de la ACL.
8. Especifique uno o más Grupos para usarlos en la verificación de la ACL.
9. Elija Siguiente.
10. En la página Detalles de aprovisionamiento, elija la Edición para desarrolladores.
11. Elija Crear para crear el índice.
12. Espere a que se cree su índice. Amazon Kendra aprovisiona el hardware para su índice. Esta operación puede llevar algún tiempo.

CLI

Para crear un índice AWS CLI con un archivo de entrada JSON, cree primero un archivo JSON con los parámetros que desee:

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",
    "UserTokenConfigurations": [
        {
            "JsonTokenTypeConfiguration": {
                "UserNameAttributeField": "user",
                "GroupAttributeField": "group"
            }
        }
    ],
    "UserContextPolicy": "USER_TOKEN"
}
```

A continuación, llame a `create-index` usando el archivo de entrada. Por ejemplo, si el nombre de su archivo JSON es `create-index-openid.json`, puede usar lo siguiente:

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

Si no utilizas Open ID AWS IAM Identity Center, puedes enviarnos el token en formato JSON. Si lo hace, debe especificar qué campo del token JSON contiene el nombre de usuario y qué campo contiene los grupos. Los valores del campo de los grupos deben ser una matriz de cadenas JSON. Por ejemplo, si utiliza SAML, su token debería ser similar al siguiente:

```
{
     "username" : "user1",
     "groups": [
        "group1",
        "group2"
     ]
}
```

La `TokenConfiguration` especificaría el nombre de usuario y los nombres del campo de los grupos:

```
{
    "UserNameAttributeField":"username",
    "GroupAttributeField":"groups"
}
```

Python

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account-id:role:/my-role',
    UserTokenConfigurations=[
        {
            "JwtTokenTypeConfiguration": {
                "UserNameAttributeField": "user",
                "GroupAttributeField": "group",
            }
        }
    ],
    UserContextPolicy='USER_TOKEN'
)
```

Creación de un conector de origen de datos

Puede crear un conector de fuente de datos para conectarse Amazon Kendra a sus documentos e indexarlos. Amazon Kendra puede conectarse a Microsoft SharePoint, Google Drive y muchos otros proveedores. Al crear un conector de fuente de datos, proporciona Amazon Kendra la información de configuración necesaria para conectarse al repositorio de fuentes. A diferencia de añadir documentos directamente a un índice, puede escanear periódicamente el origen de datos para actualizar el índice.

Por ejemplo, supongamos que tiene un repositorio de documentos fiscales almacenado en un Amazon S3 depósito. Ocasionalmente, los documentos existentes se modifican y se añaden nuevos documentos al repositorio. Si agrega el repositorio Amazon Kendra como fuente de datos, puede mantener el índice actualizado configurando sincronizaciones periódicas entre la fuente de datos y el índice.

Puede optar por actualizar un índice manualmente mediante la consola o la StartDataSourceSyncJobAPI. De lo contrario, puede configurar una programación para actualizar un índice y sincronizarlo con su origen de datos.

Un índice puede tener más de un origen de datos. Cada origen de datos puede tener su propia programación de actualizaciones. Por ejemplo, puede actualizar el índice de sus documentos de trabajo a diario, o incluso cada hora, y actualizar los documentos archivados manualmente cada vez que cambie el archivo.

Si desea modificar los metadatos o atributos del documento y el contenido durante el proceso de ingesta de documentos, consulte Custom Document Enrichment de Amazon Kendra.

> **Note**
>
> Cada ID de documento debe ser único por índice. No se puede crear un origen de datos para indexar los documentos con sus ID exclusivos y, a continuación, utilizar la API `BatchPutDocument` para indexar los mismos documentos o viceversa. No se puede crear un origen de datos para indexar los documentos con sus ID exclusivos y, a continuación, utilizar la API `BatchPutDocument` para indexar los mismos documentos o viceversa. El uso de `BatchDeleteDocument` las API `BatchPutDocument` y en combinación con un conector de fuente de Amazon Kendra datos para el mismo conjunto de documentos podría

provocar incoherencias en los datos. En su lugar, recomendamos utilizar el Amazon Kendra conector de origen de datos personalizado.

> **Note**
>
> Los archivos añadidos al índice deben estar en un flujo de bytes codificado en UTF-8. Para obtener más información sobre los documentos incluidos Amazon Kendra, consulte Documentos.

Establecimiento de un programa de actualizaciones

Configure el origen de datos para que se actualice periódicamente con la consola o mediante el parámetro `Schedule` al crear o actualizar un origen de datos. El contenido del parámetro es una cadena que contiene una cadena de programación en formato `cron` o una cadena vacía para indicar que el índice se actualiza a petición. Para conocer el formato de una expresión cron, consulte Programar expresiones para reglas en la Guía del Amazon CloudWatch Events usuario. Amazon Kendra solo admite expresiones cron. No admite expresiones de frecuencia.

Configuración del idioma

Puede indexar todos los documentos en un origen de datos en un idioma compatible. Al llamar CreateDataSource, debe especificar el código de idioma de todos los documentos en la fuente de datos. Si un documento no contiene un código de idioma especificado en un campo de metadatos, el documento se indexa utilizando el código de idioma especificado para todos los documentos en el nivel de origen de datos. Si no especifica un idioma, Amazon Kendra indexa los documentos de un origen de datos en inglés de forma predeterminada. Para obtener más información acerca de los idiomas admitidos, incluidos sus códigos, consulte Adición de documentos en idiomas distintos del inglés.

Todos los documentos de un origen de datos en un idioma compatible se indexan mediante la consola. Vaya a Orígenes de datos y edite su origen de datos o a Agregar origen de datos si va a agregar un nuevo origen de datos. En la página Especificar detalles del origen de datos, seleccione un idioma en el menú desplegable Idioma. Seleccione Actualizar o continúe introduciendo la información de configuración para conectarse a su origen de datos.

Conectores de origen de datos

En esta sección, se muestra cómo conectarse Amazon Kendra a bases de datos y repositorios de fuentes de datos compatibles mediante Amazon Kendra las API AWS Management Console y las Amazon Kendra API.

Temas

- Esquemas de plantillas de origen de datos
- Adobe Experience Manager
- Alfresco
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (Windows)
- Amazon FSx (NetApp DISPONIBLE)
- Amazon RDS/Aurora
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon S3
- Amazon Kendra Rastreador web
- Amazon WorkDocs
- Box (Cuadro)
- Confluence
- Conector de orígenes de datos personalizados
- Dropbox
- Drupal
- GitHub
- Gmail
- Google Drive
- IBM DB2
- Jira

- Microsoft Exchange
- Microsoft OneDrive
- Microsoft SharePoint
- Microsoft SQL Server
- Microsoft Teams
- Microsoft Yammer
- MySQL
- Oracle Database
- PostgreSQL
- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

Esquemas de plantillas de origen de datos

Los siguientes son esquemas de plantillas para orígenes de datos en los que se admiten plantillas.

Temas

- Esquema de plantilla de Adobe Experience Manager
- Amazon FSx Esquema de plantillas (Windows)
- Amazon FSx Esquema de plantillas (NetApp ONTAP)
- Esquema de plantilla de Alfresco
- Aurora Esquema de plantillas (MySQL)
- Aurora Esquema de plantillas (PostgreSQL)
- Amazon RDS Esquema de plantillas (Microsoft SQL Server)
- Amazon RDS Esquema de plantillas (MySQL)
- Amazon RDS Esquema de plantillas (Oracle)
- Amazon RDS Esquema de plantillas (PostgreSQL)
- Amazon S3 esquema de plantilla
- Amazon Kendra Esquema de plantillas de Web Crawler

- Esquema de plantilla de Confluence
- Esquema de plantilla de Dropbox
- Esquema de plantilla de Drupal
- GitHub esquema de plantilla
- Esquema de plantilla de Gmail
- Esquema de plantilla de Google Drive
- Esquema de plantilla de IBM DB2
- Esquema de plantilla de Microsoft Exchange
- Esquema OneDrive de plantillas de Microsoft
- Esquema SharePoint de plantillas de Microsoft
- Esquema de plantilla de Microsoft SQL Server
- Esquema de plantilla de Microsoft Teams
- Esquema de plantilla de Microsoft Yammer
- Esquema de plantilla de MySQL
- Esquema de plantilla de Oracle Database
- Esquema de plantilla de PostgreSQL
- Esquema de plantilla de Salesforce
- ServiceNow esquema de plantilla
- Esquema de plantillas de Slack
- Esquema de plantilla de Zendesk

Esquema de plantilla de Adobe Experience Manager

Incluye un JSON que contiene el esquema del origen de datos como parte del objeto TemplateConfiguration. Debe proporcionar la URL del host de Adobe Experience Manager, el tipo de autenticación y si utiliza Adobe Experience Manager (AEM) as a Cloud Service o AEM On-Premise como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Además, especifique el tipo de origen de datos como `AEM`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique `TEMPLATE` como el `Type` cuando llame a CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Para obtener más información, consulte Esquema JSON de Adobe Experience Manager.

En la siguiente tabla se describen los parámetros del esquema JSON de AEM.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
aemUrl	La URL del host de Adobe Experience Manager. Por ejemplo, si utiliza AEM On-Premise, debe incluir el nombre de host y el puerto: https://hostname:port. O bien, si usa AEM as a Cloud Service, puede usar la URL del autor: https://author-xxxxxx-xxxxxxx.adobeaemcloud.com.
authType	El tipo de autenticación que utiliza, ya sea `Basic` o `OAuth2`.
deploymentType	El tipo de Adobe Experience Manager que utiliza, ya sea `CLOUD` o `ON_PREMISE` .
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
• page • asset	Una lista de objetos que asignan los atributos o los nombres de campo de sus Adobe Experience Manager páginas y activos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.

Configuración	Descripción
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
timeZoneId	Si utiliza AEM On-Premise y la zona horaria del servidor es diferente a la zona horaria del conector o índice de Amazon Kendra AEM, puede especificar la zona horaria del servidor para alinearla con el conector o índice de AEM. La zona horaria predeterminada de AEM On-Premise es la zona horaria del conector o índice de AEM. Amazon Kendra La zona horaria predeterminada de AEM as a Cloud Service es la hora media de Greenwich.
• pageRootPaths • assetRootPaths	Una lista de rutas raíz para páginas y recursos. Por ejemplo, la ruta raíz de una página podría ser /content/sub y la ruta raíz de un recurso podría ser /content/sub/asset1.
crawlAssets	`true` para rastrear recursos.
crawlPages	`true` para rastrear páginas.
• pagePathInclusionPatrones • pageNameInclusionPatrones • assetPathInclusionPatrones • assetTypeInclusionPatrones • assetNameInclusionPatrones	Una lista de patrones de expresión regular para incluir determinadas páginas y recursos en su origen de datos de Adobe Experienc e Manager. Las páginas y recursos que coinciden con los patrones se incluyen en el índice. Las páginas y recursos que no coinciden con los patrones se excluyen del índice. Si una página o recurso coinciden con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.

Configuración	Descripción
• pagePathExclusionPatrones • pageNameExclusionPatrones • assetPathExclusionPatrones • assetTypeInclusionPatrones • assetNameInclusionPatrones	Una lista de patrones de expresión regular para excluir determinadas páginas y recursos de su origen de datos de Adobe Experienc e Manager. Las páginas y recursos que coinciden con los patrones se excluyen del índice. Las páginas y recursos que no coinciden con los patrones se incluyen en el índice. Si una página o recurso coinciden con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.
pageComponents	Una lista de nombres de componentes de página específicos que desea indexar.
contentFragmentVariations	Una lista de nombres para las variantes guardadas específicas de los fragmentos de contenido de Adobe Experience Manager que desea indexar.
type	El tipo del origen de datos. Especifica `AEM` como el tipo de origen de datos.
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
secretArn	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectars e a Adobe Experience Manager. Para obtener información sobre estos pares clave-valor, consulte las instrucciones de conexión de Adobe Experience Manager.

Configuración	Descripción
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Adobe Experience Manager

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties":
  {
    "connectionConfiguration": {
      "type": "object",
      "properties":
      {
        "repositoryEndpointMetadata":
        {
          "type": "object",
          "properties":
          {
            "aemUrl":
            {
              "type": "string",
              "pattern": "https:.*"
            },
            "authType": {
              "type": "string",
              "enum": ["Basic", "OAuth2"]
            },
            "deploymentType": {
              "type": "string",
              "enum": ["CLOUD","ON_PREMISE"]
            }
          },
          "required":
          [
            "aemUrl",
            "authType",
            "deploymentType"
          ]
        }
```

```
      },
      "required":
      [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties":
      {
        "page":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE",
                        "LONG"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
```

```
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "asset":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
                {
                  "type": "string",
                  "enum":
                  [
                    "STRING",
```

```
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    }
  }
},
"additionalProperties": {
  "type": "object",
  "properties":
  {
    "timeZoneId": {
      "type": "string",
      "enum": [
        "Africa/Abidjan",
        "Africa/Accra",
        "Africa/Addis_Ababa",
        "Africa/Algiers",
        "Africa/Asmara",
```

```
            "Africa/Asmera",
            "Africa/Bamako",
            "Africa/Bangui",
            "Africa/Banjul",
            "Africa/Bissau",
            "Africa/Blantyre",
            "Africa/Brazzaville",
            "Africa/Bujumbura",
            "Africa/Cairo",
            "Africa/Casablanca",
            "Africa/Ceuta",
            "Africa/Conakry",
            "Africa/Dakar",
            "Africa/Dar_es_Salaam",
            "Africa/Djibouti",
            "Africa/Douala",
            "Africa/El_Aaiun",
            "Africa/Freetown",
            "Africa/Gaborone",
            "Africa/Harare",
            "Africa/Johannesburg",
            "Africa/Juba",
            "Africa/Kampala",
            "Africa/Khartoum",
            "Africa/Kigali",
            "Africa/Kinshasa",
            "Africa/Lagos",
            "Africa/Libreville",
            "Africa/Lome",
            "Africa/Luanda",
            "Africa/Lubumbashi",
            "Africa/Lusaka",
            "Africa/Malabo",
            "Africa/Maputo",
            "Africa/Maseru",
            "Africa/Mbabane",
            "Africa/Mogadishu",
            "Africa/Monrovia",
            "Africa/Nairobi",
            "Africa/Ndjamena",
            "Africa/Niamey",
            "Africa/Nouakchott",
            "Africa/Ouagadougou",
            "Africa/Porto-Novo",
```

```
            "Africa/Sao_Tome",
            "Africa/Timbuktu",
            "Africa/Tripoli",
            "Africa/Tunis",
            "Africa/Windhoek",
            "America/Adak",
            "America/Anchorage",
            "America/Anguilla",
            "America/Antigua",
            "America/Araguaina",
            "America/Argentina/Buenos_Aires",
            "America/Argentina/Catamarca",
            "America/Argentina/ComodRivadavia",
            "America/Argentina/Cordoba",
            "America/Argentina/Jujuy",
            "America/Argentina/La_Rioja",
            "America/Argentina/Mendoza",
            "America/Argentina/Rio_Gallegos",
            "America/Argentina/Salta",
            "America/Argentina/San_Juan",
            "America/Argentina/San_Luis",
            "America/Argentina/Tucuman",
            "America/Argentina/Ushuaia",
            "America/Aruba",
            "America/Asuncion",
            "America/Atikokan",
            "America/Atka",
            "America/Bahia",
            "America/Bahia_Banderas",
            "America/Barbados",
            "America/Belem",
            "America/Belize",
            "America/Blanc-Sablon",
            "America/Boa_Vista",
            "America/Bogota",
            "America/Boise",
            "America/Buenos_Aires",
            "America/Cambridge_Bay",
            "America/Campo_Grande",
            "America/Cancun",
            "America/Caracas",
            "America/Catamarca",
            "America/Cayenne",
            "America/Cayman",
```

```
            "America/Chicago",
            "America/Chihuahua",
            "America/Ciudad_Juarez",
            "America/Coral_Harbour",
            "America/Cordoba",
            "America/Costa_Rica",
            "America/Creston",
            "America/Cuiaba",
            "America/Curacao",
            "America/Danmarkshavn",
            "America/Dawson",
            "America/Dawson_Creek",
            "America/Denver",
            "America/Detroit",
            "America/Dominica",
            "America/Edmonton",
            "America/Eirunepe",
            "America/El_Salvador",
            "America/Ensenada",
            "America/Fort_Nelson",
            "America/Fort_Wayne",
            "America/Fortaleza",
            "America/Glace_Bay",
            "America/Godthab",
            "America/Goose_Bay",
            "America/Grand_Turk",
            "America/Grenada",
            "America/Guadeloupe",
            "America/Guatemala",
            "America/Guayaquil",
            "America/Guyana",
            "America/Halifax",
            "America/Havana",
            "America/Hermosillo",
            "America/Indiana/Indianapolis",
            "America/Indiana/Knox",
            "America/Indiana/Marengo",
            "America/Indiana/Petersburg",
            "America/Indiana/Tell_City",
            "America/Indiana/Vevay",
            "America/Indiana/Vincennes",
            "America/Indiana/Winamac",
            "America/Indianapolis",
            "America/Inuvik",
```

```
            "America/Iqaluit",
            "America/Jamaica",
            "America/Jujuy",
            "America/Juneau",
            "America/Kentucky/Louisville",
            "America/Kentucky/Monticello",
            "America/Knox_IN",
            "America/Kralendijk",
            "America/La_Paz",
            "America/Lima",
            "America/Los_Angeles",
            "America/Louisville",
            "America/Lower_Princes",
            "America/Maceio",
            "America/Managua",
            "America/Manaus",
            "America/Marigot",
            "America/Martinique",
            "America/Matamoros",
            "America/Mazatlan",
            "America/Mendoza",
            "America/Menominee",
            "America/Merida",
            "America/Metlakatla",
            "America/Mexico_City",
            "America/Miquelon",
            "America/Moncton",
            "America/Monterrey",
            "America/Montevideo",
            "America/Montreal",
            "America/Montserrat",
            "America/Nassau",
            "America/New_York",
            "America/Nipigon",
            "America/Nome",
            "America/Noronha",
            "America/North_Dakota/Beulah",
            "America/North_Dakota/Center",
            "America/North_Dakota/New_Salem",
            "America/Nuuk",
            "America/Ojinaga",
            "America/Panama",
            "America/Pangnirtung",
            "America/Paramaribo",
```

```
            "America/Phoenix",
            "America/Port-au-Prince",
            "America/Port_of_Spain",
            "America/Porto_Acre",
            "America/Porto_Velho",
            "America/Puerto_Rico",
            "America/Punta_Arenas",
            "America/Rainy_River",
            "America/Rankin_Inlet",
            "America/Recife",
            "America/Regina",
            "America/Resolute",
            "America/Rio_Branco",
            "America/Rosario",
            "America/Santa_Isabel",
            "America/Santarem",
            "America/Santiago",
            "America/Santo_Domingo",
            "America/Sao_Paulo",
            "America/Scoresbysund",
            "America/Shiprock",
            "America/Sitka",
            "America/St_Barthelemy",
            "America/St_Johns",
            "America/St_Kitts",
            "America/St_Lucia",
            "America/St_Thomas",
            "America/St_Vincent",
            "America/Swift_Current",
            "America/Tegucigalpa",
            "America/Thule",
            "America/Thunder_Bay",
            "America/Tijuana",
            "America/Toronto",
            "America/Tortola",
            "America/Vancouver",
            "America/Virgin",
            "America/Whitehorse",
            "America/Winnipeg",
            "America/Yakutat",
            "America/Yellowknife",
            "Antarctica/Casey",
            "Antarctica/Davis",
            "Antarctica/DumontDUrville",
```

```
            "Antarctica/Macquarie",
            "Antarctica/Mawson",
            "Antarctica/McMurdo",
            "Antarctica/Palmer",
            "Antarctica/Rothera",
            "Antarctica/South_Pole",
            "Antarctica/Syowa",
            "Antarctica/Troll",
            "Antarctica/Vostok",
            "Arctic/Longyearbyen",
            "Asia/Aden",
            "Asia/Almaty",
            "Asia/Amman",
            "Asia/Anadyr",
            "Asia/Aqtau",
            "Asia/Aqtobe",
            "Asia/Ashgabat",
            "Asia/Ashkhabad",
            "Asia/Atyrau",
            "Asia/Baghdad",
            "Asia/Bahrain",
            "Asia/Baku",
            "Asia/Bangkok",
            "Asia/Barnaul",
            "Asia/Beirut",
            "Asia/Bishkek",
            "Asia/Brunei",
            "Asia/Calcutta",
            "Asia/Chita",
            "Asia/Choibalsan",
            "Asia/Chongqing",
            "Asia/Chungking",
            "Asia/Colombo",
            "Asia/Dacca",
            "Asia/Damascus",
            "Asia/Dhaka",
            "Asia/Dili",
            "Asia/Dubai",
            "Asia/Dushanbe",
            "Asia/Famagusta",
            "Asia/Gaza",
            "Asia/Harbin",
            "Asia/Hebron",
            "Asia/Ho_Chi_Minh",
```

```
            "Asia/Hong_Kong",
            "Asia/Hovd",
            "Asia/Irkutsk",
            "Asia/Istanbul",
            "Asia/Jakarta",
            "Asia/Jayapura",
            "Asia/Jerusalem",
            "Asia/Kabul",
            "Asia/Kamchatka",
            "Asia/Karachi",
            "Asia/Kashgar",
            "Asia/Kathmandu",
            "Asia/Katmandu",
            "Asia/Khandyga",
            "Asia/Kolkata",
            "Asia/Krasnoyarsk",
            "Asia/Kuala_Lumpur",
            "Asia/Kuching",
            "Asia/Kuwait",
            "Asia/Macao",
            "Asia/Macau",
            "Asia/Magadan",
            "Asia/Makassar",
            "Asia/Manila",
            "Asia/Muscat",
            "Asia/Nicosia",
            "Asia/Novokuznetsk",
            "Asia/Novosibirsk",
            "Asia/Omsk",
            "Asia/Oral",
            "Asia/Phnom_Penh",
            "Asia/Pontianak",
            "Asia/Pyongyang",
            "Asia/Qatar",
            "Asia/Qostanay",
            "Asia/Qyzylorda",
            "Asia/Rangoon",
            "Asia/Riyadh",
            "Asia/Saigon",
            "Asia/Sakhalin",
            "Asia/Samarkand",
            "Asia/Seoul",
            "Asia/Shanghai",
            "Asia/Singapore",
```

```
            "Asia/Srednekolymsk",
            "Asia/Taipei",
            "Asia/Tashkent",
            "Asia/Tbilisi",
            "Asia/Tehran",
            "Asia/Tel_Aviv",
            "Asia/Thimbu",
            "Asia/Thimphu",
            "Asia/Tokyo",
            "Asia/Tomsk",
            "Asia/Ujung_Pandang",
            "Asia/Ulaanbaatar",
            "Asia/Ulan_Bator",
            "Asia/Urumqi",
            "Asia/Ust-Nera",
            "Asia/Vientiane",
            "Asia/Vladivostok",
            "Asia/Yakutsk",
            "Asia/Yangon",
            "Asia/Yekaterinburg",
            "Asia/Yerevan",
            "Atlantic/Azores",
            "Atlantic/Bermuda",
            "Atlantic/Canary",
            "Atlantic/Cape_Verde",
            "Atlantic/Faeroe",
            "Atlantic/Faroe",
            "Atlantic/Jan_Mayen",
            "Atlantic/Madeira",
            "Atlantic/Reykjavik",
            "Atlantic/South_Georgia",
            "Atlantic/St_Helena",
            "Atlantic/Stanley",
            "Australia/ACT",
            "Australia/Adelaide",
            "Australia/Brisbane",
            "Australia/Broken_Hill",
            "Australia/Canberra",
            "Australia/Currie",
            "Australia/Darwin",
            "Australia/Eucla",
            "Australia/Hobart",
            "Australia/LHI",
            "Australia/Lindeman",
```

```
            "Australia/Lord_Howe",
            "Australia/Melbourne",
            "Australia/NSW",
            "Australia/North",
            "Australia/Perth",
            "Australia/Queensland",
            "Australia/South",
            "Australia/Sydney",
            "Australia/Tasmania",
            "Australia/Victoria",
            "Australia/West",
            "Australia/Yancowinna",
            "Brazil/Acre",
            "Brazil/DeNoronha",
            "Brazil/East",
            "Brazil/West",
            "CET",
            "CST6CDT",
            "Canada/Atlantic",
            "Canada/Central",
            "Canada/Eastern",
            "Canada/Mountain",
            "Canada/Newfoundland",
            "Canada/Pacific",
            "Canada/Saskatchewan",
            "Canada/Yukon",
            "Chile/Continental",
            "Chile/EasterIsland",
            "Cuba",
            "EET",
            "EST5EDT",
            "Egypt",
            "Eire",
            "Etc/GMT",
            "Etc/GMT+0",
            "Etc/GMT+1",
            "Etc/GMT+10",
            "Etc/GMT+11",
            "Etc/GMT+12",
            "Etc/GMT+2",
            "Etc/GMT+3",
            "Etc/GMT+4",
            "Etc/GMT+5",
            "Etc/GMT+6",
```

```
            "Etc/GMT+7",
            "Etc/GMT+8",
            "Etc/GMT+9",
            "Etc/GMT-0",
            "Etc/GMT-1",
            "Etc/GMT-10",
            "Etc/GMT-11",
            "Etc/GMT-12",
            "Etc/GMT-13",
            "Etc/GMT-14",
            "Etc/GMT-2",
            "Etc/GMT-3",
            "Etc/GMT-4",
            "Etc/GMT-5",
            "Etc/GMT-6",
            "Etc/GMT-7",
            "Etc/GMT-8",
            "Etc/GMT-9",
            "Etc/GMT0",
            "Etc/Greenwich",
            "Etc/UCT",
            "Etc/UTC",
            "Etc/Universal",
            "Etc/Zulu",
            "Europe/Amsterdam",
            "Europe/Andorra",
            "Europe/Astrakhan",
            "Europe/Athens",
            "Europe/Belfast",
            "Europe/Belgrade",
            "Europe/Berlin",
            "Europe/Bratislava",
            "Europe/Brussels",
            "Europe/Bucharest",
            "Europe/Budapest",
            "Europe/Busingen",
            "Europe/Chisinau",
            "Europe/Copenhagen",
            "Europe/Dublin",
            "Europe/Gibraltar",
            "Europe/Guernsey",
            "Europe/Helsinki",
            "Europe/Isle_of_Man",
            "Europe/Istanbul",
```

```
            "Europe/Jersey",
            "Europe/Kaliningrad",
            "Europe/Kiev",
            "Europe/Kirov",
            "Europe/Kyiv",
            "Europe/Lisbon",
            "Europe/Ljubljana",
            "Europe/London",
            "Europe/Luxembourg",
            "Europe/Madrid",
            "Europe/Malta",
            "Europe/Mariehamn",
            "Europe/Minsk",
            "Europe/Monaco",
            "Europe/Moscow",
            "Europe/Nicosia",
            "Europe/Oslo",
            "Europe/Paris",
            "Europe/Podgorica",
            "Europe/Prague",
            "Europe/Riga",
            "Europe/Rome",
            "Europe/Samara",
            "Europe/San_Marino",
            "Europe/Sarajevo",
            "Europe/Saratov",
            "Europe/Simferopol",
            "Europe/Skopje",
            "Europe/Sofia",
            "Europe/Stockholm",
            "Europe/Tallinn",
            "Europe/Tirane",
            "Europe/Tiraspol",
            "Europe/Ulyanovsk",
            "Europe/Uzhgorod",
            "Europe/Vaduz",
            "Europe/Vatican",
            "Europe/Vienna",
            "Europe/Vilnius",
            "Europe/Volgograd",
            "Europe/Warsaw",
            "Europe/Zagreb",
            "Europe/Zaporozhye",
            "Europe/Zurich",
```

```
            "GB",
            "GB-Eire",
            "GMT",
            "GMT0",
            "Greenwich",
            "Hongkong",
            "Iceland",
            "Indian/Antananarivo",
            "Indian/Chagos",
            "Indian/Christmas",
            "Indian/Cocos",
            "Indian/Comoro",
            "Indian/Kerguelen",
            "Indian/Mahe",
            "Indian/Maldives",
            "Indian/Mauritius",
            "Indian/Mayotte",
            "Indian/Reunion",
            "Iran",
            "Israel",
            "Jamaica",
            "Japan",
            "Kwajalein",
            "Libya",
            "MET",
            "MST7MDT",
            "Mexico/BajaNorte",
            "Mexico/BajaSur",
            "Mexico/General",
            "NZ",
            "NZ-CHAT",
            "Navajo",
            "PRC",
            "PST8PDT",
            "Pacific/Apia",
            "Pacific/Auckland",
            "Pacific/Bougainville",
            "Pacific/Chatham",
            "Pacific/Chuuk",
            "Pacific/Easter",
            "Pacific/Efate",
            "Pacific/Enderbury",
            "Pacific/Fakaofo",
            "Pacific/Fiji",
```

```
            "Pacific/Funafuti",
            "Pacific/Galapagos",
            "Pacific/Gambier",
            "Pacific/Guadalcanal",
            "Pacific/Guam",
            "Pacific/Honolulu",
            "Pacific/Johnston",
            "Pacific/Kanton",
            "Pacific/Kiritimati",
            "Pacific/Kosrae",
            "Pacific/Kwajalein",
            "Pacific/Majuro",
            "Pacific/Marquesas",
            "Pacific/Midway",
            "Pacific/Nauru",
            "Pacific/Niue",
            "Pacific/Norfolk",
            "Pacific/Noumea",
            "Pacific/Pago_Pago",
            "Pacific/Palau",
            "Pacific/Pitcairn",
            "Pacific/Pohnpei",
            "Pacific/Ponape",
            "Pacific/Port_Moresby",
            "Pacific/Rarotonga",
            "Pacific/Saipan",
            "Pacific/Samoa",
            "Pacific/Tahiti",
            "Pacific/Tarawa",
            "Pacific/Tongatapu",
            "Pacific/Truk",
            "Pacific/Wake",
            "Pacific/Wallis",
            "Pacific/Yap",
            "Poland",
            "Portugal",
            "ROK",
            "Singapore",
            "SystemV/AST4",
            "SystemV/AST4ADT",
            "SystemV/CST6",
            "SystemV/CST6CDT",
            "SystemV/EST5",
            "SystemV/EST5EDT",
```

```
            "SystemV/HST10",
            "SystemV/MST7",
            "SystemV/MST7MDT",
            "SystemV/PST8",
            "SystemV/PST8PDT",
            "SystemV/YST9",
            "SystemV/YST9YDT",
            "Turkey",
            "UCT",
            "US/Alaska",
            "US/Aleutian",
            "US/Arizona",
            "US/Central",
            "US/East-Indiana",
            "US/Eastern",
            "US/Hawaii",
            "US/Indiana-Starke",
            "US/Michigan",
            "US/Mountain",
            "US/Pacific",
            "US/Samoa",
            "UTC",
            "Universal",
            "W-SU",
            "WET",
            "Zulu",
            "EST",
            "HST",
            "MST",
            "ACT",
            "AET",
            "AGT",
            "ART",
            "AST",
            "BET",
            "BST",
            "CAT",
            "CNT",
            "CST",
            "CTT",
            "EAT",
            "ECT",
            "IET",
            "IST",
```

```
            "JST",
            "MIT",
            "NET",
            "NST",
            "PLT",
            "PNT",
            "PRT",
            "PST",
            "SST",
            "VST"
          ]
        },
        "pageRootPaths":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetRootPaths":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "crawlAssets":
        {
          "type": "boolean"
        },
        "crawlPages":
        {
          "type": "boolean"
        },
        "pagePathInclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
```

```
        "pagePathExclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "pageNameInclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "pageNameExclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetPathInclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetPathExclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetTypeInclusionPatterns":
        {
          "type": "array",
          "items":
```

```
          {
            "type": "string"
          }
        },
        "assetTypeExclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetNameInclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "assetNameExclusionPatterns":
        {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "pageComponents": {
          "type": "array",
          "items": {
            "type": "object"
            }
        },
        "contentFragmentVariations": {
          "type": "array",
          "items": {
            "type": "object"
          }
        },
        "cugExemptedPrincipals": {
          "type": "array",
          "items": {
            "type": "string"
```

```
          }
        }
      },
      "required":
      []
    },
    "type": {
      "type": "string",
      "pattern": "AEM"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
```

```
}
```

Amazon FSx Esquema de plantillas (Windows)

Incluye un JSON que contiene el esquema del origen de datos como parte del objeto TemplateConfiguration. El identificador del sistema de archivos se proporciona como parte de la configuración de la conexión o de los detalles del punto final del repositorio. También debe especificar el tipo de fuente de datosFSX, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique `TEMPLATE` como el `Type` cuando llame a CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Amazon FSx Esquema JSON (Windows).

En la siguiente tabla se describen los parámetros del esquema JSON Amazon FSx (Windows).

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
fileSystemId	El identificador del sistema de Amazon FSx archivos. Puede encontrar el ID del sistema de archivos en el panel de sistemas de archivos de la Amazon FSx consola.
fileSystemType	El tipo Amazon FSx de sistema de archivos. Para usarlo Windows File Server como tipo de sistema de archivos, especifique`WINDOWS`.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
Todos	Una lista de objetos que mapean los atributos o los nombres de campo de los archivos de

Configuración	Descripción
	la fuente de Amazon FSx datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
isCrawlAcl	`true`para rastrear la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documento s pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
inclusionPatterns	Una lista de patrones de expresiones regulares para incluir determinados archivos en la fuente Amazon FSx de datos. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.

Configuración	Descripción
exclusionPatterns	Una lista de patrones de expresiones regulares para excluir determinados archivos de la fuente Amazon FSx de datos. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
type	El tipo del origen de datos. Para las fuentes de datos del sistema de archivos de Windows, especifique`FSX`.

Amazon FSx Esquema JSON (Windows)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "fileSystemId": {
              "type": "string",
```

```
            "pattern": "fs-.*"
          },
          "fileSystemType": {
            "type": "string",
            "pattern": "WINDOWS"
          }
        },
        "required": ["fileSystemId", "fileSystemType"]
      }
    }
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "All": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "STRING_LIST", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
```

```
            ]
          }
        },
        "required": ["fieldMappings"]
      }
    },
    "required": ["All"]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "isCrawlAcl": {
        "type": "boolean"
      },
      "exclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      }
    },
    "required": []
  },
  "enableIdentityCrawler": {
    "type": "boolean"
  },
  "syncMode": {
    "type": "string",
    "enum": [
      "FORCED_FULL_CRAWL",
      "FULL_CRAWL"
    ]
  },
  "type" : {
    "type" : "string",
    "pattern": "FSX"
  }
},
```

```
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "enableIdentityCrawler",
    "additionalProperties",
    "type"
  ]
}
```

Amazon FSx Esquema de plantillas (NetApp ONTAP)

Incluye un JSON que contiene el esquema del origen de datos como parte del objeto TemplateConfiguration. Debe proporcionar el identificador del sistema de archivos y la máquina virtual de almacenamiento (SVM) como parte de la configuración de la conexión o de los detalles del punto final del repositorio. También debe especificar el tipo de fuente de datos`FSXONTAP`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique `TEMPLATE` como el `Type` cuando llame a CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Amazon FSx (NetApp ONTAP) Esquema JSON.

En la siguiente tabla se describen los parámetros del esquema JSON Amazon FSx (NetApp ONTAP).

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.

Configuración	Descripción
fileSystemId	El identificador del sistema de Amazon FSx archivos. Puede encontrar el ID del sistema de archivos en el panel de sistemas de archivos de la Amazon FSx consola. Para obtener información sobre cómo crear un sistema de archivos en la Amazon FSx consola de NetApp ONTAP, consulte la Guía de introducción a NetApp ONTAP en la Guía del FSx for ONTAP usuario.
fileSystemType	El tipo de sistema Amazon FSx de archivos. Para usarlo NetApp ONTAP como tipo de sistema de archivos, especifique0NTAP.
SVMid	El identificador de la máquina virtual de almacenamiento (SVM) utilizada con el sistema de Amazon FSx archivos para. NetApp ONTAP Para encontrar su ID de SVM, vaya al panel de sistemas de archivos de la Amazon FSx consola, seleccione su ID de sistema de archivos y, a continuación, seleccion e Máquinas virtuales de almacenamiento. Para obtener información sobre cómo crear un sistema de archivos en la Amazon FSx consolaNetApp ONTAP, consulte la Guía de introducción a NetApp ONTAP en la Guía del FSx for ONTAP usuario.
Tipo de protocolo	Ya sea que utilice el protocolo CIFS (Common Internet File System) para Windows o el protocolo Network File System (NFS) para Linux.

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
archivo	Una lista de objetos que mapean los atributos o los nombres de campo de los archivos de la fuente de Amazon FSx datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos. Los nombres de los campos de la fuente de datos deben estar en los metadatos personalizados de los archivos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
crawlAcl	`true`para rastrear la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documento s pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

Configuración	Descripción
inclusionPatterns	Una lista de patrones de expresiones regulares para incluir determinados archivos en la fuente Amazon FSx de datos. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
exclusionPatterns	Una lista de patrones de expresiones regulares para excluir determinados archivos de la fuente Amazon FSx de datos. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
type	El tipo del origen de datos. Para las fuentes de datos del sistema de NetApp ONTAP archivos, especifique`FSXONTAP`.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción

secretArn	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse al sistema de archivos. Amazon FSx El secreto debe contener una estructura JSON con las siguientes claves:
	`{ "username": " user@corp.example.com ", "password": " password" }`
	Si utiliza el protocolo NFS para su sistema de Amazon FSx archivos, el secreto se almacena en una estructura JSON con las siguientes claves:
	`{ "leftId": "left ID", "rightId": " right ID", "preSharedKey": " pre-shared key " }`

Amazon FSx (NetApp ONTAP) Esquema JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "fileSystemId": {
```

```
            "type": "string",
              "pattern": "^(fs-[0-9a-f]{8,21})$"
          },
          "fileSystemType": {
            "type": "string",
            "enum": ["ONTAP"]
          },
          "svmId": {
            "type": "string",
            "pattern": "^(svm-[0-9a-f]{17,21})$"
          },
          "protocolType": {
            "type": "string",
            "enum": [
              "CIFS",
              "NFS"
            ]
          }
        },
        "required": [
          "fileSystemId",
          "fileSystemType"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "file": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string",
                    "pattern": "^([a-zA-Z_]{1,20})$"
```

```
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "STRING_LIST",
                        "DATE",
                        "LONG"
                      ]
                    },
                    "dataSourceFieldName": {
                      "type": "string",
                      "pattern": "^([a-zA-Z_]{1,20})$"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ],
              "maxItems": 50
            }
          },
          "required": [
            "fieldMappings"
          ]
        }
      },
      "required": [
        "file"
      ]
    },
    "additionalProperties": {
      "type": "object",
      "properties": {
        "crawlAcl": {
          "type": "boolean"
        },
```

```
        "inclusionPatterns": {
          "type": "array",
          "items": {
            "type": "string",
            "maxLength": 30
          },
          "maxItems": 100
        },
        "exclusionPatterns": {
          "type": "array",
          "items": {
            "type": "string",
            "maxLength": 30
          },
          "maxItems": 100
        }
      }
    },
    "type": {
      "type": "string",
      "pattern": "FSXONTAP"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL"
      ]
    },
    "secretArn": {
      "type": "string",
      "pattern": "arn:aws:secretsmanager:.*"
    }
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Esquema de plantilla de Alfresco

Incluye un JSON que contiene el esquema del origen de datos como parte del objeto TemplateConfiguration. Debe proporcionar el ID del sitio de Alfresco, la URL del repositorio, la URL de la interfaz de usuario, el tipo de autenticación, si utiliza la nube o en las instalaciones y el tipo de contenido que desea rastrear. Debe proporcionarlos como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Especifique también el tipo de origen de datos como `ALFRESCO`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique `TEMPLATE` como el `Type` cuando llame a CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Alfresco.

En la siguiente tabla se describen los parámetros del esquema JSON de Alfresco.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
siteId	El identificador del sitio de Alfresco.
repoUrl	La URL de su repositorio de Alfresco. Puede obtener la URL del repositorio de su administrador de Alfresco. Por ejemplo, si utiliza Alfresco Cloud (PaaS), la URL del repositorio podría ser https://company.alfrescocloud.com. O bien, si utiliza Alfresco On-Premises, la URL del repositorio podría ser https://company-alfresco-instance.company-domain.suffix:port.
webAppUrl	La URL de la interfaz de usuario de Alfresco. Puede obtener la URL de la interfaz de usuario de Alfresco de su administrador de Alfresco. Por ejemplo, la URL de la interfaz de usuario podría ser https://example.com.

Configuración	Descripción
repositoryAdditionalProperties	Propiedades adicionales para conectarse con el punto de conexión del repositorio/origen de datos.
authType	El tipo de autenticación que utiliza, ya sea `OAuth2` o `Basic`.
type (implementación)	El tipo de Alfresco que utiliza, ya sea `PAAS` o `ON-PREM`.
crawlType	El tipo de contenido que quiere rastrear, ya sea `ASPECT` (contenido marcado con “Aspectos” en Alfresco), `SITE_ID` (contenido de un sitio de Alfresco específico) o `ALL_SITES` (contenido de todos sus sitios de Alfresco).
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
• document • comentario	Una lista de objetos que mapean los atributos o los nombres de campo de sus documentos y comentarios de Alfresco para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
aspectName	El nombre del “Aspecto” específico que desea indexar.
aspectProperties	Una lista de propiedades de contenido de “Aspecto” específicas que desea indexar.

Configuración	Descripción
enableFineGrainedControl	`true` para rastrear “Aspectos”.
isCrawlComment	`true`para rastrear los comentarios.
• inclusionFileNamePatrones • inclusionFileTypePatrones • inclusionFilePathPatrones	Una lista de patrones de expresión regular para incluir determinados archivos en su origen de datos de Alfresco. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
• exclusionFileNamePatrones • exclusionFileTypePatrones • exclusionFilePathPatrones	Una lista de patrones de expresión regular para excluir determinados archivos en su origen de datos de Alfresco. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
type	El tipo del origen de datos. Especifica `ALFRESCO` como el tipo de origen de datos.

Configuración	Descripción

secretArn	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectars e a su. Alfresco El secreto debe contener una estructura JSON con las siguientes claves:
	
	Si utiliza la autenticación básica:
	
	`{ "username": " user name", "password": " password" }`
	
	Si utiliza la autenticación OAuth 2.0:
	
	`{ "clientId": " client ID", "clientSecret": " client secret", "tokenUrl": " token URL" }`

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Alfresco

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "siteId": {
              "type": "string"
            },
            "repoUrl": {
              "type": "string"
            },
            "webAppUrl": {
              "type": "string"
            },
            "repositoryAdditionalProperties": {
              "type": "object",
              "properties": {
                "authType": {
                  "type": "string",
                  "enum": [
                    "OAuth2",
                    "Basic"
                  ]
                },
                "type": {
                  "type": "string",
                  "enum": [
                    "PAAS",
                    "ON_PREM"
                  ]
                },
                "crawlType": {
                  "type": "string",
                  "enum": [
                    "ASPECT",
                    "SITE_ID",
                    "ALL_SITES"
```

```
                ]
              }
            }
          }
        }
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "DATE",
                        "STRING_LIST",
                        "LONG"
                      ]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
```

```
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "comment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "STRING_LIST",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
```

```
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "aspectName": {
        "type": "string"
      },
      "aspectProperties": {
        "type": "array"
      },
      "enableFineGrainedControl": {
        "type": "boolean"
      },
      "isCrawlComment": {
        "type": "boolean"
      },
      "inclusionFileNamePatterns": {
        "type": "array"
      },
      "exclusionFileNamePatterns": {
        "type": "array"
      },
      "inclusionFileTypePatterns": {
        "type": "array"
      },
      "exclusionFileTypePatterns": {
        "type": "array"
      },
      "inclusionFilePathPatterns": {
```

```
          "type": "array"
        },
        "exclusionFilePathPatterns": {
          "type": "array"
        }
      }
    },
    "type": {
      "type": "string",
      "pattern": "ALFRESCO"
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL"
      ]
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "type",
    "secretArn"
  ]
}
```

Aurora Esquema de plantillas (MySQL)

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `mysql`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Aurora Esquema JSON (MySQL).

En la siguiente tabla se describen los parámetros del esquema JSON Aurora (MySQL).

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea`mysql`,`db2`, `postgresql` `oracle`, o. `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar

Configuración	Descripción
	los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .

Configuración	Descripción
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Aurora Esquema JSON (MySQL)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Aurora Esquema de plantillas (PostgreSQL)

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `postgresql`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Aurora Esquema JSON (PostgreSQL).

En la siguiente tabla se describen los parámetros del esquema Aurora JSON (PostgreSQL).

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea, `mysqldb2`, `postgresql` o. `oracle` `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

Configuración	Descripción
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.

Configuración	Descripción
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Aurora Esquema JSON (PostgreSQL)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon RDS Esquema de plantillas (Microsoft SQL Server)

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `sqlserver`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Amazon RDS Esquema JSON (Microsoft SQL Server).

En la siguiente tabla se describen los parámetros del esquema JSON Amazon RDS (Microsoft SQL Server).

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea `mysqldb2`, `postgresql oracle`, o. `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

Configuración	Descripción
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.

Configuración	Descripción
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Amazon RDS Esquema JSON (Microsoft SQL Server)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon RDS Esquema de plantillas (MySQL)

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `mysql`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Amazon RDS Esquema JSON (MySQL).

En la siguiente tabla se describen los parámetros del esquema JSON Amazon RDS (MySQL).

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea`mysql`,`db2`, `postgresql` `oracle`, o. `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

Configuración	Descripción
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.

Configuración	Descripción
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Amazon RDS Esquema JSON (MySQL)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon RDS Esquema de plantillas (Oracle)

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `oracle`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Amazon RDS (Oracle) Esquema JSON.

En la siguiente tabla se describen los parámetros del esquema JSON Amazon RDS (Oracle).

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea `mysql`,`db2`, `postgresql oracle`, o. `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

Configuración	Descripción
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.

Configuración	Descripción
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Amazon RDS (Oracle) Esquema JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon RDS Esquema de plantillas (PostgreSQL)

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `postgresql`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Amazon RDS Esquema JSON (PostgreSQL).

En la siguiente tabla se describen los parámetros del esquema Amazon RDS JSON (PostgreSQL).

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea, `mysqldb2`, `postgresql` o. `oracle` `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

Configuración	Descripción
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.

Configuración	Descripción
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Amazon RDS Esquema JSON (PostgreSQL)

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Amazon S3 esquema de plantilla

Incluye un JSON que contiene el esquema del origen de datos como parte de la configuración de la plantilla. Debe proporcionar el nombre del bucket de S3 como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Especifique también el tipo de origen de datos como `S3` y otras configuraciones necesarias. A continuación, especifique `TEMPLATE` como `Type` cuando llame CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de S3.

En la siguiente tabla se describen los parámetros del esquema Amazon S3 JSON.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
BucketName	El nombre de tu Amazon S3 depósito.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos

Configuración	Descripción
• inclusionPatterns • exclusionPatterns • inclusionPrefixes • exclusionPrefixes	Una lista de patrones de expresiones regulares para incluir o excluir archivos específicos de la fuente Amazon S3 de datos. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
aclConfigurationFileRuta	La ruta del archivo que controla el acceso a los documentos en un índice de Amazon Kendra .
metadataFilesPrefix	La ubicación dentro del bucket para los archivos de metadatos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
type	El tipo del origen de datos. Especifica S3 como el tipo de origen de datos.
versión	La versión de la plantilla admitida.

Esquema JSON de S3

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "BucketName": {
              "type": "string"
            }
          },
          "required": [
            "BucketName"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "document": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
```

```
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
      "document"
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "inclusionPatterns": {
        "type": "array"
      },
      "exclusionPatterns": {
        "type": "array"
      },
      "inclusionPrefixes": {
        "type": "array"
```

```
        },
        "exclusionPrefixes": {
          "type": "array"
        },
        "aclConfigurationFilePath": {
          "type": "string"
        },
        "metadataFilesPrefix": {
          "type": "string"
        }
      }
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FULL_CRAWL",
        "FORCED_FULL_CRAWL"
      ]
    },
    "type": {
      "type": "string",
      "pattern": "S3"
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "required": [
    "connectionConfiguration",
    "type",
    "syncMode",
    "repositoryConfigurations"
  ]
}
```

Amazon Kendra Esquema de plantillas de Web Crawler

Incluye un JSON que contiene el esquema del origen de datos como parte del objeto TemplateConfiguration.

Debe proporcionar las URL semilla o de punto de partida, o puede proporcionar las URL del mapa del sitio, como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. En lugar de enumerar manualmente todas las URL, puede proporcionar la ruta al Amazon S3 depósito que almacena un archivo de texto para su lista de direcciones URL iniciales o archivos XML de mapa del sitio, que puede agrupar en un archivo ZIP en S3.

También puede especificar el tipo de fuente de datos`WEBCRAWLERV2`, las credenciales de autenticación del sitio web y el tipo de autenticación si sus sitios web requieren autenticación, además de otras configuraciones necesarias.

A continuación, especifique `TEMPLATE` como el `Type` cuando llame a CreateDataSource.

> **Important**
>
> La creación de conectores Web Crawler v2.0 no es compatible con. AWS CloudFormation Utilice el conector Web Crawler v1.0 si necesita asistencia. AWS CloudFormation

Al seleccionar los sitios web que se van a indexar, se debe respetar la Política de uso aceptable de Amazon y todas las demás condiciones de Amazon. Recuerde que solo debe usar Amazon Kendra Web Crawler para indexar sus propias páginas web o páginas web para las que tenga autorización para indexar. Para obtener información sobre cómo impedir que el rastreador web de Amazon Kendra indexe sus sitios web, consulte Configuración del archivo *`robots.txt`* para el rastreador web de Amazon Kendra.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Amazon Kendra Esquema JSON de Web Crawler.

En la siguiente tabla se describen los parámetros del esquema JSON del Amazon Kendra Web Crawler.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
siteMapUrls	La lista de URL del mapa del sitio web de los sitios web que quiere rastrear. Puede enumerar hasta tres URL de mapa del sitio.
s3 SeedUrl	La ruta de S3 al archivo de texto que almacena la lista de URL semilla o de punto de partida. Por ejemplo, s3://bucket-name/directory/. Cada URL del archivo de texto debe estar formatead a en una línea independiente. Puede enumerar hasta 100 URL semilla en un archivo.
s3 SiteMapUrl	La ruta S3 a los archivos XML de mapa del sitio. Por ejemplo, s3://bucket-name/directory/ . Puede enumerar hasta tres archivos XML de mapa del sitio. Puedes agrupar varios archivos de mapa del sitio en un archivo ZIP y almacenar el archivo ZIP en tu Amazon S3 depósito.
seedUrlConnections	La lista de URL semilla o de punto de partida de los sitios web que desea rastrear. Puede enumerar hasta 100 URL semilla.
seedUrl	La URL semilla o de punto de partida.
authentication	El tipo de autenticación si sus sitios web requieren la misma autenticación; en caso contrario, especifique `NoAuthentication` .
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos

Configuración	Descripción
	específicos de contenido y asignaciones de campos.
• webPage • attachment	Una lista de objetos que mapean los atributos o los nombres de campo de sus páginas web y archivos de páginas web para Amazon Kendra indexar los nombres de los campos. Por ejemplo, la etiqueta de título de la página web HTML se puede asignar al campo de índice de `_document_title` . Para obtener más información, consulte Asignación de campos de origen de datos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
rateLimit	El número de direcciones URL rastreadas por host de sitio web por minuto.

Configuración	Descripción
maxFileSize	Tamaño máximo (en MB) de una página web o un archivo adjunto que se van a rastrear.
crawlDepth	El número de niveles desde la URL semilla que se va a rastrear. Por ejemplo, la página URL semilla tiene la profundidad 1 y todos los hipervínculos de esta página que también se rastreen tienen la profundidad 2.
maxLinksPerURL	El número máximo de URL de una página web que se deben incluir al rastrear un sitio web. Este número es por página web. A medida que se rastrean las páginas web de un sitio web, también se rastrean las URL a las que enlazan las páginas web. Las URL de una página web se rastrean en orden de aparición.
crawlSubDomain	`true` para rastrear los dominios del sitio web con subdominios. Por ejemplo, si la URL semilla es “abc.example.com”, también se rastrearán “a.abc.example.com ” y “b.abc.example.com”. Si no lo configuras `crawlSubDomain` o `crawlAllDomain` no lo hace`true`, Amazon Kendra solo rastreará los dominios de los sitios web que quieras rastrear.
crawlAllDomain	`true` para rastrear los dominios del sitio web con subdominios y otros dominios a los que enlazan las páginas web. Si no lo configuras `crawlSubDomain` ni `crawlAllDomain` lo hace`true`, Amazon Kendra solo rastreará los dominios de los sitios web que quieras rastrear.

Configuración	Descripción
honorRobots	`true` para respetar las directivas de robots.txt de los sitios web que desea rastrear. Estas directivas controlan la forma en que Amazon Kendra Web Crawler rastrea los sitios web, ya sea que solo Amazon Kendra pueda rastrear contenido específico o no rastrear ningún contenido.
crawlAttachments	`true` para rastrear los archivos a los que enlazan las páginas web.
• URL de inclusión CrawlPatterns • URL de inclusión IndexPatterns	Una lista de patrones de expresiones regulares para incluir el rastreo de determinadas URL y la indexación de los hipervínculos de estas páginas web con URL. Las URL que coinciden con los patrones se incluyen en el índice. Las URL que no coinciden con los patrones se excluyen del índice. Si una URL coincide con un patrón de exclusión y un patrón de inclusión, el patrón de exclusión tiene prioridad y la URL/páginas web del sitio web no se incluyen en el índice.
• URL de exclusión CrawlPatterns • URL de exclusión IndexPatterns	Una lista de patrones de expresiones regulares para excluir el rastreo de determinadas URL y la indexación de los hipervínculos de estas páginas web con URL. Las URL que coinciden con los patrones se excluyen del índice. Las URL que no coinciden con los patrones se incluyen en el índice. Si una URL coincide con un patrón de exclusión y un patrón de inclusión, el patrón de exclusión tiene prioridad y la URL/páginas web del sitio web no se incluyen en el índice.

Configuración	Descripción
inclusionFileIndexPatrones	Una lista de patrones de expresión regular para incluir determinados archivos de página web. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
exclusionFileIndexPatrones	Una lista de patrones de expresión regular para excluir determinados archivos de página web. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
proxy	Información de configuración necesaria para conectarse a sus sitios web internos a través de un proxy web.
host	El nombre del host del servidor proxy que desea utilizar para conectarse a sitios web internos. Por ejemplo, el nombre de host de https://a.example.com/page1.html es “a.example.com”.
port	El número de puerto del servidor proxy que desea utilizar para conectarse a sitios web internos. Por ejemplo, 443 es el puerto estándar para HTTPS.

Configuración	Descripción
secretArn (proxy)	Si se requieren credenciales de proxy web para conectarse a un servidor de sitios web, puede crear un AWS Secrets Manager secreto que almacene las credenciales. Proporciona el nombre de recurso de Amazon (ARN) del secreto.
type	El tipo del origen de datos. Especifica `WEBCRAWLERV2` como el tipo de origen de datos.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que se utiliza si sus sitios web requieren autenticación para acceder a los sitios web. Las credenciales de autenticación del sitio web se almacenan en el secreto que contiene los pares clave-valor de JSON. Si utiliza Basic o NTLM/Kerberos, introduzca el nombre de usuario y la contraseña. Las claves JSON del secreto deben ser `userName` y `password`. El protocolo de autenticación NTLM incluye el hash de contraseñas y el protocolo de autenticación de Kerberos incluye el cifrado de contraseñas. Si utiliza la autenticación mediante SAML o mediante formulario, introduzca el nombre de usuario y la contraseña, XPath para el campo del nombre de usuario (y el botón del nombre de usuario si utiliza SAML), XPaths para el campo y el botón de la contraseña y la URL de la página de inicio de sesión. Las claves JSON del secreto deben ser `userName`, `password`, `userNameFieldXpath` , `userNameButtonXpath` , `passwordFieldXpath` , `passwordButtonXpath` y `loginPageUrl` . Puede encontrar los XPaths (lenguaje de rutas XML) de los elementos utilizando las herramientas para desarrolladores de su navegador web. Los XPaths suelen seguir este formato: `//tagname[@Attribute='Value']` .

Configuración	Descripción
	Amazon Kendra también comprueba si la información de punto final (URL iniciales) incluida en el secreto es la misma que la información de punto final especificada en los detalles de configuración del punto final de la fuente de datos.
versión	La versión de esta plantilla que se admite actualmente.

Amazon Kendra Esquema JSON de Web Crawler

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "siteMapUrls": {
              "type": "array",
              "items":{
                "type": "string",
                "pattern": "https://.*"
              }
            },
            "s3SeedUrl": {
              "type": "string",
              "pattern": "s3:.*"
            },
            "s3SiteMapUrl": {
              "type": "string",
              "pattern": "s3:.*"
            },
            "seedUrlConnections": {
              "type": "array",
```

```
            "items": [
              {
                "type": "object",
                "properties": {
                  "seedUrl":{
                    "type": "string",
                    "pattern": "https://.*"
                  }
                },
                "required": [
                  "seedUrl"
                ]
              }
            ]
          },
          "authentication": {
            "type": "string",
            "enum": [
              "NoAuthentication",
              "BasicAuth",
              "NTLM_Kerberos",
              "Form",
              "SAML"
            ]
          }
        }
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "webPage": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
```

```
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "attachment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
```

```
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "syncMode": {
    "type": "string",
    "enum": [
      "FORCED_FULL_CRAWL",
      "FULL_CRAWL"
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "rateLimit": {
        "type": "string",
        "default": "300"
```

```
        },
        "maxFileSize": {
          "type": "string",
          "default": "50"
        },
        "crawlDepth": {
          "type": "string",
          "default": "2"
        },
        "maxLinksPerUrl": {
          "type": "string",
          "default": "100"
        },
        "crawlSubDomain": {
          "type": "boolean",
          "default": false
        },
        "crawlAllDomain": {
          "type": "boolean",
          "default": false
        },
        "honorRobots": {
          "type": "boolean",
          "default": false
        },
        "crawlAttachments": {
          "type": "boolean",
          "default": false
        },
        "inclusionURLCrawlPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionURLCrawlPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionURLIndexPatterns": {
          "type": "array",
          "items": {
```

```
            "type": "string"
          }
        },
        "exclusionURLIndexPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileIndexPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileIndexPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "proxy": {
          "type": "object",
          "properties": {
            "host": {
              "type": "string"
            },
            "port": {
              "type": "string"
            },
            "secretArn": {
              "type": "string",
              "minLength": 20,
              "maxLength": 2048
            }
          }
        }
      },
      "required": [
        "rateLimit",
        "maxFileSize",
        "crawlDepth",
        "crawlSubDomain",
        "crawlAllDomain",
```

```
        "maxLinksPerUrl",
        "honorRobots"
      ]
    },
    "type": {
      "type": "string",
      "pattern": "WEBCRAWLERV2"
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "type",
    "additionalProperties"
  ]
}
```

Esquema de plantilla de Confluence

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Proporciona la URL del host de Confluence, el método de alojamiento y el tipo de autenticación como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Además, especifique el tipo de origen de datos como `CONFLUENCEV2`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Confluence.

En la siguiente tabla se describen los parámetros del esquema JSON de Confluence.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
hostUrl	La dirección URL de la instancia de Confluence. Por ejemplo, *https://example.confluence.com* .
type	El método de alojamiento de su instancia de Confluence, ya sea `SAAS` o `ON_PREM`.
authType	El método de autenticación de su instancia de Confluence, ya sea `Basic`, `OAuth2` o `Personal-token` .
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
• espacio • page • blog • comentario • attachment	Una lista de objetos que mapean los atributos o los nombres de campo de tus espacios, páginas, blogs, comentarios y archivos adjuntos de Confluence para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos). Los nombres de los campos del origen de datos de Confluence deben existir en sus metadatos personalizados de Confluence.

Configuración	Descripción
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
isCrawlAcl	`true`para rastrear la información de la lista de control de acceso (ACL) de tus documentos, si tienes una ACL y quieres usarla para el control de acceso. La ACL especifica a qué documento s pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
fieldForUserID	Especifique `email` si desea utilizar el correo electrónico del usuario como ID de usuario. `email`se usa de forma predeterminada y actualmente es el único tipo de ID de usuario compatible. | • inclusionSpaceKeyFiltrar • exclusionSpaceKeyFiltro • pageTitleRegEX • blogTitleRegEX • commentTitleRegEX • attachmentTitleRegEX • inclusionFileTypePatrones • exclusionFileTypePatrones • inclusionUrlPatterns • exclusionUrlPatterns	Una lista de patrones de expresión regular para incluir o excluir determinados archivos en su origen de datos de Confluence. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
proxyHost	El nombre de host del proxy web que utilizas, sin el `https://` protocolo `http://` o.

Configuración	Descripción
proxyPort	El número de puerto utilizado por el protocolo de transporte de URL del host. Debe ser un valor numérico entre 0 y 65535.
• isCrawlPersonalEspacio • isCrawlArchivedEspacio • isCrawlArchivedPágina • isCrawlPage • isCrawlBlog • isCrawlPageComentario • isCrawlPageAdjunto • isCrawlBlogComentario • isCrawlBlogAdjunto	`true`para rastrear los archivos de tus espacios personales, páginas, blogs, comentarios de página, adjuntos de página, comentarios de blog y archivos adjuntos de blog de Confluenc e.
maxFileSizeInMegaBytes	Especifica el límite de tamaño de los archivos en MB que se Amazon Kendra pueden rastrear. Amazon Kendra rastrea solo los archivos dentro del límite de tamaño que usted defina. El tamaño predeterminado del archivo es de 50 MB. El tamaño máximo del archivo debe ser superior a 0 MB e inferior o igual a 50 MB.
type	El tipo del origen de datos. Especifica `CONFLUENCEV2` como el tipo de origen de datos.

Configuración	Descripción
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a Confluence. Para obtener información sobre estos pares clave-valor, consulta las instrucciones de conexión de Confluence.
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Confluence

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "hostUrl": {
              "type": "string",
              "pattern": "https:.*"
            },
            "type": {
              "type": "string",
              "enum": [
                "SAAS",
                "ON_PREM"
              ]
            },
            "authType": {
              "type": "string",
              "enum": [
                "Basic",
                "OAuth2",
                "Personal-token"
```

```
            ]
          }
        },
        "required": [
          "hostUrl",
          "type",
          "authType"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "space": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
```

```
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "page": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
```

```
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "blog": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
```

```
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "comment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
```

```
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "attachment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
```

```
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "usersAclS3FilePath": {
        "type": "string"
      },
      "isCrawlAcl": {
        "type": "boolean"
      },
      "fieldForUserId": {
        "type": "string"
      },
      "inclusionSpaceKeyFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionSpaceKeyFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "pageTitleRegEX": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "blogTitleRegEX": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
```

```
        "commentTitleRegEX": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "attachmentTitleRegEX": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "isCrawlPersonalSpace": {
          "type": "boolean"
        },
        "isCrawlArchivedSpace": {
          "type": "boolean"
        },
        "isCrawlArchivedPage": {
          "type": "boolean"
        },
        "isCrawlPage": {
          "type": "boolean"
        },
        "isCrawlBlog": {
          "type": "boolean"
        },
        "isCrawlPageComment": {
          "type": "boolean"
        },
        "isCrawlPageAttachment": {
          "type": "boolean"
        },
        "isCrawlBlogComment": {
          "type": "boolean"
        },
        "isCrawlBlogAttachment": {
          "type": "boolean"
        },
        "maxFileSizeInMegaBytes":  {
          "type":"string"
        },
        "inclusionFileTypePatterns": {
          "type": "array",
```

```
          "items": {
            "type": "string"
          }
        },
        "exclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionUrlPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionUrlPatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "proxyHost": {
          "type": "string"
        },
        "proxyPort": {
          "type": "string"
        }
      },
      "required": []
    },
    "type": {
      "type": "string",
      "pattern": "CONFLUENCEV2"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FULL_CRAWL",
        "FORCED_FULL_CRAWL"
      ]
```

```
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Esquema de plantilla de Dropbox

Incluyes un JSON que contiene el esquema de la fuente de datos como parte del objeto. TemplateConfiguration Proporciona la clave de la aplicación, el secreto de la aplicación y el token de acceso de Dropbox como parte del secreto que almacena sus credenciales de autenticación. Especifica también el tipo de origen de datos como `DROPBOX`, el tipo de token de acceso que quiere usar (temporal o permanente) y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Dropbox.

En la siguiente tabla se describen los parámetros del esquema JSON de Dropbox.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos. Este origen de datos no especifica un punto de conexión en `repositoryEndpointMetadata` . Por el contrario, la información de conexión se incluye en un AWS Secrets Manager secreto que tú proporcionas`secretArn` .
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
• archivo • paper • papert • shortcut	Una lista de objetos que mapean los atributos o los nombres de campo de tus archivos de Dropbox, Dropbox Paper y atajos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
syncMode	Especifica cómo Amazon Kendra debes actualizar tu índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic

Configuración	Descripción
	e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

Configuración	Descripción
secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectars e a tu Dropbox. El secreto debe contener una estructura JSON con las siguientes claves: `{ "appKey": "Dropbox app key", "appSecret": " Dropbox app secret", "accesstoken": " temporary access token or refresh access token" }`
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
isCrawlAcl	`true`para rastrear la información de la lista de control de acceso (ACL) de tus documentos, si tienes una ACL y quieres usarla para controlar el acceso. La ACL especifica a qué documento s pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

Configuración	Descripción
• inclusionFileNamePatrones • inclusionFileTypePatrones	Una lista de patrones de expresión regular para incluir determinados nombres y tipos de archivo en su origen de datos de Dropbox. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
• exclusionFileNamePatrones • exclusionFileTypePatrones	Una lista de patrones de expresión regular para excluir determinados nombres y tipos de archivo en su origen de datos de Dropbox. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
• crawlFile • crawlPaper • crawlPapert • crawlShortcut	`true`para rastrear los archivos de tu Dropbox, los documentos de Dropbox Paper, las plantilla s de Dropbox Paper y los atajos de páginas web almacenados en tu Dropbox.
type	El tipo del origen de datos. Especifica `DROPBOX` como el tipo de origen de datos.

Configuración	Descripción
tokenType	Especifica el tipo de token de acceso: token de acceso permanente o temporal. Se recomiend a crear un token de acceso actualizado que no caduque nunca en Dropbox, en lugar de utilizar un token de acceso único que caduca a las 4 horas. Debe crear una aplicación y un token de acceso de actualización en la consola para desarrolladores de Dropbox y proporcionar el token de acceso en su secreto.
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Dropbox

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
          }
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "file": {
          "type": "object",
          "properties": {
```

```
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": [
                          "STRING",
                          "STRING_LIST",
                          "LONG",
                          "DATE"
                        ]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "paper": {
          "type": "object",
          "properties": {
```

```
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": [
                          "STRING",
                          "STRING_LIST",
                          "LONG",
                          "DATE"
                        ]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "papert": {
          "type": "object",
          "properties": {
```

```
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": [
                          "STRING",
                          "STRING_LIST",
                          "LONG",
                          "DATE"
                        ]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "shortcut": {
          "type": "object",
          "properties": {
```

```
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": [
                          "STRING",
                          "STRING_LIST",
                          "LONG",
                          "DATE"
                        ]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        }
      }
    },
    "syncMode": {
```

```
      "type": "string",
      "enum": [
        "FULL_CRAWL",
        "FORCED_FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "secretArn": {
      "type": "string"
    },
    "additionalProperties": {
      "type": "object",
      "properties": {
        "isCrawlAcl": {
          "type": "boolean"
        },
        "inclusionFileNamePatterns": {
          "type": "array"
        },
        "exclusionFileNamePatterns": {
          "type": "array"
        },
        "inclusionFileTypePatterns": {
          "type": "array"
        },
        "exclusionFileTypePatterns": {
          "type": "array"
        },
        "crawlFile": {
          "type": "boolean"
        },
        "crawlPaper": {
          "type": "boolean"
        },
        "crawlPapert": {
          "type": "boolean"
        },
        "crawlShortcut": {
          "type": "boolean"
        }
      }
```

```
    },
    "type": {
      "type": "string",
      "pattern": "DROPBOX"
    },
    "tokenType": {
      "type": "string",
      "enum": [
        "PERMANENT",
        "TEMPORARY"
      ]
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "additionalProperties": false,
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "syncMode",
    "enableIdentityCrawler",
    "secretArn",
    "type",
    "tokenType"
  ]
}
```

Esquema de plantilla de Drupal

Como parte del objeto, incluyes un JSON que contiene el esquema de la TemplateConfigurationfuente de datos. Debe proporcionar la URL del host de Drupal y el tipo de autenticación como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Especifique también el tipo de origen de datos como DRUPAL, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Drupal.

La siguiente tabla describe los parámetros del esquema JSON de Drupal.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
hostUrl	La URL del host de su sitio web de Drupal. Por ejemplo, *https://<hostname>/<drupalsitename>* .
repositoryConfigurations	Información de configuración del contenido del origen de datos.
• content • comentario • attachment	Una lista de objetos que asignan los atributos o los nombres de campo de sus archivos de Drupal. Para obtener más información, consulte Asignación de campos de origen de datos. Los nombres de los campos del origen de datos de Drupal deben existir en los metadatos personali zados de Drupal.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
• inclusionFileNamePatrones • articleTitleInclusionPatrones • pageTitleInclusionPatrones • customContentTitleInclusionPatterns • basicBlockTitleInclusionPatterns • customBlockTitleInclusionPatterns	Una lista de patrones de expresión regular para incluir determinados archivos en su origen de datos de Drupal. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón

Configuración	Descripción
	de exclusión tiene prioridad y el archivo no se incluye en el índice.
• exclusionFileNamePatrones • articleTitleExclusionPatrones • pageTitleExclusionPatrones • customContentTitleExclusionPatterns • basicBlockTitleExclusionPatterns • customBlockTitleExclusionPatterns	Una lista de patrones de expresión regular para excluir determinados archivos en su origen de datos de Drupal. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
contentDefinitions • contentType • fieldDefinition • isCrawlComments • isCrawlFiles • isCrawlArticle • isCrawlBasicPágina • isCrawlBasicBloquear • isCrawlCustomContentTypesList	Especifica los tipos de contenido que desea rastrear y si desea rastrear los comentarios y los archivos adjuntos de los tipos de contenido seleccionados.
type	El tipo del origen de datos. Especifica `DRUPAL` como el tipo de origen de datos.
authType	El tipo de autenticación que utiliza, ya sea `BASIC-AUTH` o `OAUTH2`.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectars e a tu Drupal. El secreto debe contener una estructura JSON con las siguientes claves:
	
	Si utiliza la autenticación básica:
	
	`{ "username": "user name", "passwords": "password" }`
	
	Si utiliza la autenticación OAuth 2.0:
	
	`{ "username": "user name", "password": "password" , "clientId": "client id", "clientSecret": "client secret" }`

versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Drupal

```
{
 "$schema": "http://json-schema.org/draft-04/schema#",
 "type": "object",
 "properties": {
  "connectionConfiguration": {
   "type": "object",
   "properties": {
    "repositoryEndpointMetadata": {
     "type": "object",
     "properties": {
      "hostUrl": {
       "type": "string",
       "pattern": "https:.*"
      }
     },
     "required": [
      "hostUrl"
     ]
    }
   },
   "required": [
    "repositoryEndpointMetadata"
   ]
  },
  "repositoryConfigurations": {
   "type": "object",
   "properties": {
    "content": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
```

```
            "STRING",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
    "comment": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "DATE"
           ]
          },
```

```
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
    "attachment": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
```

```
          "type": "string",
          "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
         }
        },
        "required": [
         "indexFieldName",
         "indexFieldType",
         "dataSourceFieldName"
        ]
       }
      ]
     }
    },
    "required": [
     "fieldMappings"
    ]
   }
  }
 },
 "additionalProperties": {
  "type": "object",
  "properties": {
   "isCrawlArticle": {
    "type": "boolean"
   },
   "isCrawlBasicPage": {
    "type": "boolean"
   },
   "isCrawlBasicBlock": {
    "type": "boolean"
   },
   "crawlCustomContentTypesList": {
    "type": "array",
    "items": {
     "type": "string"
    }
   },
   "crawlCustomBlockTypesList": {
    "type": "array",
    "items": {
     "type": "string"
    }
   },
   "filePath": {
```

```
      "anyOf": [
       {
        "type": "string",
        "pattern": "s3:.*"
       },
       {
        "type": "string",
        "pattern": ""
       }
      ]
     },
     "inclusionFileNamePatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "exclusionFileNamePatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "articleTitleInclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "articleTitleExclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "pageTitleInclusionPatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "pageTitleExclusionPatterns": {
      "type": "array",
      "items": {
```

```
      "type": "string"
     }
    },
    "customContentTitleInclusionPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "customContentTitleExclusionPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "basicBlockTitleInclusionPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "basicBlockTitleExclusionPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "customBlockTitleInclusionPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "customBlockTitleExclusionPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "contentDefinitions": {
     "type": "array",
     "items": {
      "properties": {
       "contentType": {
```

```
        "type": "string"
       },
       "fieldDefinition": {
        "type": "array",
        "items": [
         {
          "type": "object",
          "properties": {
           "machineName": {
            "type": "string"
           },
           "type": {
            "type": "string"
           }
          },
          "required": [
           "machineName",
           "type"
          ]
         }
        ]
       },
       "isCrawlComments": {
        "type": "boolean"
       },
       "isCrawlFiles": {
        "type": "boolean"
       }
      }
     },
     "required": [
      "contentType",
      "fieldDefinition",
      "isCrawlComments",
      "isCrawlFiles"
     ]
    }
   },
   "required": []
  },
  "type": {
   "type": "string",
   "pattern": "DRUPAL"
  },
```

```
  "authType": {
   "type": "string",
   "enum": [
    "BASIC-AUTH",
    "OAUTH2"
   ]
  },
  "syncMode": {
   "type": "string",
   "enum": [
    "FORCED_FULL_CRAWL",
    "FULL_CRAWL",
    "CHANGE_LOG"
   ]
  },
  "enableIdentityCrawler": {
   "type": "boolean"
  },
  "secretArn": {
   "type": "string",
   "minLength": 20,
   "maxLength": 2048
  }
 },
 "version": {
  "type": "string",
  "anyOf": [
   {
    "pattern": "1.0.0"
   }
  ]
 },
 "required": [
  "connectionConfiguration",
  "repositoryConfigurations",
  "syncMode",
  "additionalProperties",
  "secretArn",
  "type"
 ]
}
```

GitHub esquema de plantilla

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Debe proporcionar la URL del GitHub host, el nombre de la organización y si utiliza la GitHub GitHub nube o de forma local como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Además, especifique el tipo de origen de datos como `GITHUB`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE Type` especifíquelo cuando llame CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte GitHub Esquema JSON.

En la siguiente tabla se describen los parámetros del esquema GitHub JSON.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
type	Especifique el tipo como `SAAS` o`ON_PREMIS E` .
hostUrl	La URL del GitHub host. Por ejemplo, si utiliza GitHub SaaS/Enterprise Cloud:. https://a pi.github.com O, si usa un servidor GitHub local o empresarial:. https://on-prem-host-url/api/v3/
organizationName	Puede encontrar el nombre de su organizac ión al iniciar sesión en el GitHub escritorio y ir a Sus organizaciones en el menú desplegable de su imagen de perfil.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.

Configuración	Descripción
• Repositorio GH • ghCommit • ghIssueDocument • ghIssueComment • ghIssueAttachment • Documento GHPR • Comentario GHPR • Adjunto GHPR	Una lista de objetos que asignan los atributos o los nombres de campo del GitHub contenido para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
isCrawlAcl	`true`para rastrear la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica los documento s a los que los usuarios y los grupos pueden acceder y buscar. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más informaci ón, consulte Filtrado de contexto de usuario.
fieldForUserID	Especifique el tipo de ID de usuario que desea usar para el rastreo de las ACL. Especifique `email` si desea usar el correo electrónico del usuario como ID de usuario o `username` si desea usar el nombre de usuario como ID de usuario. Si no especifica ninguna opción, `email` se utiliza de forma predeterminada.
Filtro de repositorio	Una lista de los nombres de los repositorios y ramas específicos que quieres indexar.

Configuración	Descripción
CrawlRepository	`true`para rastrear repositorios.
crawlRepositoryDocuments	`true`para rastrear los documentos del repositor io.
Problema de rastreo	`true`a problemas de rastreo.
crawlIssueComment	`true`para rastrear los comentarios de los problemas.
crawlIssueCommentAdjunto	`true`para rastrear los archivos adjuntos a los comentarios de un problema.
crawlPullRequest	`true`para rastrear las solicitudes de incorpora ción de cambios.
crawlPullRequestComentario	`true`para rastrear los comentarios de las solicitudes de extracción.
crawlPullRequestCommentAttachment	`true`para rastrear los archivos adjuntos de los comentarios de las solicitudes de extracción.
• inclusionFolderNamePatrones • inclusionFileTypePatrones • inclusionFileNamePatrones	Una lista de patrones de expresiones regulares para incluir cierto contenido en la fuente GitHub de datos. El contenido que coincide con los patrones se incluye en el índice. El contenido que no coincide con los patrones se excluye del índice. Si algún contenido coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.

Configuración	Descripción
• exclusionFolderNamePatrones • exclusionFileTypePatrones • exclusionFileNamePatrones	Una lista de patrones de expresiones regulares para excluir cierto contenido de la fuente GitHub de datos. El contenido que coincide con los patrones se excluye del índice. El contenido que no coincide con los patrones se incluye en el índice. Si algún contenido coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.
type	El tipo del origen de datos. Especifica `GITHUB` como el tipo de origen de datos.
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a su. GitHub El secreto debe contener una estructura JSON con las siguientes claves: `{ "personalToken": " token" }`
versión	La versión de esta plantilla que se admite actualmente.

GitHub Esquema JSON

El siguiente es el esquema GitHub JSON:

```
{
    "$schema": "http://json-schema.org/draft-04/schema#",
    "type": "object",
    "properties": {
        "connectionConfiguration": {
            "type": "object",
            "properties": {
                "repositoryEndpointMetadata": {
                    "type": "object",
                    "properties": {
                        "type": {
                            "type": "string"
                        },
                        "hostUrl": {
                            "type": "string",
                            "pattern": "https://.*"
                        },
                        "organizationName": {
                            "type": "string"
                        }
                    },
```

```
                "required": [
                    "type",
                    "hostUrl",
                    "organizationName"
                ]
            }
        },
        "required": [
            "repositoryEndpointMetadata"
        ]
    },
    "repositoryConfigurations": {
        "type": "object",
        "properties": {
            "ghRepository": {
                "type": "object",
                "properties": {
                    "fieldMappings": {
                        "type": "array",
                        "items": [
                            {
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": [
                                            "STRING",
                                            "STRING_LIST",
                                            "DATE"
                                        ]
                                    },
                                    "dataSourceFieldName": {
                                        "type": "string"
                                    },
                                    "dateFieldFormat": {
                                        "type": "string",
                                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                    }
                                },
                                "required": [
                                    "indexFieldName",
```

```
                                "indexFieldType",
                                "dataSourceFieldName"
                            ]
                        }
                    ]
                }
            },
            "required": [
                "fieldMappings"
            ]
        },
        "ghCommit": {
            "type": "object",
            "properties": {
                "fieldMappings": {
                    "type": "array",
                    "items": [
                        {
                            "type": "object",
                            "properties": {
                                "indexFieldName": {
                                    "type": "string"
                                },
                                "indexFieldType": {
                                    "type": "string",
                                    "enum": [
                                        "STRING",
                                        "STRING_LIST",
                                        "DATE"
                                    ]
                                },
                                "dataSourceFieldName": {
                                    "type": "string"
                                },
                                "dateFieldFormat": {
                                    "type": "string",
                                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                }
                            },
                            "required": [
                                "indexFieldName",
                                "indexFieldType",
                                "dataSourceFieldName"
                            ]
```

```
                        }
                    ]
                }
            },
            "required": [
                "fieldMappings"
            ]
        },
        "ghIssueDocument": {
            "type": "object",
            "properties": {
                "fieldMappings": {
                    "type": "array",
                    "items": [
                        {
                            "type": "object",
                            "properties": {
                                "indexFieldName": {
                                    "type": "string"
                                },
                                "indexFieldType": {
                                    "type": "string",
                                    "enum": [
                                        "STRING",
                                        "STRING_LIST",
                                        "DATE"
                                    ]
                                },
                                "dataSourceFieldName": {
                                    "type": "string"
                                },
                                "dateFieldFormat": {
                                    "type": "string",
                                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                }
                            },
                            "required": [
                                "indexFieldName",
                                "indexFieldType",
                                "dataSourceFieldName"
                            ]
                        }
                    ]
                }
```

```
                    },
                    "required": [
                        "fieldMappings"
                    ]
                },
                "ghIssueComment": {
                    "type": "object",
                    "properties": {
                        "fieldMappings": {
                            "type": "array",
                            "items": [
                                {
                                    "type": "object",
                                    "properties": {
                                        "indexFieldName": {
                                            "type": "string"
                                        },
                                        "indexFieldType": {
                                            "type": "string",
                                            "enum": [
                                                "STRING",
                                                "STRING_LIST",
                                                "DATE"
                                            ]
                                        },
                                        "dataSourceFieldName": {
                                            "type": "string"
                                        },
                                        "dateFieldFormat": {
                                            "type": "string",
                                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                        }
                                    },
                                    "required": [
                                        "indexFieldName",
                                        "indexFieldType",
                                        "dataSourceFieldName"
                                    ]
                                }
                            ]
                        }
                    },
                    "required": [
                        "fieldMappings"
```

```
                    ]
                },
                "ghIssueAttachment": {
                    "type": "object",
                    "properties": {
                        "fieldMappings": {
                            "type": "array",
                            "items": [
                                {
                                    "type": "object",
                                    "properties": {
                                        "indexFieldName": {
                                            "type": "string"
                                        },
                                        "indexFieldType": {
                                            "type": "string",
                                            "enum": [
                                                "STRING",
                                                "STRING_LIST",
                                                "DATE"
                                            ]
                                        },
                                        "dataSourceFieldName": {
                                            "type": "string"
                                        },
                                        "dateFieldFormat": {
                                            "type": "string",
                                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                        }
                                    },
                                    "required": [
                                        "indexFieldName",
                                        "indexFieldType",
                                        "dataSourceFieldName"
                                    ]
                                }
                            ]
                        }
                    },
                    "required": [
                        "fieldMappings"
                    ]
                },
                "ghPRDocument": {
```

```
                    "type": "object",
                    "properties": {
                        "fieldMappings": {
                            "type": "array",
                            "items": [
                                {
                                    "type": "object",
                                    "properties": {
                                        "indexFieldName": {
                                            "type": "string"
                                        },
                                        "indexFieldType": {
                                            "type": "string",
                                            "enum": [
                                                "STRING",
                                                "STRING_LIST",
                                                "DATE"
                                            ]
                                        },
                                        "dataSourceFieldName": {
                                            "type": "string"
                                        },
                                        "dateFieldFormat": {
                                            "type": "string",
                                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                        }
                                    },
                                    "required": [
                                        "indexFieldName",
                                        "indexFieldType",
                                        "dataSourceFieldName"
                                    ]
                                }
                            ]
                        }
                    },
                    "required": [
                        "fieldMappings"
                    ]
                },
                "ghPRComment": {
                    "type": "object",
                    "properties": {
                        "fieldMappings": {
```

```
                        "type": "array",
                        "items": [
                            {
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": [
                                            "STRING",
                                            "STRING_LIST",
                                            "DATE"
                                        ]
                                    },
                                    "dataSourceFieldName": {
                                        "type": "string"
                                    },
                                    "dateFieldFormat": {
                                        "type": "string",
                                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                    }
                                },
                                "required": [
                                    "indexFieldName",
                                    "indexFieldType",
                                    "dataSourceFieldName"
                                ]
                            }
                        ]
                    }
                },
                "required": [
                    "fieldMappings"
                ]
            },
            "ghPRAttachment": {
                "type": "object",
                "properties": {
                    "fieldMappings": {
                        "type": "array",
                        "items": [
                            {
```

```
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": [
                                            "STRING",
                                            "STRING_LIST",
                                            "DATE"
                                        ]
                                    },
                                    "dataSourceFieldName": {
                                        "type": "string"
                                    },
                                    "dateFieldFormat": {
                                        "type": "string",
                                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                    }
                                },
                                "required": [
                                    "indexFieldName",
                                    "indexFieldType",
                                    "dataSourceFieldName"
                                ]
                            }
                        ]
                    }
                },
                "required": [
                    "fieldMappings"
                ]
            }
        }
    },
    "additionalProperties": {
        "type": "object",
        "properties": {
            "isCrawlAcl": {
                "type": "boolean"
            },
            "fieldForUserId": {
                "type": "string"
```

```
                },
                "crawlRepository": {
                    "type": "boolean"
                },
                "crawlRepositoryDocuments": {
                    "type": "boolean"
                },
                "crawlIssue": {
                    "type": "boolean"
                },
                "crawlIssueComment": {
                    "type": "boolean"
                },
                "crawlIssueCommentAttachment": {
                    "type": "boolean"
                },
                "crawlPullRequest": {
                    "type": "boolean"
                },
                "crawlPullRequestComment": {
                    "type": "boolean"
                },
                "crawlPullRequestCommentAttachment": {
                    "type": "boolean"
                },
                "repositoryFilter": {
                    "type": "array",
                    "items": [
                        {
                            "type": "object",
                            "properties": {
                                "repositoryName": {
                                    "type": "string"
                                },
                                "branchNameList": {
                                    "type": "array",
                                    "items": {
                                        "type": "string"
                                    }
                                }
                            }
                        }
                    ]
                },
```

```
                "inclusionFolderNamePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "inclusionFileTypePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "inclusionFileNamePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "exclusionFolderNamePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "exclusionFileTypePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                },
                "exclusionFileNamePatterns": {
                    "type": "array",
                    "items": {
                        "type": "string"
                    }
                }
            },
            "required": []
        },
        "type": {
            "type": "string",
            "pattern": "GITHUB"
        },
        "syncMode": {
```

```
            "type": "string",
            "enum": [
                "FULL_CRAWL",
                "FORCED_FULL_CRAWL",
                "CHANGE_LOG"
            ]
        },
        "enableIdentityCrawler": {
            "type": "boolean"
        },
        "secretArn": {
            "type": "string",
            "minLength": 20,
            "maxLength": 2048
        }
    },
    "version": {
        "type": "string",
        "anyOf": [
            {
                "pattern": "1.0.0"
            }
        ]
    },
    "required": [
        "connectionConfiguration",
        "repositoryConfigurations",
        "syncMode",
        "additionalProperties",
        "enableIdentityCrawler"
    ]
}
```

Esquema de plantilla de Gmail

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `GMAIL`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Gmail.

En la siguiente tabla se describen los parámetros del esquema JSON de Gmail.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos. Este origen de datos no especifica un punto de conexión en `repositoryEndpointMetadata` . Por el contrario, la información de conexión se incluye en un AWS Secrets Manager secreto que usted proporciona`secretArn` .
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
• message • attachments	Una lista de objetos que asignan los atributos o los nombres de campo de tus mensajes y archivos adjuntos de Gmail para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
• inclusionLabelNamePatrones • exclusionLabelNamePatrones • inclusionAttachmentTypePatrones • exclusionAttachmentTypePatrones • inclusionAttachmentNamePatrones • exclusionAttachmentNamePatrones	Una lista de patrones de expresión regular para incluir o excluir mensajes con nombres de asuntos específicos en su origen de datos de Gmail. Los archivos que coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de inclusión

Configuración	Descripción
• inclusionSubjectFilter • exclusionSubjectFilter • isSubjectAnd • inclusionFromFilter • exclusionFromFilter • inclusionToFilter • exclusionToFilter • inclusionCcFilter • exclusionCcFilter • inclusionBccFilter • exclusionBccFilter	y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
beforeDateFilter	Especifica los mensajes y archivos adjuntos que se incluirán antes de una fecha determina da.
afterDateFilter	Especifica los mensajes y archivos adjuntos que se incluirán después de una fecha determinada.
isCrawlAttachment	Un valor booleano para elegir si desea rastrear los archivos adjuntos. Los mensajes se rastrean automáticamente.
type	El tipo del origen de datos. Especifica `GMAIL` como el tipo de origen de datos.
shouldCrawlDraftMensajes	Un valor booleano para elegir si desea rastrear los borradores de mensajes.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. **Important** Como no existe una API para actualizar los mensajes de Gmail eliminados permanentemente, cualquier contenido nuevo, modificado o eliminado se sincroniza: • No eliminará de tu Amazon Kendra índice los mensajes que se hayan eliminado permanentemente de Gmail • No sincronizará los cambios en las etiquetas de correo de Gmail

Configuración	Descripción
	Para sincronizar los cambios en la etiqueta de la fuente de datos de Gmail y los mensajes de correo electrónico eliminados permanentemente con tu Amazon Kendra índice, debes realizar rastreos completos de forma periódica.
secretARN	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene los pares clave/valor necesarios para conectars e a su Gmail. El secreto debe contener una estructura JSON con las siguientes claves: `{ "adminAccountEmailId": " service account email", "clientEmailId": " user account email", "privateKey": " private key" }`
versión	La versión de la plantilla que se admite actualmente.

Esquema JSON de Gmail

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
      }
    },
```

```
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "message": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "attachments": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
```

```
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        }
      }
    },
    "required": []
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "inclusionLabelNamePatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionLabelNamePatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionAttachmentTypePatterns": {
        "type": "array",
        "items": {
          "type": "string"
```

```
        }
      },
      "exclusionAttachmentTypePatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionAttachmentNamePatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionAttachmentNamePatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionSubjectFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionSubjectFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "isSubjectAnd": {
        "type": "boolean"
      },
      "inclusionFromFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionFromFilter": {
        "type": "array",
        "items": {
```

```
            "type": "string"
          }
        },
        "inclusionToFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionToFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionCcFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionCcFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionBccFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionBccFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "beforeDateFilter": {
          "anyOf": [
            {
              "type": "string",
              "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
```

```
            },
            {
              "type": "string",
              "pattern": ""
            }
          ]
        },
        "afterDateFilter": {
          "anyOf": [
            {
              "type": "string",
              "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
            },
            {
              "type": "string",
              "pattern": ""
            }
          ]
        },
        "isCrawlAttachment": {
          "type": "boolean"
        },
        "shouldCrawlDraftMessages": {
          "type": "boolean"
        }
      },
      "required": [
        "isCrawlAttachment",
        "shouldCrawlDraftMessages"
      ]
    },
    "type" : {
      "type" : "string",
      "pattern": "GMAIL"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL"
      ]
    },
    "secretArn": {
      "type": "string"
```

```
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "syncMode",
    "secretArn",
    "type"
  ]
}
```

Esquema de plantilla de Google Drive

Incluyes un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `GOOGLEDRIVE2`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Google Drive.

En la siguiente tabla se describen los parámetros del esquema JSON de Google Drive.

Configuración	Descripción
connectionConfiguration	Información de configuración del origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos. Este origen de datos no especifica un punto de conexión. Usted elige el tipo de autenticación: `serviceAccount` y

Configuración	Descripción
	`0Auth2`. La información de conexión se incluye en un AWS Secrets Manager secreto que usted proporciona`secretArn` .
authType	Elija entre `serviceAccount` y `0Auth2` en función de su caso de uso.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
• archivo • comentario	Una lista de objetos que asignan los atributos o los nombres de campo de su Google Drive a los nombres de campo del índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos
• maxFileSizeInMegaBytes	Especifique un límite de tamaño de archivo en MB que Amazon Kendra debe rastrearse.
• iscrawlComment	`true`para rastrear los comentarios de su fuente de datos de Google Drive.
• isCrawlMyDriveAndSharedWithMe	`true`para rastrear MyDrive y compartir conmigo las unidades de tu fuente de datos de Google Drive.
• isCrawlSharedUnidades	`true`para rastrear las unidades compartidas de tu fuente de datos de Google Drive.

Configuración	Descripción
isCrawlAcl	`true`para rastrear la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y quiere usarla para el control de acceso. La ACL especifica los documento s a los que los usuarios y los grupos pueden acceder y buscar. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más informaci ón, consulte Filtrado de contexto de usuario.
• excludeUserAccounts • excludeSharedDrives • excludeMimeTypes • exclusionFileTypePatrones • exclusionFileNamePatrones • exclusionFilePathFiltro	Una lista de patrones de expresión regular para excluir determinados archivos en su origen de datos de Google Drive. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
• includeUserAccounts • includeSharedDrives • includeMimeTypes • inclusionFileTypePatrones • inclusionFileNamePatrones • inclusionFilePathFiltro	Una lista de patrones de expresión regular para incluir determinados archivos en su origen de datos de Google Drive. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
type	El tipo del origen de datos. Especifica `GOOGLEDRIVEV2` como el tipo de origen de datos.

Configuración	Descripción
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción

secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a Google Drive. El secreto debe contener una estructura JSON con las siguientes claves:
	
	Si utiliza la autenticación de la cuenta de servicio de Google:
	
	`{ "clientEmail": " user account email", "adminAccountEmail": " service account email", "privateKey": " private key" }`
	
	Si utiliza la autenticación OAuth 2.0:
	
	`{ "clientID": " OAuth client ID", "clientSecret": " client secret", "refreshToken": " refresh token" }`

versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Google Drive

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
```

```
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "authType": {
              "type": "string",
              "enum": [
                "serviceAccount",
                "OAuth2"
              ]
            }
          },
          "required": [
            "authType"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "file": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "DATE",
                        "STRING_LIST",
                        "LONG"
                      ]
                    },
```

```
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "comment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "DATE",
                    "STRING_LIST"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
```

```
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "maxFileSizeInMegaBytes": {
        "type": "string"
      },
      "isCrawlComment": {
        "type": "boolean"
      },
      "isCrawlMyDriveAndSharedWithMe": {
        "type": "boolean"
      },
      "isCrawlSharedDrives": {
        "type": "boolean"
      },
      "isCrawlAcl": {
        "type": "boolean"
      },
      "excludeUserAccounts": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
```

```
        "excludeSharedDrives": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "excludeMimeTypes": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "includeUserAccounts": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "includeSharedDrives": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "includeMimeTypes": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "includeTargetAudienceGroup": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileNamePatterns": {
          "type": "array",
```

```
          "items": {
            "type": "string"
          }
        },
        "exclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFilePathFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFilePathFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        }
      }
    },
    "type": {
      "type": "string",
      "pattern": "GOOGLEDRIVEV2"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
```

```
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Esquema de plantilla de IBM DB2

Incluyes un JSON que contiene el esquema de la fuente de datos como parte del objeto. TemplateConfiguration Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como db2, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de IBM DB2.

La siguiente tabla describe los parámetros del esquema JSON de IBM DB2.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.

Configuración	Descripción
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea `mysql`, `db2` `postgresql` , `oracle` o. `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.

Configuración	Descripción
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.

Configuración	Descripción
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Esquema JSON de IBM DB2

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Esquema de plantilla de Microsoft Exchange

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Debe proporcionar el ID de inquilino como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Además, especifique el tipo de origen de datos como `MSEXCHANGE`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Microsoft Exchange.

En la siguiente tabla se describen los parámetros del esquema JSON de Microsoft Exchange.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
tenantId	El ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos

Configuración	Descripción
	específicos de contenido y asignaciones de campos.
• email • attachment • calendar • contacts • notes	Una lista de objetos que asignan los atributos o nombres de campo de la fuente de datos de Microsoft Exchange a los campos de Amazon Kendra indexación. Para obtener más informaci ón, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos
inclusionPatterns	Una lista de patrones de expresión regular para incluir determinados archivos en su origen de datos de Microsoft Exchange. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
exclusionPatterns	Una lista de patrones de expresión regular para excluir determinados archivos en su origen de datos de Microsoft Exchange. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.

Configuración	Descripción
• inclusionUsersList • inclusionUsersFileNombre • inclusionDomainUsers	Una lista de patrones de expresión regular para incluir determinados usuarios y archivos de usuario en su origen de datos de Microsoft Exchange. Los usuarios que coinciden con los patrones se incluyen en el índice. Los usuarios que no coinciden con los patrones se excluyen del índice. Si un usuario coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el usuario no se incluye en el índice.
• exclusionUsersList • exclusionUsersFileNombre • exclusionDomainUsers	Una lista de patrones de expresión regular para excluir determinados usuarios y archivos de usuario en su origen de datos de Microsoft Exchange. Los usuarios que coinciden con los patrones se excluyen del índice. Los usuarios que no coinciden con los patrones se incluyen en el índice. Si un usuario coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el usuario no se incluye en el índice.
s3bucketName	El nombre del bucket de S3, si lo desea usar.
• crawlCalendar • crawlNotes • crawlContacts • crawlFolderAcl	`true`para rastrear estos tipos de contenido e información de control de acceso a su fuente de datos de Microsoft Exchange.
startCalendarDateHora	Puede configurar una fecha y hora de inicio específicas para el contenido de su calendario.
endCalendarDateHora	Puede configurar una fecha y hora de finalizac ión específicas para el contenido del calendari o.

Configuración	Descripción
subject	Puede configurar una línea de asunto específica para el contenido de su correo.
emailFrom	Puede configurar un correo electrónico específico para el contenido del correo del remitente.
emailTo	Puede configurar un correo electrónico específico para el contenido del correo del destinatario.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
type	El tipo del origen de datos. Especifica `MSEXCHANGE` como el tipo de origen de datos.

Configuración	Descripción
secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectars e a Microsoft Exchange. Esto incluye su ID de cliente y su secreto de cliente, que se genera al crear una aplicación OAuth en el portal de Azure.
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Microsoft Exchange

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "tenantId": {
              "type": "string",
              "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]
{12}$",
              "minLength": 36,
              "maxLength": 36
            }
          },
          "required": ["tenantId"]
        }
      }
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
```

```
        "email": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "attachment": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
```

```
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "DATE","LONG"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "calendar": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
```

```
                    "enum": ["STRING", "STRING_LIST", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "contacts": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "STRING_LIST", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
```

```
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "notes": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
```

```
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": ["email"
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "inclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionUsersList": {
        "type": "array",
        "items": {
          "type": "string",
          "format": "email"
        }
      },
      "exclusionUsersList": {
        "type": "array",
        "items": {
          "type": "string",
          "format": "email"
        }
      },
      "s3bucketName": {
        "type": "string"
```

```
        },
        "inclusionUsersFileName": {
          "type": "string"
        },
        "exclusionUsersFileName": {
          "type": "string"
        },
        "inclusionDomainUsers": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionDomainUsers": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "crawlCalendar": {
          "type": "boolean"
        },
        "crawlNotes": {
          "type": "boolean"
        },
        "crawlContacts": {
          "type": "boolean"
        },
        "crawlFolderAcl": {
          "type": "boolean"
        },
        "startCalendarDateTime": {
          "anyOf": [
            {
              "type": "string",
              "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
            },
            {
              "type": "string",
              "pattern": ""
            }
          ]
        },
        "endCalendarDateTime": {
```

```
          "anyOf": [
            {
            "type": "string",
            "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
            },
            {
              "type": "string",
              "pattern": ""
            }
          ]
        },
        "subject": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "emailFrom": {
          "type": "array",
          "items": {
            "type": "string",
            "format": "email"
          }
        },
        "emailTo": {
          "type": "array",
          "items": {
            "type": "string",
            "format": "email"
          }
        }
      },
      "required": [
      ]
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "type" : {
```

```
      "type" : "string",
      "pattern": "MSEXCHANGE"
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Esquema OneDrive de plantillas de Microsoft

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Debe proporcionar el ID de inquilino como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Especifique también el tipo de origen de datos como `ONEDRIVEV2`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema OneDrive JSON de Microsoft.

En la siguiente tabla se describen los parámetros del esquema OneDrive JSON de Microsoft.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
tenantId	El ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
archivo	Una lista de objetos que asignan los atributos o los nombres de campo de los OneDrive archivos de Microsoft a los nombres de los campos de Amazon Kendra indexación. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos
• userNameFilter • userFilterPath • inclusionFileTypePatrones • exclusionFileTypePatrones • inclusionFileNamePatrones • exclusionFileNamePatrones • inclusionFilePathPatrones • exclusionFilePathPatrones	Puede elegir indexar archivos, OneNote secciones y OneNote páginas específicos y filtrar por nombre de usuario.

Configuración	Descripción
• inclusionOneNoteSectionNamePatterns • exclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns • exclusionOneNotepageNamePatterns	
isUserNameEn S3	`true` para proporcionar una lista de nombres de usuario en un archivo almacenado en un Amazon S3.
type	El tipo del origen de datos. Especifica `ONEDRIVEV2` como el tipo de origen de datos.
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.
type	El tipo del origen de datos. Especifica `ONEDRIVEV2` como el tipo de origen de datos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a Microsoft. OneDrive El secreto debe contener una estructura JSON con las siguientes claves: `{ "clientId": " client ID", "clientSecret": " client secret" }`
versión	La versión de esta plantilla que se admite actualmente.

Esquema OneDrive JSON de Microsoft

```
{
 "$schema": "http://json-schema.org/draft-04/schema#",
 "type": "object",
 "properties": {
  "connectionConfiguration": {
   "type": "object",
   "properties": {
    "repositoryEndpointMetadata": {
     "type": "object",
     "properties": {
      "tenantId": {
       "type": "string",
       "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}$",
       "minLength": 36,
       "maxLength": 36
      }
     },
     "required": [
      "tenantId"
     ]
    }
```

```
    },
    "required": [
     "repositoryEndpointMetadata"
    ]
   },
   "repositoryConfigurations": {
    "type": "object",
    "properties": {
     "file": {
      "type": "object",
      "properties": {
       "fieldMappings": {
        "type": "array",
        "items": [
         {
          "type": "object",
          "properties": {
           "indexFieldName": {
            "type": "string"
           },
           "indexFieldType": {
            "type": "string",
            "enum": [
             "STRING",
             "STRING_LIST",
             "DATE",
             "LONG"
            ]
           },
           "dataSourceFieldName": {
            "type": "string"
           },
           "dateFieldFormat": {
            "type": "string",
            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
           }
          },
          "required": [
           "indexFieldName",
           "indexFieldType",
           "dataSourceFieldName"
          ]
         }
        ]
```

```
      }
     },
     "required": [
      "fieldMappings"
     ]
    }
   }
  },
  "additionalProperties": {
   "type": "object",
   "properties": {
    "userNameFilter": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "userFilterPath": {
     "type": "string"
    },
    "isUserNameOnS3": {
     "type": "boolean"
    },
    "inclusionFileTypePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionFileTypePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "inclusionFileNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionFileNamePatterns": {
     "type": "array",
     "items": {
```

```
      "type": "string"
     }
    },
    "inclusionFilePathPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionFilePathPatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "inclusionOneNoteSectionNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionOneNoteSectionNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "inclusionOneNotePageNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionOneNotePageNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    }
   },
   "required": []
  },

  "enableIdentityCrawler": {
```

```
    "type": "boolean"
   },
   "type": {
    "type": "string",
    "pattern": "ONEDRIVEV2"
   },
   "syncMode": {
    "type": "string",
    "enum": [
     "FULL_CRAWL",
     "FORCED_FULL_CRAWL",
     "CHANGE_LOG"
    ]
   },
   "secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
   }
  },
  "version": {
   "type": "string",
   "anyOf": [
    {
     "pattern": "1.0.0"
    }
   ]
  },
  "required": [
   "connectionConfiguration",
   "repositoryConfigurations",
   "syncMode",
   "additionalProperties",
   "secretArn",
   "type"
  ]
}
```

Esquema SharePoint de plantillas de Microsoft

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Como parte de la configuración de la conexión o de los detalles del punto final del repositorio, debe proporcionar la dirección URL o las direcciones URL del SharePoint

sitio, el dominio y, si es necesario, un ID de inquilino. Además, especifique el tipo de origen de datos como `SHAREPOINTV2`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` especifíquelo como tipo cuando llame. CreateDataSource

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte SharePoint Esquema JSON.

En la siguiente tabla se describen los parámetros del esquema SharePoint JSON de Microsoft.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos
tenantId	El identificador de inquilino de tu SharePoint cuenta.
Dominio	El dominio de tu SharePoint cuenta.
siteUrls	Las direcciones URL del servidor de tu SharePoint cuenta.
repositoryAdditionalProperties	Propiedades adicionales para conectarse con el punto de conexión del repositorio/origen de datos.
s3bucketName	El nombre del Amazon S3 depósito que almacena el certificado X.509 autofirmado de Azure AD.
s3certificateName	El nombre del certificado X.509 autofirmado de Azure AD almacenado en el depósito. Amazon S3
authType	El tipo de autenticación que usa,`0Auth2`,, `0Auth2Certificate 0Auth2App` ,

Configuración	Descripción
	`Basic` `OAuth2_RefreshToken` `NTLM`, o. `Kerberos`
versión	La SharePoint versión que utiliza, ya sea `Server` o`Online`.
onPremVersion	La versión SharePoint del servidor que utiliza, ya sea `2013` `20162019`, o`SubscriptionEdition` .
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
• evento • page • archivo • link • attachment • comentario	Una lista de objetos que asignan los atributos o los nombres de campo del SharePoint contenido para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.

Configuración	Descripción
• eventTitleFilterRegEx • pageTitleFilterRegEx • linkTitleFilterRegEx • inclusionFilePath • exclusionFilePath • inclusionFileTypePatrones • exclusionFileTypePatrones • inclusionFileNamePatrones • exclusionFileNamePatrones • inclusionOneNoteSectionNamePatterns • exclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns • exclusionOneNotePageNamePatterns	Una lista de patrones de expresiones regulares para incluir/excluir cierto contenido de la fuente de SharePoint datos. Los elementos de contenido que coinciden con los patrones de inclusión se incluyen en el índice. Los elementos de contenido que no coinciden con los patrones de inclusión se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
• crawlFiles • crawlPages • crawlEvents • crawlComments • crawlLinks • crawlAttachments	`true`para rastrear este tipo de contenido.
crawlAcl	`true`para rastrear la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica los documento s a los que los usuarios y los grupos pueden acceder y buscar. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más informaci ón, consulte Filtrado de contexto de usuario.

Configuración	Descripción
fieldForUserID	Especifique `email` si desea usar el correo electrónico del usuario como ID de usuario o `userPrincipalName` si desea usar un nombre de usuario para el ID de usuario. Si no especifica ninguna opción, `email` se utiliza de forma predeterminada.
aclConfiguration	Especifique una de `ACLWithLDAPEmailFmt` las dos `ACLWithManualEmailFmt` opciones: o`ACLWithUsernameFmtM` .
emailDomain	El dominio del correo electrónico. Por ejemplo, *`“amazon.com”`* . | • isCrawlLocalGroupMapping • isCrawlAdGroupMapping	`true`para rastrear la información de mapeo de grupos.
proxyHost	El nombre de host del proxy web que utiliza, sin el protocolo http://o https://.
proxyPort	El número de puerto utilizado por el protocolo de transporte de URL del host. Debe ser un valor numérico entre 0 y 65535.
type	Especifica `SHAREPOINTV2` como el tipo de origen de datos

Configuración	Descripción
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectars e a su. SharePoint Para obtener informaci ón sobre estos pares clave-valor, consulte las instrucciones de conexión en línea y en servidor. SharePoint SharePoint

Configuración	Descripción
versión	La versión de esta plantilla que se admite actualmente.

SharePoint Esquema JSON

```
{
 "$schema": "http://json-schema.org/draft-04/schema#",
 "type": "object",
 "properties": {
  "connectionConfiguration": {
   "type": "object",
   "properties": {
    "repositoryEndpointMetadata": {
     "type": "object",
     "properties": {
      "tenantId": {
       "type": "string",
       "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}$",
       "minLength": 36,
       "maxLength": 36
      },
      "domain": {
       "type": "string"
      },
      "siteUrls": {
       "type": "array",
       "items": {
        "type": "string",
        "pattern": "https://.*"
       }
      },
      "repositoryAdditionalProperties": {
       "type": "object",
       "properties": {
        "s3bucketName": {
         "type": "string"
        },
        "s3certificateName": {
         "type": "string"
        },
```

```
        "authType": {
         "type": "string",
         "enum": [
          "OAuth2",
          "OAuth2Certificate",
          "OAuth2App",
          "Basic",
          "OAuth2_RefreshToken",
          "NTLM",
          "Kerberos"
         ]
        },
        "version": {
         "type": "string",
         "enum": [
          "Server",
          "Online"
         ]
        },
        "onPremVersion": {
         "type": "string",
         "enum": [
          "",
          "2013",
          "2016",
          "2019",
          "SubscriptionEdition"
         ]
        }
       },
       "required": [
        "authType",
        "version"
       ]
      }
     },
     "required": [
      "siteUrls",
      "domain",
      "repositoryAdditionalProperties"
     ]
    }
   },
   "required": [
```

```
    "repositoryEndpointMetadata"
   ]
  },
  "repositoryConfigurations": {
   "type": "object",
   "properties": {
    "event": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "STRING_LIST",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
```

```
      "fieldMappings"
     ]
    },
    "page": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "DATE",
            "LONG"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
```

```
    "file": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "DATE",
            "LONG"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
    "link": {
     "type": "object",
     "properties": {
```

```
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "STRING_LIST",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
    "attachment": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
```

```
        {
         "type": "object",
         "properties": {
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "STRING_LIST",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    },
    "comment": {
     "type": "object",
     "properties": {
      "fieldMappings": {
       "type": "array",
       "items": [
        {
         "type": "object",
         "properties": {
```

```
          "indexFieldName": {
           "type": "string"
          },
          "indexFieldType": {
           "type": "string",
           "enum": [
            "STRING",
            "STRING_LIST",
            "DATE"
           ]
          },
          "dataSourceFieldName": {
           "type": "string"
          },
          "dateFieldFormat": {
           "type": "string",
           "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
          }
         },
         "required": [
          "indexFieldName",
          "indexFieldType",
          "dataSourceFieldName"
         ]
        }
       ]
      }
     },
     "required": [
      "fieldMappings"
     ]
    }
   }
  },
  "additionalProperties": {
   "type": "object",
   "properties": {
    "eventTitleFilterRegEx": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "pageTitleFilterRegEx": {
```

```
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "linkTitleFilterRegEx": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "inclusionFilePath": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "exclusionFilePath": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "inclusionFileTypePatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "exclusionFileTypePatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "inclusionFileNamePatterns": {
      "type": "array",
      "items": {
       "type": "string"
      }
     },
     "exclusionFileNamePatterns": {
      "type": "array",
      "items": {
```

```
      "type": "string"
     }
    },
    "inclusionOneNoteSectionNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionOneNoteSectionNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "inclusionOneNotePageNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "exclusionOneNotePageNamePatterns": {
     "type": "array",
     "items": {
      "type": "string"
     }
    },
    "crawlFiles": {
     "type": "boolean"
    },
    "crawlPages": {
     "type": "boolean"
    },
    "crawlEvents": {
     "type": "boolean"
    },
    "crawlComments": {
     "type": "boolean"
    },
    "crawlLinks": {
     "type": "boolean"
    },
    "crawlAttachments": {
     "type": "boolean"
```

```
    },
    "crawlListData": {
     "type": "boolean"
    },
    "crawlAcl": {
     "type": "boolean"
    },
    "fieldForUserId": {
     "type": "string"
    },
    "aclConfiguration": {
     "type": "string",
     "enum": [
      "ACLWithLDAPEmailFmt",
      "ACLWithManualEmailFmt",
      "ACLWithUsernameFmt"
     ]
    },
    "emailDomain": {
     "type": "string"
    },
    "isCrawlLocalGroupMapping": {
     "type": "boolean"
    },
    "isCrawlAdGroupMapping": {
     "type": "boolean"
    },
    "proxyHost": {
     "type": "string"
    },
    "proxyPort": {
     "type": "string"
    }
   },
   "required": [
   ]
  },
  "type": {
   "type": "string",
   "pattern": "SHAREPOINTV2"
  },
  "enableIdentityCrawler": {
   "type": "boolean"
  },
```

```
  "syncMode": {
   "type": "string",
   "enum": [
    "FULL_CRAWL",
    "FORCED_FULL_CRAWL",
    "CHANGE_LOG"
   ]
  },
  "secretArn": {
   "type": "string",
   "minLength": 20,
   "maxLength": 2048
  }
 },
 "version": {
  "type": "string",
  "anyOf": [
   {
    "pattern": "1.0.0"
   }
  ]
 },
 "required": [
  "connectionConfiguration",
  "repositoryConfigurations",
  "enableIdentityCrawler",
  "syncMode",
  "additionalProperties",
  "secretArn",
  "type"
 ]
}
```

Esquema de plantilla de Microsoft SQL Server

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `sqlserver`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Microsoft SQL Server.

En la siguiente tabla se describen los parámetros del esquema JSON de Microsoft SQL Server.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • DBType: el tipo de base de datos Java que utiliza, ya sea,`mysql`, `db2` o. `postgresql` `oracle` `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.

Configuración	Descripción
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.

Configuración	Descripción
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Esquema JSON de Microsoft SQL Server

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Esquema de plantilla de Microsoft Teams

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Debe proporcionar el ID de inquilino como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Además, especifique el tipo de origen de datos como `MSTEAMS`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Microsoft Teams.

En la siguiente tabla se describen los parámetros del esquema JSON de Microsoft Teams.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
tenantId	El ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos

Configuración	Descripción
	específicos de contenido y asignaciones de campos.
• chatMessage • chatAttachment • channelPost • channelWiki • channelAttachment • meetingChat • meetingFile • meetingNote • calendarMeeting	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de Microsoft Teams para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
paymentModel	Especifica el tipo de modelo de pago que se debe utilizar con el origen de datos de Microsoft Teams. Los modelos de pago del modelo A están restringidos a los modelos de licencia y pago que requieren el cumplimiento de las normas de seguridad. Los modelos de pago del modelo B son adecuados para los modelos de licencia y pago que no requieren el cumplimiento de las normas de seguridad.

Configuración	Descripción
• inclusionTeamNameFiltrar • inclusionChannelNameFiltro • inclusionFileNamePatrones • inclusionFileTypePatrones • inclusionUserEmailFiltro • inclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns	Una lista de patrones de expresión regular para incluir determinado contenido en su origen de datos de Microsoft Teams. El contenido que coincide con los patrones se incluye en el índice. El contenido que no coincide con los patrones se excluye del índice. Si el contenido coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.
• exclusionTeamNameFiltro • exclusionChannelNameFiltro • exclusionFileNamePatrones • exclusionFileTypePatrones • exclusionUserEmailFiltro • exclusionOneNoteSectionNamePatterns • exclusionOneNotePageNamePatterns	Una lista de patrones de expresión regular para excluir determinado contenido en su origen de datos de Microsoft Teams. El contenido que coincide con los patrones se excluye del índice. El contenido que no coincide con los patrones se incluye en el índice. Si el contenido coincide con un patrón de inclusión y de exclusión , el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.
• isCrawlChatMensaje • isCrawlChatAdjunto • isCrawlChannelPublicar • isCrawlChannelAdjunto • isCrawlChannelWiki • isCrawlCalendarReunión • isCrawlMeetingCharla • isCrawlMeetingArchivo • isCrawlMeetingNota	`true`para rastrear este tipo de contenido en la fuente de datos de Microsoft Teams.
startCalendarDate¿Hora	Puede configurar una fecha y hora de inicio específicas para el contenido de su calendario.

Configuración	Descripción
endCalendarDateHora	Puede configurar una fecha y hora de finalización específicas para el contenido del calendario.
type	El tipo del origen de datos. Especifica `MSTEAMS` como el tipo de origen de datos.
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la información principal o de identidad de los usuarios y grupos con acceso a determinados documentos. Si el rastreador de identidad está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
secretArn	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectars e a Microsoft Teams. Esto incluye su ID de cliente y su secreto de cliente, que se genera al crear una aplicación OAuth en el portal de Azure.

Configuración	Descripción
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Microsoft Teams

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "tenantId": {
              "type": "string",
              "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]
{12}$",
              "minLength": 36,
              "maxLength": 36
            }
          },
          "required": [
            "tenantId"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "chatMessage": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
```

```
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "chatAttachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
```

```
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "channelPost": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
```

```
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "STRING_LIST",
                    "DATE"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "channelWiki": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
```

```
                    "enum": [
                      "STRING",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "channelAttachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
```

```
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "meetingChat": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "STRING_LIST",
                    "DATE"
                  ]
                },
```

```
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "meetingFile": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
```

```
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "meetingNote": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "DATE"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
```

```
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "calendarMeeting": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
```

```
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
   "additionalProperties": {
    "type": "object",
    "properties": {
      "paymentModel": {
        "type": "string",
        "enum": [
          "A",
          "B",
          "Evaluation Mode"
        ]
      },
      "inclusionTeamNameFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionTeamNameFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionChannelNameFilter": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionChannelNameFilter": {
        "type": "array",
        "items": {
```

```
            "type": "string"
          }
        },
        "inclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionUserEmailFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionOneNoteSectionNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionOneNoteSectionNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
```

```
        },
        "inclusionOneNotePageNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionOneNotePageNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "isCrawlChatMessage": {
          "type": "boolean"
        },
        "isCrawlChatAttachment": {
          "type": "boolean"
        },
        "isCrawlChannelPost": {
          "type": "boolean"
        },
        "isCrawlChannelAttachment": {
          "type": "boolean"
        },
        "isCrawlChannelWiki": {
          "type": "boolean"
        },
        "isCrawlCalendarMeeting": {
          "type": "boolean"
        },
        "isCrawlMeetingChat": {
          "type": "boolean"
        },
        "isCrawlMeetingFile": {
          "type": "boolean"
        },
        "isCrawlMeetingNote": {
          "type": "boolean"
        },
        "startCalendarDateTime": {
          "anyOf": [
            {
              "type": "string",
```

```
            "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
          },
          {
            "type": "string",
            "pattern": ""
          }
        ]
      },
      "endCalendarDateTime": {
        "anyOf": [
          {
            "type": "string",
            "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
          },
          {
            "type": "string",
            "pattern": ""
          }
        ]
      }
    },
    "required": []
  },
  "type": {
    "type": "string",
    "pattern": "MSTEAMS"
  },
  "enableIdentityCrawler": {
    "type": "boolean"
  },
  "syncMode": {
    "type": "string",
    "enum": [
      "FORCED_FULL_CRAWL",
      "FULL_CRAWL",
      "CHANGE_LOG"
    ]
  },
  "secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
  }
},
```

```
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Esquema de plantilla de Microsoft Yammer

Incluye un JSON que contiene el esquema de la fuente de datos como parte del objeto. TemplateConfiguration Especifique el tipo de origen de datos como `YAMMER`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifíquelo `TEMPLATE` como Tipo cuando llame CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores.

En la siguiente tabla se describen los parámetros del esquema JSON de Microsoft Yammer.

Configuración	Descripción
connectionConfiguration	Información de configuración del origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos. Este origen de datos no especifica un punto de conexión en `repositoryEndpointMetadata` . Por el contrario, la información de conexión se incluye en un AWS Secrets Manager secreto que usted proporciona. `secretArn`

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
• community • usuario • message • attachment	Una lista de objetos que asignan atributos o nombres de campo de Microsoft Yammer a los nombres de campo del índice de Amazon Kendra. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos
inclusionPatterns	Una lista de patrones de expresión regular para incluir determinados archivos en su origen de datos de Microsoft Yammer. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coinciden con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
exclusionPatterns	Una lista de patrones de expresión regular para excluir determinados archivos en su origen de datos de Microsoft Yammer. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.

Configuración	Descripción
sinceDate	Puede optar por configurar un parámetro `sinceDate` para que el conector de Microsoft Yammer rastree el contenido en función de una `sinceDate` específica.
communityNameFilter	Puede elegir indexar contenido específico de la comunidad.
• isCrawlMessage • isCrawlAttachment • isCrawlPrivateMensaje	`true`para rastrear mensajes, archivos adjuntos de mensajes y mensajes privados.
type	Especifica `YAMMER` como el tipo de origen de datos.
secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectars e a Microsoft Yammer. Esto incluye su nombre de usuario y contraseña de Microsoft Yammer, su ID de cliente y su secreto de cliente, que se genera al crear una aplicación OAuth en el portal de Azure.
useChangeLog	`true`para usar el registro de cambios de Microsoft Yammer para determinar qué documentos del índice deben actualizarse.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

Esquema JSON de Microsoft Yammer

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
          }
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "community": {
          "type": "object",
          "properties": {
            "fieldMappings": {
```

```
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "user": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
```

```
            "anyOf": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "message": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
```

```
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "attachment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
                "type": "object",
                "properties": {
```

```
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": [
                        "STRING",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "inclusionPatterns": {
        "type": "array"
      },
      "exclusionPatterns": {
        "type": "array"
      },
      "sinceDate": {
```

```
          "type": "string",
          "pattern": "^(19|2[0-9])[0-9]{2}-(0[1-9]|1[012])-(0[1-9]|[12][0-9]|
3[01])T(0[0-9]|1[0-9]|2[0-3]):([0-5][0-9]):([0-5][0-9])((\\+|-)(0[0-9]|1[0-9]|2[0-3]):
([0-5][0-9]))?$"
        },
        "communityNameFilter": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "isCrawlMessage": {
          "type": "boolean"
        },
        "isCrawlAttachment": {
          "type": "boolean"
        },
        "isCrawlPrivateMessage": {
          "type": "boolean"
        }
      },
      "required": [
        "sinceDate"
      ]
    },
    "type": {
      "type": "string",
      "pattern": "YAMMER"
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    },
    "useChangeLog": {
      "type": "string",
      "enum": [
        "true",
        "false"
      ]
    },
    "syncMode": {
      "type": "string",
      "enum": [
```

```
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "version": {
      "type": "string",
      "anyOf": [
        {
          "pattern": "1.0.0"
        }
      ]
    }
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "type",
    "secretArn",
    "syncMode"
  ]
}
```

Esquema de plantilla de MySQL

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `mysql`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de MySQL.

En la siguiente tabla se describen los parámetros del esquema JSON de MySQL.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea `mysql`, `db2`, `postgresql` `oracle`, o. `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.

Configuración	Descripción
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas

Configuración	Descripción
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Esquema JSON de MySQL

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Esquema de plantilla de Oracle Database

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `oracle`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Oracle Database.

En la siguiente tabla se describen los parámetros del esquema JSON de Oracle Database.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea `mysql`,`db2`, `postgresql` `oracle`, o. `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

Configuración	Descripción
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.

Configuración	Descripción
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Esquema JSON de Oracle Database

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Esquema de plantilla de PostgreSQL

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Especifique el tipo de origen de datos como `JDBC`, el tipo de base de datos como `postgresql`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de PostgreSQL.

En la siguiente tabla se describen los parámetros del esquema JSON de PostgreSQL.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	Información de configuración necesaria para conectarse a su origen de datos. • dbType: el tipo de base de datos Java que utiliza, ya sea,, `mysql` `db2`, `postgresql` o. `oracle` `sqlserver` • dbHost: el nombre del host de la base de datos. • dbPort: el puerto de la base de datos. • dbInstance: la instancia de base de datos.

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos. Especifique el tipo de origen de datos y el ARN secreto.
revisión	Una lista de objetos que asignan los atributos o los nombres de campo del contenido de la base de datos para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos. Se utiliza para incluir o excluir contenido específico en el origen de datos de la base de datos.
primaryKey	Proporciona la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
titleColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
bodyColumn	Proporciona el nombre de la columna del título del documento en la tabla de la base de datos.
sqlQuery	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

Configuración	Descripción
timestampColumn	Introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de la marca de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
timestampFormat	Introduce el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido .
timezone	Introduce el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
changeDetectingColumns	Introduce los nombres de las columnas que Amazon Kendra se utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas
allowedUsersColumns	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
allowedGroupsColumn	Introduce el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
sourceURIColumn	Introduce el nombre de la columna que contiene las URL de origen que se van a indexar.

Configuración	Descripción
isSslEnabled	Introduce instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
type	El tipo del origen de datos. Especifica `JDBC` como el tipo de origen de datos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción
secretArn	El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene el nombre de usuario y la contraseña necesarios para conectarse a su base de datos. El secreto debe contener una estructura JSON con las siguientes claves: `{ "user name": "database user name", "password": " password" }`
versión	La versión de la plantilla que se admite actualmente.

Esquema JSON de PostgreSQL

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "dbType": {
              "type": "string",
              "enum": [
                "mysql",
                "db2",
                "postgresql",
                "oracle",
                "sqlserver"
              ]
            },
            "dbHost": {
```

```
            "type": "string"
          },
          "dbPort": {
            "type": "string"
          },
          "dbInstance": {
            "type": "string"
          }
        },
        "required": [
          "dbType",
          "dbHost",
          "dbPort",
          "dbInstance"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "document": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string"
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  }
                },
                "required": [
```

```
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "primaryKey": {
        "type": "string"
      },
      "titleColumn": {
        "type": "string"
      },
      "bodyColumn": {
        "type": "string"
      },
      "sqlQuery": {
        "type": "string",
        "not": {
          "pattern": ";+"
        }
      },
      "timestampColumn": {
        "type": "string"
      },
      "timestampFormat": {
        "type": "string"
      },
      "timezone": {
        "type": "string"
      },
      "changeDetectingColumns": {
```

```
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "allowedUsersColumn": {
          "type": "string"
        },
        "allowedGroupsColumn": {
          "type": "string"
        },
        "sourceURIColumn": {
          "type": "string"
        },
        "isSslEnabled": {
          "type": "boolean"
        }
      },
      "required": ["primaryKey", "titleColumn", "bodyColumn", "sqlQuery"]
    },
    "type" : {
      "type" : "string",
      "pattern": "JDBC"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
```

```
  "required": [
      "connectionConfiguration",
      "repositoryConfigurations",
      "syncMode",
      "additionalProperties",
      "secretArn",
      "type"
  ]
}
```

Esquema de plantilla de Salesforce

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Debe proporcionar la URL del host de Salesforce como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Además, especifique el tipo de origen de datos como `SALESFORCEV2`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Salesforce.

En la siguiente tabla se describen los parámetros del esquema JSON de Salesforce.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
hostUrl	La URL de la instancia de Salesforce que se va a indexar.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.

Configuración	Descripción
• cuenta • contact • campaña • caso • product • lead • contrato • partner • profile • idea • pricebook • tarea • solución • attachment • usuario • revisión • knowledgeArticles • grupo • opportunity • chatter • customEntity	Una lista de objetos que mapean los atributos o los nombres de campo de sus entidades de Salesforce para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos.

Configuración	Descripción

secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a su Salesforce. El secreto debe contener una estructura JSON con las siguientes claves:
	
	{
	"authenticationUrl": " *OAUTH endpoint that Amazon Kendra connects to get an OAUTH token*",
	"consumerKey": " *Application public key generated when you created your Salesforce application* ",
	"consumerSecret": " *Application private key generated when you created your Salesforce application* ",
	"password": " *Password associated with the user logging in to the Salesforce instance* ",
	"securityToken": " *Token associated with the user account logging in to the Salesforce instance* ",
	"username": " *User name of the user logging in to the Salesforce instance*"
	}

additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos

Configuración	Descripción
• accountFilter • contactFilter • caseFilter • campaignFilter • contractFilter • groupFilter • leadFilter • productFilter • opportunityFilter • partnerFilter • pricebookFilter • ideaFilter • profileFilter • taskFilter • solutionFilter • userFilter • chatterFilter • documentFilter • knowledgeArticleFilter • customEntities	Una colección de cadenas que especifica qué entidades filtrar.

Configuración	Descripción
inclusionPatterns • inclusionDocumentFileTypePatterns • inclusionDocumentFileNamePatterns • inclusionAccountFileTypePatterns • inclusionCampaignFileTypePatterns • inclusionDocumentFileNamePatterns • inclusionCampaignFileNamePatterns • inclusionCaseFileTypePatterns • inclusionCaseFileNamePatterns • inclusionContactFileTypePatterns • inclusionContractFileNamePatterns • inclusionLeadFileTypePatterns • inclusionLeadFileNamePatterns • inclusionOpportunityFileTypePatterns • inclusionOpportunityFileNamePatterns • inclusionSolutionFileTypePatterns • inclusionSolutionFileNamePatterns • inclusionTaskFileTypePatterns • inclusionTaskFileNamePatterns • inclusionGroupFileTypePatterns • inclusionGroupFileNamePatterns • inclusionChatterFileTypePatterns • inclusionChatterFileNamePatterns • inclusionCustomEntityFileTypePatterns • inclusionCustomEntityFileNamePatterns	Una lista de patrones de expresión regular para incluir determinados archivos en su origen de datos de Salesforce. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.

Configuración	Descripción
exclusionPatterns • exclusionDocumentFileTypePatterns • exclusionDocumentFileNamePatterns • exclusionAccountFileTypePatterns • exclusionCampaignFileTypePatterns • exclusionCampaignFileNamePatterns • exclusionCaseFileTypePatterns • exclusionCaseFileNamePatterns • exclusionContactFileTypePatterns • exclusionContractFileNamePatterns • exclusionLeadFileTypePatterns • exclusionLeadFileNamePatterns • exclusionOpportunityFileTypePatterns • exclusionOpportunityFileNamePatterns • exclusionSolutionFileTypePatterns • exclusionSolutionFileNamePatterns • exclusionTaskFileTypePatterns • exclusionTaskFileNamePatterns • exclusionGroupFileTypePatterns • exclusionGroupFileNamePatterns • exclusionChatterFileTypePatterns • exclusionChatterFileNamePatterns • exclusionCustomEntityFileTypePatterns • exclusionCustomEntityFileNamePatterns	Una lista de patrones de expresión regular para excluir determinados archivos en su origen de datos de Salesforce. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.

Configuración	Descripción
• isCrawlAccount • isCrawlContact • isCrawlCase • isCrawlCampaign • isCrawlProduct • isCrawlLead • isCrawlContract • isCrawlPartner • isCrawlProfile • isCrawlIdea • isCrawlPricebook • isCrawlDocument • crawlSharedDocument • isCrawlGroup • isCrawlOpportunity • isCrawlChatter • isCrawlUser • isCrawlSolution • isCrawlTask • isCrawlAccountAdjuntos • isCrawlContactAdjuntos • isCrawlCaseAdjuntos • isCrawlCampaignAdjuntos • isCrawlLeadAdjuntos • isCrawlContractAdjuntos • isCrawlGroupAdjuntos • isCrawlOpportunityAdjuntos • isCrawlChatterAdjuntos • isCrawlSolutionAdjuntos	`true`para rastrear estos tipos de archivos en su cuenta de Salesforce.

Configuración	Descripción

• isCrawlTaskAdjuntos
• isCrawlCustomEntityAttachments
• isCrawlKnowledgeArtículos
  • isCrawlDraft
  • isCrawlPublish
  • isCrawlArchived	

type	El tipo del origen de datos. Especifica `SALESFORCEV2` como el tipo de origen de datos.
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.

Configuración	Descripción
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Salesforce

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties":
```

```
{
  "connectionConfiguration": {
    "type": "object",
    "properties":
    {
      "repositoryEndpointMetadata":
      {
        "type": "object",
        "properties":
        {
          "hostUrl":
          {
            "type": "string",
            "pattern": "https:.*"
          }
        },
        "required":
        [
          "hostUrl"
        ]
      }
    },
    "required":
    [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties":
    {
      "account":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
```

```
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "contact":
      {
        "type": "object",
```

```
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
```

```
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "campaign":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
                {
                  "type": "string",
                  "enum":
                  [
                    "STRING",
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
```

```
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "case":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
```

```
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "product":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
```

```
                {
                  "type": "string",
                  "enum":
                  [
                    "STRING",
                    "STRING_LIST",
                    "DATE"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "lead":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
```

```
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE",
                        "LONG"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          },
          "required":
          [
            "fieldMappings"
          ]
        },
```

```
        "contract":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
```

```
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "partner":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
                {
                  "type": "string",
                  "enum":
                  [
                    "STRING",
                    "STRING_LIST",
                    "DATE"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
```

```
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "profile":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
```

```
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "idea":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
```

```
                },
                "indexFieldType":
                {
                  "type": "string",
                  "enum":
                  [
                    "STRING",
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "pricebook":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
```

```
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
```

```
          ]
        },
        "task":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
```

```
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "solution":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
                {
                  "type": "string",
                  "enum":
                  [
                    "STRING",
                    "STRING_LIST",
                    "DATE"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
```

```
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "attachment":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
                "indexFieldName":
                {
                  "type": "string"
                },
                "indexFieldType":
                {
                  "type": "string",
                  "enum":
                  [
                    "STRING",
```

```
                    "STRING_LIST",
                    "DATE",
                    "LONG"
                  ]
                },
                "dataSourceFieldName":
                {
                  "type": "string"
                },
                "dateFieldFormat":
                {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required":
              [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required":
      [
        "fieldMappings"
      ]
    },
    "user":
    {
      "type": "object",
      "properties":
      {
        "fieldMappings":
        {
          "type": "array",
          "items":
          [
            {
              "type": "object",
              "properties":
              {
```

```
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "document":
      {
        "type": "object",
        "properties":
        {
```

```
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE",
                        "LONG"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          },
```

```
          "required":
          [
            "fieldMappings"
          ]
        },
        "knowledgeArticles":
        {
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
```

```
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "group":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName":
                  {
```

```
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "opportunity":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
```

```
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "chatter":
      {
        "type": "object",
        "properties":
        {
          "fieldMappings":
          {
            "type": "array",
            "items":
            [
              {
```

```
                "type": "object",
                "properties":
                {
                  "indexFieldName":
                  {
                    "type": "string"
                  },
                  "indexFieldType":
                  {
                    "type": "string",
                    "enum":
                    [
                      "STRING",
                      "STRING_LIST",
                      "DATE"
                    ]
                  },
                  "dataSourceFieldName":
                  {
                    "type": "string"
                  },
                  "dateFieldFormat":
                  {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required":
                [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required":
        [
          "fieldMappings"
        ]
      },
      "customEntity":
      {
```

```
          "type": "object",
          "properties":
          {
            "fieldMappings":
            {
              "type": "array",
              "items":
              [
                {
                  "type": "object",
                  "properties":
                  {
                    "indexFieldName":
                    {
                      "type": "string"
                    },
                    "indexFieldType":
                    {
                      "type": "string",
                      "enum":
                      [
                        "STRING",
                        "STRING_LIST",
                        "DATE"
                      ]
                    },
                    "dataSourceFieldName":
                    {
                      "type": "string"
                    },
                    "dateFieldFormat":
                    {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required":
                  [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
```

```
            }
          },
          "required":
          [
            "fieldMappings"
          ]
        }
      }
    },
    "additionalProperties": {
      "type": "object",
      "properties":
      {
        "accountFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "contactFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "caseFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "campaignFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "contractFilter":{
          "type": "array",
          "items":
```

```
          {
            "type": "string"
          }
        },
        "groupFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "leadFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "productFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "opportunityFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "partnerFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "pricebookFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
```

```
        }
      },
      "ideaFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "profileFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "taskFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "solutionFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "userFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "chatterFilter":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
```

```
        "documentFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "knowledgeArticleFilter":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "customEntities":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "isCrawlAccount": {
          "type": "boolean"
        },
        "isCrawlContact": {
          "type": "boolean"
        },
        "isCrawlCase": {
          "type": "boolean"
        },
        "isCrawlCampaign": {
          "type": "boolean"
        },
        "isCrawlProduct": {
          "type": "boolean"
        },
        "isCrawlLead": {
          "type": "boolean"
        },
        "isCrawlContract": {
          "type": "boolean"
        },
        "isCrawlPartner": {
          "type": "boolean"
```

```
        },
        "isCrawlProfile": {
          "type": "boolean"
        },
        "isCrawlIdea": {
          "type": "boolean"
        },
        "isCrawlPricebook": {
          "type": "boolean"
        },
        "isCrawlDocument": {
          "type": "boolean"
        },
        "crawlSharedDocument": {
          "type": "boolean"
        },
        "isCrawlGroup": {
          "type": "boolean"
        },
        "isCrawlOpportunity": {
          "type": "boolean"
        },
        "isCrawlChatter": {
          "type": "boolean"
        },
        "isCrawlUser": {
          "type": "boolean"
        },
        "isCrawlSolution":{
          "type": "boolean"
        },
        "isCrawlTask":{
          "type": "boolean"
        },

        "isCrawlAccountAttachments": {
          "type": "boolean"
        },
        "isCrawlContactAttachments": {
          "type": "boolean"
        },
        "isCrawlCaseAttachments": {
          "type": "boolean"
        },
```

```
        "isCrawlCampaignAttachments": {
          "type": "boolean"
        },
        "isCrawlLeadAttachments": {
          "type": "boolean"
        },
        "isCrawlContractAttachments": {
          "type": "boolean"
        },
        "isCrawlGroupAttachments": {
          "type": "boolean"
        },
        "isCrawlOpportunityAttachments": {
          "type": "boolean"
        },
        "isCrawlChatterAttachments": {
          "type": "boolean"
        },
        "isCrawlSolutionAttachments":{
          "type": "boolean"
        },
        "isCrawlTaskAttachments":{
          "type": "boolean"
        },
        "isCrawlCustomEntityAttachments":{
          "type": "boolean"
        },
        "isCrawlKnowledgeArticles": {
          "type": "object",
          "properties":
          {
            "isCrawlDraft": {
              "type": "boolean"
            },
            "isCrawlPublish": {
              "type": "boolean"
            },
            "isCrawlArchived": {
              "type": "boolean"
            }
          }
        },
        "inclusionDocumentFileTypePatterns":{
          "type": "array",
```

```
          "items":
          {
            "type": "string"
          }
        },
        "exclusionDocumentFileTypePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionDocumentFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionDocumentFileNamePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionAccountFileTypePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionAccountFileTypePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionAccountFileNamePatterns":{
          "type": "array",
          "items":
          {
```

```
            "type": "string"
          }
        },
        "exclusionAccountFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionCampaignFileTypePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionCampaignFileTypePatterns": {
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionCampaignFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionCampaignFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionCaseFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
```

```
        },
        "exclusionCaseFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionCaseFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionCaseFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionContactFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionContactFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionContactFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionContactFileNamePatterns":{
```

```
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionContractFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionContractFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionContractFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionContractFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionLeadFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionLeadFileTypePatterns":{
          "type": "array",
          "items":
```

```
          {
            "type": "string"
          }
        },
        "inclusionLeadFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionLeadFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionOpportunityFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionOpportunityFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionOpportunityFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionOpportunityFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
```

```
        }
      },
      "inclusionSolutionFileTypePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "exclusionSolutionFileTypePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "inclusionSolutionFileNamePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "exclusionSolutionFileNamePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "inclusionTaskFileTypePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
      "exclusionTaskFileTypePatterns":{
        "type": "array",
        "items":
        {
          "type": "string"
        }
      },
```

```
        "inclusionTaskFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionTaskFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionGroupFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionGroupFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionGroupFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionGroupFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionChatterFileTypePatterns":{
          "type": "array",
```

```
          "items":
          {
            "type": "string"
          }
        },
        "exclusionChatterFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionChatterFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionChatterFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionCustomEntityFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "exclusionCustomEntityFileTypePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        },
        "inclusionCustomEntityFileNamePatterns":{
          "type": "array",
          "items":
          {
```

```
            "type": "string"
          }
        },
        "exclusionCustomEntityFileNamePatterns":{
          "type": "array",
          "items":
          {
            "type": "string"
          }
        }
      },
      "required":
      []
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "type": {
      "type": "string",
      "pattern": "SALESFORCEV2"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FULL_CRAWL",
        "FORCED_FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
```

```
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

ServiceNow esquema de plantilla

Incluye un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Debe proporcionar la URL del ServiceNow host, el tipo de autenticación y la versión de la instancia como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Además, especifique el tipo de origen de datos como `SERVICENOWV2`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, debe `TEMPLATE` especificarlo `Type` cuando llame CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte ServiceNow Esquema JSON.

En la siguiente tabla se describen los parámetros del esquema ServiceNow JSON.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
hostUrl	La URL del ServiceNow host. Por ejemplo, *your-domain.service-now.com* .
authType	El tipo de autenticación que utiliza, ya sea `basicAuth` o `OAuth2`.
servicenowInstanceVersion	La ServiceNow versión que utilizas. Puede elegir entre`Tokyo`, `SandiegoRome`, y`Others`.

Configuración	Descripción
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
• knowledgeArticle • attachment • serviceCatalog • incident	Una lista de objetos que mapean los atributos o nombres de campo de sus artículos de ServiceNow conocimiento, archivos adjuntos, catálogo de servicios e incidentes para Amazon Kendra indexar los nombres de los campos. Para obtener más información, consulte Asignación de campos de origen de datos. Los nombres de los campos de la fuente de ServiceNow datos deben estar en sus metadatos ServiceNow personalizados.
additional properties	Opciones de configuración adicionales para el contenido del origen de datos.
maxFileSizeInMegaBytes	Especifique el límite de tamaño del archivo en MB que Amazon Kendra rastreará. Amazon Kendra rastreará solo los archivos dentro del límite de tamaño que usted defina. El tamaño predeterminado del archivo es de 50 MB. El tamaño máximo del archivo debe ser superior a 0 MB e inferior o igual a 50 MB.

Configuración	Descripción
• knowledgeArticleFilter • incidentQueryFilter • serviceCatalogQueryFiltro • knowledgeArticleTitleRegExp • serviceCatalogTitleRegExp • incidentTitleRegExp • inclusionFileTypePatrones • exclusionFileTypePatrones • inclusionFileNamePatrones • exclusionFileNamePatrones • incidentStateType	Una lista de patrones de expresiones regulares para incluir o excluir determinados archivos de la fuente ServiceNow de datos. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
• isCrawlKnowledgeArtículo • isCrawlKnowledgeArticleAttachment • includePublicArticlesSolo • isCrawlServiceCatálogo • isCrawlServiceCatalogAttachment • isCrawlActiveServiceCatalog • isCrawlInactiveServiceCatalog • isCrawlIncident • isCrawlIncidentAdjunto • isCrawlActiveIncidente • isCrawlInactiveIncidente • Aplicar una CL ForKnowledgeArticle • Aplicar una ACL ForServiceCatalog • Aplicar una ACL ForIncident	`true`para rastrear artículos de ServiceNow conocimie nto, catálogos de servicios, incidentes y archivos adjuntos.
type	El tipo del origen de datos. Especifica `SERVICENOWV2` como el tipo de origen de datos.

Configuración	Descripción
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la información principal o de identidad de los usuarios y grupos con acceso a determina dos documentos. Si el rastreador de identidad está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappin gAPI para cargar la información de acceso de usuarios y grupos.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronic e con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

Configuración	Descripción

secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a su. ServiceNow El secreto debe contener una estructura JSON con las siguientes claves:
	
	`{ "username": " user name", "password": " password" }`
	
	Si utiliza la autenticación OAuth2, su secreto debe contener una estructura JSON con las siguientes claves:
	
	`{ "username": " user name", "password": " password", "clientId": " client id", "clientSecret": " client secret" }`

versión	La versión de la plantilla que se admite actualmente.

ServiceNow Esquema JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "hostUrl": {
              "type": "string",
```

```
              "pattern": "^(?!(^(https?|ftp|file):\/\/))[a-z0-9-]+(.service-
now.com|.servicenowservices.com)$",
              "minLength": 1,
              "maxLength": 2048
            },
            "authType": {
              "type": "string",
              "enum": [
                "basicAuth",
                "OAuth2"
              ]
            },
            "servicenowInstanceVersion": {
              "type": "string",
              "enum": [
                "Tokyo",
                "Sandiego",
                "Rome",
                "Others"
                ]
            }
          },
          "required": [
            "hostUrl",
            "authType",
            "servicenowInstanceVersion"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "knowledgeArticle": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
```

```
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "STRING_LIST"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "attachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
```

```
                },
                "indexFieldType": {
                  "type": "string",
                  "enum": [
                    "STRING",
                    "LONG",
                    "DATE",
                    "STRING_LIST"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "serviceCatalog": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": [
            {
              "type": "object",
              "properties": {
                "indexFieldName": {
                  "type": "string"
                },
                "indexFieldType": {
```

```
                    "type": "string",
                    "enum": [
                      "STRING",
                      "DATE",
                      "STRING_LIST"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "incident": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
```

```
                    "DATE",
                    "STRING_LIST"
                  ]
                },
                "dataSourceFieldName": {
                  "type": "string"
                },
                "dateFieldFormat": {
                  "type": "string",
                  "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                }
              },
              "required": [
                "indexFieldName",
                "indexFieldType",
                "dataSourceFieldName"
              ]
            }
          ]
        }
      },
      "required": [
        "fieldMappings"
      ]
    }
  }
},
"additionalProperties": {
  "type": "object",
  "properties": {
    "maxFileSizeInMegaBytes": {
      "type": "string"
    },
    "isCrawlKnowledgeArticle": {
      "type": "boolean"
    },
    "isCrawlKnowledgeArticleAttachment": {
      "type": "boolean"
    },
    "includePublicArticlesOnly": {
      "type": "boolean"
    },
    "knowledgeArticleFilter": {
      "type": "string"
```

```
        },
        "incidentQueryFilter": {
          "type": "string"
        },
        "serviceCatalogQueryFilter": {
          "type": "string"
        },
        "isCrawlServiceCatalog": {
          "type": "boolean"
        },
        "isCrawlServiceCatalogAttachment": {
          "type": "boolean"
        },
        "isCrawlActiveServiceCatalog": {
          "type": "boolean"
        },
        "isCrawlInactiveServiceCatalog": {
          "type": "boolean"
        },
        "isCrawlIncident": {
          "type": "boolean"
        },
        "isCrawlIncidentAttachment": {
          "type": "boolean"
        },
        "isCrawlActiveIncident": {
          "type": "boolean"
        },
        "isCrawlInactiveIncident": {
          "type": "boolean"
        },
        "applyACLForKnowledgeArticle": {
          "type": "boolean"
        },
        "applyACLForServiceCatalog": {
          "type": "boolean"
        },
        "applyACLForIncident": {
          "type": "boolean"
        },
        "incidentStateType": {
          "type": "array",
          "items": {
            "type": "string",
```

```
            "enum": [
              "Open",
              "Open - Unassigned",
              "Resolved",
              "All"
            ]
          }
        },
        "knowledgeArticleTitleRegExp": {
          "type": "string"
        },
        "serviceCatalogTitleRegExp": {
          "type": "string"
        },
        "incidentTitleRegExp": {
          "type": "string"
        },
        "inclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileTypePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "inclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        },
        "exclusionFileNamePatterns": {
          "type": "array",
          "items": {
            "type": "string"
          }
        }
      },
      "required": []
    },
```

```
    "type": {
      "type": "string",
      "pattern": "SERVICENOWV2"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "syncMode": {
      "type": "string",
      "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL"
      ]
    },
    "secretArn": {
      "type": "string",
      "minLength": 20,
      "maxLength": 2048
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
  ]
}
```

Esquema de plantillas de Slack

Incluyes un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Debe proporcionar la URL del host como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Además, especifique el

tipo de origen de datos como `SLACK`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Slack.

En la siguiente tabla se describen los parámetros del esquema JSON de Slack.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
ID de equipo	El ID del equipo de Slack que copiaste de la URL de tu página principal de Slack.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
Todos	Una lista de objetos que mapean los atributos o los nombres de campo de tu Slack contenido para Amazon Kendra indexar los nombres de los campos.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos.
inclusionPatterns	Una lista de patrones de expresiones regulares para incluir contenido específico desde su origen de datos de Slack. El contenido que coincide con los patrones se incluye en el índice. El contenido que no coincide con los patrones se excluye del índice. Si algún

Configuración	Descripción
	contenido coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.
exclusionPatterns	Una lista de patrones de expresiones regulares para excluir contenido específico en su origen de datos de Slack. El contenido que coincide con los patrones se excluye del índice. El contenido que no coincide con los patrones se incluye en el índice. Si algún contenido coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.
crawlBotMessages	`true`para rastrear los mensajes de los bots.
Excluir archivados	`true`para excluir el rastreo de los mensajes archivados.
Tipo de conversación	El tipo de conversación que desea indexar`PUBLIC_CHANNEL` , ya sea`PRIVATE_CHANNEL` , `GROUP_MESSAGE` y. `DIRECT_MESSAGE`
Filtro de canal	El tipo de canal que quiere indexar si `private_channel` o `public_channel` .
sinceDate	Puede optar por configurar un parámetro `sinceDate` para que el conector de Slack rastree el contenido en una `sinceDate` específica.

Configuración	Descripción
Mira hacia atrás	Puede configurar un `lookBack` parámetro para que el Slack conector rastree el contenido actualizado o eliminado hasta un número específico de horas antes de la última sincroniz ación del conector.
syncMode	Especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Puede elegir entre las siguientes opciones: • `FORCED_FULL_CRAWL` para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice. • `FULL_CRAWL` para indexar únicamente el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronic e con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización. • `CHANGE_LOG` para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
type	El tipo del origen de datos. Especifica `SLACK` como el tipo de origen de datos.

Configuración	Descripción
enableIdentityCrawler	`true`utilizar el rastreador Amazon Kendra de identidad para sincronizar la informaci ón principal o de identidad de los usuarios y grupos con acceso a determinados documento s. Si el rastreador de identidad está desactiva do, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar la información de acceso de usuarios y grupos.
secretArn	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectars e a su. Slack El secreto debe contener una estructura JSON con las siguientes claves: `{ "slackToken": " token" }`
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Slack

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
```

```
          "properties": {
            "teamId": {
              "type": "string"
            }
          },
          "required": ["teamId"]
        }
      }
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "All": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "DATE","LONG"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
```

```
        },
        "required": [
          "fieldMappings"
        ]
      }
    },
    "required": [
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "exclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "crawlBotMessages": {
        "type": "boolean"
      },
      "excludeArchived": {
        "type": "boolean"
      },
      "conversationType": {
        "type": "array",
        "items": {
          "type": "string",
          "enum": [
            "PUBLIC_CHANNEL",
            "PRIVATE_CHANNEL",
            "GROUP_MESSAGE",
            "DIRECT_MESSAGE"
          ]
        }
      },
      "channelFilter": {
          "type": "object",
```

```
            "properties": {
              "private_channel": {
                "type": "array",
                "items": {
                  "type": "string"
                }
              },
              "public_channel": {
                "type": "array",
                "items": {
                  "type": "string"
                }
              }
            }
          },
          "channelIdFilter": {
            "type": "array",
            "items": {
              "type": "string"
            }
          },
          "sinceDate": {
            "anyOf": [
              {
                "type": "string",
                "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
              },
              {
                "type": "string",
                "pattern": ""
              }
            ]
          },
          "lookBack": {
            "type": "string",
            "pattern": "^[0-9]*$"
          }
        },
        "required": [
        ]
      },
      "syncMode": {
        "type": "string",
        "enum": [
```

```
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
      ]
    },
    "type" : {
      "type" : "string",
      "pattern": "SLACK"
    },
    "enableIdentityCrawler": {
      "type": "boolean"
    },
    "secretArn": {
      "type": "string"
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type",
    "enableIdentityCrawler"
  ]
}
```

Esquema de plantilla de Zendesk

Incluyes un JSON que contiene el esquema de la fuente de datos como parte del TemplateConfigurationobjeto. Debe proporcionar la URL del host como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Además, especifique el tipo de origen de datos como ZENDESK, un secreto para sus credenciales de autenticación y otras

configuraciones necesarias. A continuación, `TEMPLATE` se especifica como `Type` cuando se llama CreateDataSource.

Puede usar la plantilla que se proporciona en esta guía para desarrolladores. Consulte Esquema JSON de Zendesk.

En la siguiente tabla se describen los parámetros del esquema JSON de Zendesk.

Configuración	Descripción
connectionConfiguration	Información de configuración para el punto de conexión para el origen de datos.
repositoryEndpointMetadata	La información del punto de conexión para el origen de datos.
hostURL	La URL del host de Zendesk. Por ejemplo, https://yoursubdomain.zendesk.com.
repositoryConfigurations	Información de configuración del contenido del origen de datos. Por ejemplo, configurar tipos específicos de contenido y asignaciones de campos.
• ticket • ticketComment • ticketCommentAttachment • article • articleComment • articleAttachment • communityTopic • communityPostComment	Una lista de objetos que asignan atributos de origen de datos o nombres de campo de tickets de Zendesk a los nombres de campo del índice de Amazon Kendra. Para obtener más información, consulte Asignación de campos de origen de datos.
secretARN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a su cuenta de Zendesk. El secreto debe contener una estructura JSON

Configuración	Descripción
	con las siguientes claves: URL de host, ID de cliente, secreto del cliente, nombre de usuario y contraseña.
additionalProperties	Opciones de configuración adicionales para el contenido del origen de datos
organizationNameFilter	Puede optar por indexar los tickets que existen en una organización específica.
sinceDate	Puede optar por configurar un parámetro `sinceDate` para que el conector de Zendesk rastree el contenido en función de una `sinceDate` específica.
inclusionPatterns	Una lista de patrones de expresión regular para incluir determinados archivos en su origen de datos de Zendesk. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
exclusionPatterns	Una lista de patrones de expresión regular para excluir determinados archivos en su origen de datos de Zendesk. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coincidan con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.

Configuración	Descripción
• isCrawlTicket • isCrawlTicketComentario • isCrawlTicketCommentAttachment • isCrawlArticle • isCrawlArticleComentario • isCrawlArticleAdjunto • isCrawlCommunityTema • isCrawlCommunityPublicar • isCrawlCommunityPostComment	Introduce `true` "» para rastrear estos tipos de contenido.
type	Especifica `ZENDESK` como el tipo de origen de datos.
useChangeLog	Introduzca `true` "» para usar el registro de cambios de Zendesk y determinar qué documentos del índice deben actualizarse. Según el tamaño del registro de cambios, podría ser más rápido escanear los documento s en Zendesk. Si está sincronizando el origen de datos de Zendesk con su índice por primera vez, se escanean todos los documentos.

Esquema JSON de Zendesk

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "hostUrl": {
```

```
            "type": "string",
            "pattern": "https:.*"
          }
        },
        "required": [
          "hostUrl"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "ticket": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
```

```
                    "dataSourceFieldName"

                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "ticketComment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"

                  ]
                }
```

```
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "ticketCommentAttachment": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "dd-MM-yyyy HH:mm:ss"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        }
      },
      "required": [
```

```
            "fieldMappings"
          ]
        },
        "article": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "communityPostComment": {
          "type": "object",
```

```
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "articleComment": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
```

```
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "articleAttachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
```

```
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "communityTopic": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
```

```
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "organizationNameFilter": {
        "type": "array"
      },
      "sinceDate": {
        "type": "string",
        "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2} [0-9]{2}:[0-9]{2}:[0-9]{2}$"
      },
      "inclusionPatterns": {
        "type": "array"
      },
      "exclusionPatterns": {
        "type": "array"
```

```
        },
        "isCrawTicket": {
          "type": "string"
        },
        "isCrawTicketComment": {
          "type": "string"
        },
        "isCrawTicketCommentAttachment": {
          "type": "string"
        },
        "isCrawlArticle": {
          "type": "string"
        },
        "isCrawlArticleAttachment": {
          "type": "string"
        },
        "isCrawlArticleComment": {
          "type": "string"
        },
        "isCrawlCommunityTopic": {
          "type": "string"
        },
        "isCrawlCommunityPost": {
          "type": "string"
        },
        "isCrawlCommunityPostComment": {
          "type": "string"
        }
      }
    },
    "type": {
      "type": "string",
      "pattern": "ZENDESK"
    },
    "useChangeLog": {
      "type": "string",
      "enum": ["true", "false"]
    }
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
```

```
      }
    ]
  },
  "additionalProperties": false,
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "useChangeLog",
    "secretArn",
    "type"
  ]
}
```

Adobe Experience Manager

Adobe Experience Manager es un sistema de gestión de contenido que se utiliza para crear contenido de sitios web o aplicaciones móviles. Puede usarlo Amazon Kendra para conectarse a sus páginas Adobe Experience Manager y activos de contenido e indexarlos.

Amazon Kendra admite Adobe Experience Manager (AEM) como instancia de autor de Cloud Service y como instancia de autor y publicación Adobe Experience Manager local.

Puede conectarse Amazon Kendra a su fuente de Adobe Experience Manager datos mediante la Amazon Kendra consola o la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuentes de datos de Amazon Kendra Adobe Experience Manager, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión

Características admitidas

El conector de origen de datos de Adobe Experience Manager admite las siguientes características:

- Asignaciones de campo
- Control de acceso de usuarios

- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Autenticación OAuth 2.0 y básica
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de Adobe Experience Manager datos, realice estos cambios en sus AWS cuentas Adobe Experience Manager y.

En Adobe Experience Manager, asegúrese de que:

- Tiene acceso a una cuenta con privilegios administrativos o un usuario administrador.
- Ha copiado la URL del host de Adobe Experience Manager.

> **Note**
>
> (local o en el servidor) Amazon Kendra comprueba si la información de punto final incluida AWS Secrets Manager es la misma que la información de punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a evitar el problema del suplente confuso, que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, pero utiliza Amazon Kendra como proxy para acceder al secreto configurado y realizar la acción. Si más adelante cambia la información de punto de conexión, debe crear un nuevo secreto para sincronizar esta información.

- Ha apuntado las credenciales de autenticación básica del nombre de usuario y la contraseña del administrador.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Opcional: se configuraron las credenciales de OAuth 2.0 en Adobe Experience Manager (AEM) como un servicio en la nube o en las instalaciones de AEM. Si utiliza AEM On-Premise,

las credenciales incluyen el ID de cliente, el secreto del cliente y la clave privada. Si utiliza AEM as a Cloud Service, las credenciales incluyen el ID de cliente, el secreto del cliente, la clave privada, el ID de la organización, el ID de la cuenta técnica y el host de Adobe Identity Management System (IMS). Para obtener más información sobre cómo generar estas credenciales para AEM as a Cloud Service, consulte la documentación de Adobe Experience Manager.
En el caso de AEM On-Premise, la implementación del servidor OAuth 2.0 de Adobe Granite (com.adobe.granite.oauth.server) es compatible con las funcionalidades del servidor OAuth 2.0 en AEM.

- Ha comprobado que cada documento es único en Adobe Experience Manager y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En las suyas Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Adobe Experience Manager en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes un IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y Secrets Manager secreto al conectar tu fuente de datos de Adobe Experience Manager a Amazon

Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de Adobe Experience Manager datos, debe proporcionar los detalles necesarios de la fuente de Adobe Experience Manager datos para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configurado Adobe Experience Manager Amazon Kendra, consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Adobe Experience Manager

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, seleccione el conector Adobe Experience Manager y, a continuación, seleccione Añadir conector. Si utiliza la versión 2 (si corresponde), elija el conector Adobe Experience Manager con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. Origen: elija AEM On-Premise o AEM as a Cloud Service.

      Introduzca la URL del host de Adobe Experience Manager. Por ejemplo, si utiliza AEM On-Premise, debe incluir el nombre de host y el puerto: https://hostname:port. O bien, si usa AEM as a Cloud Service, puede usar la URL del autor: https://author-xxxxxx-xxxxxxx.adobeaemcloud.com.

   b. Ubicación del certificado SSL: introduzca la ruta al certificado SSL almacenado en un bucket de Amazon S3 . Se utiliza para conectarse a AEM On-Premise mediante una conexión SSL segura.

   c. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

   d. Autenticación: elija la autenticación básica o la autenticación OAuth 2.0. A continuación, elija un AWS Secrets Manager secreto existente o cree uno nuevo para almacenar sus Adobe Experience Manager credenciales. Si decides crear un secreto nuevo, se abrirá una ventana AWS Secrets Manager secreta.

      Si ha elegido la autenticación básica, introduzca un nombre para el secreto, el nombre de usuario del sitio de Adobe Experience Manager y la contraseña. El usuario debe tener permiso de administrador o ser un usuario administrador.

      Si ha elegido la autenticación OAuth 2.0 y utiliza AEM On-Premise, introduzca un nombre para el secreto, el ID de cliente, el secreto de cliente y la clave privada. Si utiliza AEM as a Cloud Service, introduzca un nombre para el secreto, el ID de cliente, el secreto del cliente, la clave privada, el ID de la organización, el ID de la cuenta técnica y el host de Adobe Identity Management System (IMS).

      Guarda y añade tu secreto.

   e. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

   f. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control

de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

g. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

h. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Ámbito de sincronización: establezca límites para rastrear determinados tipos de contenido, componentes de páginas y rutas raíz y filtre el contenido mediante patrones de expresiones regex.

      i. Tipos de contenido: elija si deseas rastrear solo las páginas o los activos, o ambos.

      ii. (Opcional) Configuración adicional: configure los siguientes ajustes:

      - Componentes de página: los nombres específicos de los componentes de página. El componente de página es un componente de página extensible diseñado para funcionar con el editor de plantillas de Adobe Experience Manager y permite ensamblar los componentes de encabezado/pie de página y estructura con el editor de plantillas.
      - Variaciones de fragmentos de contenido: los nombres específicos de las variaciones de fragmentos de contenido. Los fragmentos de contenido le permiten diseñar, crear, seleccionar y publicar contenido independiente de la página en

Adobe Experience Manager. Le permiten preparar contenido listo para su uso en múltiples ubicaciones o en múltiples canales.

- Rutas raíz: las rutas raíz a contenido específico.
- Patrones regex: patrones de expresiones regulares para incluir o excluir determinadas páginas y recursos.

b. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

c. ID de zona horaria: si utiliza AEM On-Premise y la zona horaria del servidor es diferente a la zona horaria del conector o índice de AEM de Amazon Kendra , puede especificar la zona horaria del servidor para alinearla con el conector o índice de AEM. La zona horaria predeterminada de AEM On-Premise es la zona horaria del conector o índice de AEM de Amazon Kendra . La zona horaria predeterminada de AEM as a Cloud Service es la hora media de Greenwich.

d. Calendario de ejecución de sincronización, para frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione uno de los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice. Para agregar campos de origen de datos personalizados, cree un nombre de campo de índice para asignarlos y el tipo de datos del campo.

b. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Adobe Experience Manager

Debe especificar un JSON del esquema del origen de datos mediante la API TemplateConfiguration. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `AEM` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- URL del host de AEM: especifique la URL del host de Adobe Experience Manager. Por ejemplo, si utiliza AEM On-Premise, debe incluir el nombre de host y el puerto: https://hostname:port. O bien, si usa AEM as a Cloud Service, puede usar la URL del autor: https://author-xxxxxx-xxxxxxx.adobeaemcloud.com.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Tipo de autenticación: especifique qué tipo de autenticación quiere usar, ya sea `Basic` o `OAuth2`.
- Tipo de AEM: especifique qué tipo de Adobe Experience Manager va a utilizar, ya sea `CLOUD` o `ON_PREMISE`.
- Nombre de recurso de Amazon (ARN) secreto: si desea utilizar la autenticación básica para AEM On-Premise o Cloud, debe proporcionar un secreto que almacene las credenciales de autenticación de su nombre de usuario y contraseña. Usted proporciona el nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "aemUrl": "Adobe Experience Manager On-Premise host URL",
    "username": "user name with admin permissions",
    "password": "password with admin permissions"
}
```

Si quiere utilizar la autenticación OAuth 2.0 para AEM On-Premise, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "aemUrl": "Adobe Experience Manager host URL",
    "clientId": "client ID",
    "clientSecret": "client secret",
    "privateKey": "private key"
}
```

Si quiere utilizar la autenticación OAuth 2.0 para AEM as a Cloud Service, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "clientId": "client ID",
    "clientSecret": "client secret",
    "privateKey": "private key",
    "orgId": "organization ID",
```

```
    "technicalAccountId": "technical account ID",
    "imsHost": "Adobe Identity Management System (IMS) host"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas necesarias para el conector de Adobe Experience Manager y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Adobe Experience Manager.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- ID de zona horaria: si utiliza AEM On-Premise y la zona horaria del servidor es diferente a la zona horaria del conector o índice de Amazon Kendra AEM, puede especificar la zona horaria del servidor para alinearla con el conector o índice de AEM.

  La zona horaria predeterminada de AEM On-Premise es la zona horaria del conector o índice de AEM. Amazon Kendra La zona horaria predeterminada de AEM as a Cloud Service es la hora media de Greenwich.

  Para obtener información sobre los ID de zonas horarias compatibles, consulte Esquema JSON de Adobe Experience Manager.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinadas páginas y activos.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Rastreador de identidades: especifique si se debe activar el rastreador de identidades. Amazon Kendra El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.
- Asignaciones de campos: elija asignar los campos del origen de datos de Adobe Experience Manager a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el Esquema de plantilla de Adobe Experience Manager.

Alfresco

Alfresco es un servicio de administración de contenido que ayuda a los clientes a almacenar y administrar su contenido. Puede usarlo Amazon Kendra para indexar su biblioteca de Alfresco documentos, wiki y blog.

Amazon Kendra es compatible con el Alfresco entorno local y Alfresco en la nube (plataforma como servicio).

Puede conectarse Amazon Kendra a su fuente de Alfresco datos mediante la Amazon Kendra consola o la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Alfresco, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

El conector de origen de datos de Amazon Kendra Alfresco admite las siguientes características:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Autenticación OAuth 2.0 y básica
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar su fuente de datos de Alfresco, realice estos cambios en su Alfresco sitio web. Cuentas de AWS

En Alfresco, asegúrese de que:

- Ha copiado la URL del repositorio de Alfresco y la URL de la aplicación web. Si solo quiere indexar un sitio de Alfresco específico, copie también el ID del sitio.
- Ha apuntado sus credenciales de autenticación de Alfresco, que incluyen un nombre de usuario y una contraseña con al menos permisos de lectura. Si desea utilizar la autenticación OAuth 2.0, debe añadir el usuario al grupo de administradores de Alfresco.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Opcional: se configuraron las credenciales de OAuth 2.0 en. Alfresco Las credenciales incluyen el ID de cliente, el secreto del cliente y la URL del token. Para obtener más información sobre cómo configurar los clientes para Alfresco On-Premises, consulte la documentación de Alfresco. Si utiliza Alfresco Cloud (PaaS), debe ponerse en contacto con el servicio de asistencia de Hyland para obtener la autenticación OAuth 2.0 de Alfresco.
- Ha comprobado que cada documento es único en Alfresco y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En su interior Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Alfresco en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de datos de Alfresco. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Alfresco, debe proporcionar los detalles necesarios de su fuente de datos de Alfresco para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Alfresco para Amazon Kendra, consulte. Requisitos previos

Console

Para conectarse a Amazon Kendra Alfresco

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, seleccione el conector de Alfresco y, a continuación, seleccione Añadir conector. Si utiliza la versión 2 (si procede), elija el conector Alfresco con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. Alfrescoescriba: elija si va a utilizar un entorno Alfresco local, en servidor o en la Alfresco nube (plataforma como servicio).

   b. URL del repositorio de Alfresco: introduzca la URL del repositorio de Alfresco. Por ejemplo, si utiliza Alfresco Cloud (PaaS), la URL del repositorio podría ser https://company.alfrescocloud.com. O bien, si utiliza Alfresco On-Premises, la URL del repositorio podría ser https://company-alfresco-instance.company-domain.suffix:port.

   c. Aplicación de usuario de Alfresco. URL: introduzca la URL de la interfaz de usuario de Alfresco. Puede obtener la URL del repositorio de su administrador de Alfresco. Por ejemplo, la URL de la interfaz de usuario podría ser https://example.com.

   d. Ubicación del certificado SSL: introduzca la ruta al certificado SSL almacenado en un depósito. Amazon S3 Se utiliza para conectarse a Alfresco On-Premises mediante una conexión SSL segura.

   e. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

   f. Autenticación: elija la autenticación básica o la autenticación OAuth 2.0. A continuación, elija un secreto de Secrets Manager existente o cree uno nuevo para almacenar sus credenciales de Alfresco. Si decide crear un secreto nuevo, se abre una ventana AWS Secrets Manager secreta.

      Si ha elegido la autenticación básica, introduzca un nombre para el secreto, el nombre de usuario de Alfresco y la contraseña.

      Si ha elegido la autenticación OAuth 2.0, introduzca un nombre para el secreto, el ID de cliente, el secreto de cliente y la URL del token.

   g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

   h. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en

función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

i. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

j. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Ámbito de sincronización: establezca límites para rastrear determinado contenido y filtre el contenido mediante patrones de expresiones regex.

   b. i. Contenido: elija si rastrear contenido marcado con “Aspectos” en Alfresco, contenido de un sitio de Alfresco específico o contenido de todos sus sitios de Alfresco.

      ii. (Opcional) Configuración adicional: configure los siguientes ajustes:

      - Incluir comentarios: elija incluir comentarios en la biblioteca de documentos y en el blog de Alfresco.
      - Patrones regex: patrones de expresiones regulares para incluir o excluir determinados archivos.

   c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Seleccione uno de los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

   b. Para agregar campos de origen de datos personalizados, cree un nombre de campo de índice para asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Alfresco

Debe especificar un JSON del esquema del origen de datos mediante la API TemplateConfiguration. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `ALFRESCO` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- ID del sitio de Alfresco: especifique el ID del sitio de Alfresco.
- URL del repositorio de Alfresco: especifique la URL del repositorio de Alfresco. Puede obtener la URL del repositorio de su administrador de Alfresco. Por ejemplo, si utiliza Alfresco Cloud

(PaaS), la URL del repositorio podría ser https://company.alfrescocloud.com. O bien, si utiliza Alfresco On-Premises, la URL del repositorio podría ser https://company-alfresco-instance.company-domain.suffix:port.

- URL de la aplicación web de Alfresco: especifique la URL de la interfaz de usuario de Alfresco. Puede obtener la URL del repositorio de su administrador de Alfresco. Por ejemplo, la URL de la interfaz de usuario podría ser https://example.com.
- Tipo de autenticación: especifique el tipo de autenticación que desea usar, ya sea `OAuth2` o `Basic`.
- Tipo de Alfresco: especifique qué tipo de Alfresco que utiliza, ya sea `PAAS` (Cloud/Platform as a Service) o `ON_PREM` (On-Premises).
- Nombre de recurso de Amazon (ARN) secreto: si desea utilizar la autenticación básica, debe proporcionar un secreto que almacene las credenciales de autenticación de su nombre de usuario y contraseña. Usted proporciona el nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "username": "user name",
    "password": "password"
}
```

Si quiere utilizar la autenticación OAuth 2.0, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "clientId": "client ID",
    "clientSecret": "client secret",
    "tokenUrl": "token URL"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionarle a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas necesarias para el conector de Alfresco y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Alfresco.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Tipo de contenido: el tipo de contenido que quiere rastrear, ya sea contenido marcado con “Aspectos” en Alfresco, contenido de un sitio de Alfresco específico o contenido de todos sus sitios de Alfresco. También puede incluir contenido de “Aspectos” específicos.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse su índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en

todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

- Asignaciones de campos: elija asignar los campos del origen de datos de Alfresco a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el Esquema de plantilla de Alfresco.

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Alfresco, consulte:

- Busque contenido de forma inteligente mediante AlfrescoAmazon Kendra

Aurora (MySQL)

Aurora es un sistema de administración de bases de datos relacionales (RDBMS) creado para la nube. Si es un Aurora usuario, puede usarlo Amazon Kendra para indexar su Aurora (MySQL) fuente de datos. El conector Amazon Kendra Aurora (MySQL) de fuente de datos es compatible con Aurora MySQL 3 y Aurora Serverless MySQL 8.0.

Puede conectarse Amazon Kendra a su fuente de Aurora (MySQL) datos mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de la fuente de Amazon Kendra Aurora (MySQL) datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de Aurora (MySQL) datos, realice estos cambios en sus AWS cuentas Aurora (MySQL) y.

En Aurora (MySQL), asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

  > **Important**
  >
  > Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos. Puede encontrar esta información en la Amazon RDS consola.
- Ha comprobado que cada documento es único en Aurora (MySQL) y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Aurora (MySQL) en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de Aurora (MySQL) datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de Aurora (MySQL) datos, debe proporcionar los detalles de sus Aurora (MySQL) credenciales para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configuradoAurora (MySQL), Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Aurora (MySQL)

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el Aurora (MySQL)conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el Aurora (MySQL)conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. En Origen, introduzca la siguiente información:
   b. Host: ingrese la URL del host de la base de datos, por ejemplo: `http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`.
   c. Puerto: ingrese el puerto de la base de datos, por ejemplo, `5432`.
   d. Instancia: introduzca la instancia de la base de datos.
   e. En Autenticación, introduzca la siguiente información:
      - AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de Aurora (MySQL) autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Aurora (MySQL) -' se añade automáticamente a tu nombre secreto.

II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

B. Seleccione Guardar.

f. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

g. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

h. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

- Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Las consultas SQL deben tener menos de 32 KB y no contener puntos y comas (;). Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
- Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
- Columna de cuerpo: proporcione el nombre de la columna del cuerpo del documento en la tabla de la base de datos.

b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

- Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.
- Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
- Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
- Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.
- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Aurora (MySQL)

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `mySql`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra

por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:

- `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. Aurora (MySQL) El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. Aurora (MySQL) Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Aurora (MySQL).

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para sus documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos del origen de datos de Aurora (MySQL) a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Aurora Esquema de plantillas (MySQL).

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.

- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.
- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Aurora (PostgreSQL)

Aurora es un sistema de administración de bases de datos relacionales (RDBMS) creado para la nube. Si es un Aurora usuario, puede usarlo Amazon Kendra para indexar su Aurora (PostgreSQL) fuente de datos. El conector Amazon Kendra Aurora (PostgreSQL) de fuente de datos es compatible con Aurora PostgreSQL 1.

Puede conectarse Amazon Kendra a su fuente de Aurora (PostgreSQL) datos mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de la fuente de Amazon Kendra Aurora (PostgreSQL) datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de Aurora (PostgreSQL) datos, realice estos cambios en sus AWS cuentas Aurora (PostgreSQL) y.

En Aurora (PostgreSQL), asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos.
- Ha comprobado que cada documento es único en Aurora (PostgreSQL) y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Aurora (PostgreSQL) en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de Aurora (PostgreSQL) datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de Aurora (PostgreSQL) datos, debe proporcionar los detalles de sus Aurora (PostgreSQL) credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configuradoAurora (PostgreSQL), Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Aurora (PostgreSQL)

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el Aurora (PostgreSQL)conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el Aurora (PostgreSQL)conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. En Origen, introduzca la siguiente información:

b. Host: ingrese la URL del host de la base de datos, por ejemplo: `http://instance URL.region.rds.amazonaws.com`.

c. Puerto: ingrese el puerto de la base de datos, por ejemplo, `5432`.

d. Instancia: ingrese la instancia de la base de datos, por ejemplo `postgres`.

e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.

f. En Autenticación, introduzca la siguiente información:

- AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de Aurora (PostgreSQL) autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

  A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

     I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Aurora (PostgreSQL) -' se añade automáticamente a tu nombre secreto.

     II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

  B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

      - Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Las consultas SQL deben tener menos de 32 KB y no contener puntos y comas (;). Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
      - Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
      - Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
      - Columna de cuerpo: proporcione el nombre de la columna del cuerpo del documento en la tabla de la base de datos.

   b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

      - Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.
      - Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
      - Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
      - Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.

- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Aurora (PostgreSQL)

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `postgresql`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de

datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. Aurora (PostgreSQL) El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. Aurora (PostgreSQL) Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Aurora (PostgreSQL).

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para sus documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

- Asignaciones de campos: elija asignar los campos del origen de datos de Aurora (PostgreSQL) a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Aurora Esquema de plantillas (PostgreSQL).

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.
- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.
- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Amazon FSx (Windows)

Amazon FSx (Windows) es un sistema de servidor de archivos totalmente gestionado y basado en la nube que ofrece capacidades de almacenamiento compartido. Si es usuario de Amazon FSx (Windows), puede utilizarlo Amazon Kendra para indexar su fuente de datos Amazon FSx (Windows).

> **Note**
>
> Amazon Kendra ahora es compatible con un conector actualizado Amazon FSx (Windows). La consola se ha actualizado automáticamente. Todos los conectores nuevos que cree en la consola utilizarán la arquitectura actualizada. Si usa la API, ahora debe usar el TemplateConfigurationobjeto en lugar del `FSxConfiguration` objeto para configurar el conector.
> Los conectores configurados con la antigua arquitectura de consola y API seguirán funcionando tal y como estaban configurados. Sin embargo, no podrá editarlos ni actualizarlos. Si desea editar o actualizar la configuración del conector, debe crear uno nuevo.
> Se recomienda migrar el flujo de trabajo del conector a la versión actualizada. Está previsto que el soporte para los conectores configurados con la arquitectura anterior finalice en junio de 2024.

Puede conectarse Amazon Kendra a su fuente de datos Amazon FSx (Windows) mediante la Amazon Kendra consola o la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuente de datos Amazon Kendra Amazon FSx (Windows), consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra Amazon FSx El conector de fuente de datos (Windows) admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Rastreo de identidad de usuario

- Filtros de inclusión y exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar su fuente de datos Amazon FSx (Windows), compruebe los detalles de su fuente de datos Amazon FSx (Windows) y Cuentas de AWS.

Para Amazon FSx (Windows), asegúrese de tener:

- Configure Amazon FSx (Windows) con permisos de lectura y montaje.
- Apuntó el ID de su sistema de archivos. Puede encontrar el ID de su sistema de archivos en el panel de sistemas de archivos de la consola Amazon FSx (Windows).
- Configuró una nube privada virtual utilizando el Amazon VPC lugar donde reside su sistema de archivos Amazon FSx (Windows).
- Apuntó sus credenciales de autenticación Amazon FSx (de Windows) para una cuenta Active Directory de usuario. Esto incluye su nombre de usuario de Active Directory con su nombre de dominio DNS (por ejemplo, user@corp.example.com) y contraseña.

  > **Note**
  >
  > Utilice únicamente las credenciales necesarias para que el conector funcione. No utilice credenciales privilegiadas como las de administrador del dominio.

  > **Note**
  >
  > Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha comprobado que cada documento es único en Amazon FSx (Windows) y en otras fuentes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Guardó sus credenciales de autenticación Amazon FSx (Windows) en un AWS Secrets Manager secreto y, si utiliza la API, anotó el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y Secrets Manager secreto al conectar su fuente de datos Amazon FSx (Windows) a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos Amazon FSx (Windows), debe proporcionar los detalles necesarios de su fuente de datos Amazon FSx (Windows) para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configurado Amazon FSx (Windows) Amazon Kendra, consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a su sistema de archivos Amazon FSx (Windows)

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.

2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector Amazon FSx (Windows) y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el conector Amazon FSx (Windows) con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

   e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. Amazon FSx ID del sistema de archivos (Windows): seleccione en el menú desplegable su ID de sistema de archivos existente, obtenido de Amazon FSx (Windows). O bien, cree un sistema de Amazon FSx archivos (Windows). Puede encontrar el identificador del sistema de archivos en el panel de sistemas de archivos de la consola Amazon FSx (Windows).

   b. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

c. Autenticación: elija un AWS Secrets Manager secreto existente o cree uno nuevo para almacenar las credenciales del sistema de archivos. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

   Proporcione un secreto que almacene las credenciales de autenticación de su nombre de usuario y contraseña. El nombre de usuario debe incluir su nombre de dominio DNS. Por ejemplo, user@corp.example.com.

   Guarda y añade tu secreto.

d. Virtual Private Cloud (VPC): debe seleccionar un Amazon VPC lugar donde resida su Amazon FSx (Windows). Incluye la subred de la VPC y los grupos de seguridad. Consulte Configurar un. Amazon VPC

e. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

f. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Ámbito de sincronización, patrones de expresiones regulares: añada patrones de expresiones regulares para incluir o excluir determinados archivos.

   b. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

      - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
      - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice.

Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

c. Calendario de ejecución de la sincronización: en Frecuencia, elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualice el índice.

d. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione entre los campos predeterminados Amazon Kendra generados de sus archivos que desee asignar a su índice. Para agregar campos de origen de datos personalizados, cree un nombre de campo de índice para asignarlos y el tipo de datos del campo.

b. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a su sistema de archivos Amazon FSx (Windows)

Debe especificar un JSON del esquema del origen de datos mediante la API TemplateConfiguration. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `FSX` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- ID del sistema de archivos: el identificador del sistema de archivos Amazon FSx (Windows). Puede encontrar el ID del sistema de archivos en el panel de sistemas de archivos de la consola Amazon FSx (Windows).
- Tipo de sistema de archivos: especifique el tipo de sistema de archivos como `WINDOWS`.
- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.

> **Note**
>
> Debe seleccionar el Amazon VPC lugar donde reside su Amazon FSx (Windows). Incluye la subred de la VPC y los grupos de seguridad.

- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.
- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su Amazon FSx cuenta (Windows). El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "username": "user@corp.example.com",
    "password": "password"
```

```
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionarle a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector Amazon FSx (Windows) y. Amazon Kendra Para obtener más información, consulte las IAM funciones de las fuentes de datos Amazon FSx (Windows).

También puede añadir las siguientes características opcionales:

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Lista de control de acceso (ACL): especifique si desea rastrear la información de la ACL de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

> **Note**
>
> Para probar el filtrado de contexto de usuario con un usuario, debe incluir el nombre de dominio DNS como parte del nombre de usuario al realizar la consulta. Debe disponer de permisos administrativos del dominio de Active Directory. También puede probar el filtrado de contexto de usuario con el nombre de un grupo.

- Asignaciones de campos: elija asignar los campos de la fuente de datos Amazon FSx (Windows) a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice `_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el esquema de plantillas Amazon FSx (Windows).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos Amazon FSx (Windows), consulte:

- Busque de forma segura datos no estructurados en los sistemas de archivos de Windows con el Amazon Kendra conector para Amazon FSx (Windows) para Windows File Server.

Amazon FSx (NetApp DISPONIBLE)

Amazon FSx (NetApp ONTAP) es un sistema de servidor de archivos totalmente gestionado y basado en la nube que ofrece capacidades de almacenamiento compartido. Si es usuario de Amazon FSx (NetApp ONTAP), puede utilizarlo Amazon Kendra para indexar su fuente de datos Amazon FSx (NetApp ONTAP).

Puede conectarse Amazon Kendra a su fuente de datos Amazon FSx (NetApp ONTAP) mediante la Amazon Kendra consola o la API. TemplateConfiguration

Para solucionar problemas del conector de fuente de datos Amazon Kendra Amazon FSx (NetApp ONTAP), consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión

Características admitidas

Amazon Kendra Amazon FSx El conector de fuente de datos (NetApp ONTAP) admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión y exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar su fuente de datos Amazon FSx (NetApp ONTAP), compruebe los detalles de su Amazon FSx (NetApp ONTAP) y. Cuentas de AWS

Para Amazon FSx (NetApp ONTAP), asegúrese de tener:

- Configure Amazon FSx (NetApp ONTAP) con permisos de lectura y montaje.
- Apuntó el ID de su sistema de archivos. Puede encontrar el ID de su sistema de archivos en el panel de sistemas de archivos de la consola Amazon FSx (NetApp ONTAP).
- Apuntó el ID de la máquina virtual de almacenamiento (SVM) utilizado con su sistema de archivos. Para encontrar su ID de SVM, vaya al panel de sistemas de archivos de la consola Amazon FSx (NetApp ONTAP), seleccione su ID de sistema de archivos y, a continuación, seleccione Máquinas virtuales de almacenamiento.
- Configuró una nube privada virtual utilizando el Amazon VPC lugar donde reside su sistema de archivos Amazon FSx (NetApp ONTAP).
- Apuntó sus credenciales de autenticación Amazon FSx (NetApp ONTAP) para una cuenta de Active Directory usuario. Esto incluye su nombre de usuario de Active Directory con su nombre de dominio DNS (por ejemplo, user@corp.example.com) y contraseña. Si utiliza el protocolo Network File System (NFS) para su sistema de archivos Amazon FSx (NetApp ONTAP), las credenciales de autenticación incluyen un identificador izquierdo, un identificador derecho y una clave previamente compartida.

> **Note**
>
> Utilice únicamente las credenciales necesarias para que el conector funcione. No utilice credenciales privilegiadas como las de administrador del dominio.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha comprobado que cada documento es único en Amazon FSx (NetApp ONTAP) y en otras fuentes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Guardó sus credenciales de autenticación Amazon FSx (NetApp ONTAP) en un AWS Secrets Manager secreto y, si utiliza la API, anotó el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y Secrets Manager secreto al conectar su fuente de datos Amazon FSx (NetApp ONTAP). Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos Amazon FSx (NetApp ONTAP), debe proporcionar los detalles necesarios de su fuente de datos Amazon FSx (NetApp ONTAP) para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Amazon FSx (NetApp ONTAP) para Amazon Kendra, consulte. Requisitos previos

Console

Para conectarse Amazon Kendra a su sistema de Amazon FSx archivos (NetApp ONTAP)

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector Amazon FSx (NetApp ONTAP) y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el conector Amazon FSx (NetApp ONTAP) con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. Fuente: proporcione la información de su sistema de archivos.

- Protocolo del sistema de archivos: elija el protocolo de su sistema de archivos Amazon FSx (NetApp ONTAP). Puede elegir el protocolo Common Internet File System (CIFS) o el protocolo Network File System (NFS) para Linux.
- Amazon FSx ID del sistema de archivos (NetApp ONTAP): seleccione en el menú desplegable su ID de sistema de archivos actual, obtenido de (ONTAP). Amazon FSx NetApp O bien, cree un sistema de archivos Amazon FSx (ONTAP) NetApp . Puede encontrar el ID del sistema de archivos en el panel de sistemas de archivos de la consola Amazon FSx (NetApp ONTAP).
- ID de SVM (Amazon FSx (NetApp ONTAP) NetApp ONTAP solo para): proporcione el ID de la máquina virtual de almacenamiento (SVM) de su (ONTAP). Amazon FSx NetApp NetApp ONTAP Para encontrar su ID de SVM, vaya al panel de sistemas de archivos de la consola Amazon FSx (NetApp ONTAP), seleccione su ID de sistema de archivos y seleccione Máquinas virtuales de almacenamiento.

b. Autorización: active o desactive la información de la lista de control de acceso (ACL) en sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

c. Autenticación: elija un AWS Secrets Manager secreto existente o cree uno nuevo para almacenar las credenciales del sistema de archivos. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

Proporcione un secreto que almacene las credenciales de autenticación de su nombre de usuario y contraseña. El nombre de usuario debe incluir su nombre de dominio DNS. Por ejemplo, user@corp.example.com.

Si utiliza el protocolo NFS para su sistema de archivos Amazon FSx (NetApp ONTAP), proporcione un secreto que almacene sus credenciales de autenticación: el identificador izquierdo, el identificador derecho y la clave previamente compartida.

Guarde y añada su secreto.

d. Virtual Private Cloud (VPC): debe seleccionar un Amazon VPC lugar en el que resida su Amazon FSx (ONTAP). NetApp Incluye la subred de la VPC y los grupos de seguridad. Consulte Configurar un. Amazon VPC

e. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

f. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. Ámbito de sincronización, patrones de expresiones regulares: añada patrones de expresiones regulares para incluir o excluir determinados archivos.

b. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

c. Calendario de ejecución de la sincronización: en Frecuencia, elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualice el índice.

d. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Seleccione entre los campos predeterminados Amazon Kendra generados de sus archivos que desee asignar a su índice. Para agregar campos de origen de datos personalizados, cree un nombre de campo de índice para asignarlos y el tipo de datos del campo.

   b. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a su sistema de archivos Amazon FSx (NetApp ONTAP)

Debe especificar un JSON del esquema del origen de datos mediante la API TemplateConfiguration. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `FSXONTAP` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- ID del sistema de archivos: el identificador del sistema de archivos Amazon FSx (NetApp ONTAP). Puede encontrar el ID del sistema de archivos en el panel de sistemas de archivos de la consola Amazon FSx (NetApp ONTAP).
- ID de SVM: el ID de la máquina virtual de almacenamiento (SVM) que se utiliza con el sistema de archivos. Para encontrar su ID de SVM, vaya al panel de sistemas de archivos de la consola Amazon FSx (NetApp ONTAP), seleccione su ID de sistema de archivos y, a continuación, seleccione Máquinas virtuales de almacenamiento.
- Tipo de protocolo: especifique si utiliza el protocolo Common Internet File System (CIFS) o el protocolo Network File System (NFS) para Linux.
- Tipo de sistema de archivos: especifique el tipo de sistema de archivos como uno de los dos. `FSXONTAP`

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.

  > **Note**
  >
  > Debe seleccionar un Amazon VPC lugar en el que resida su Amazon FSx (NetApp ONTAP). Incluye la subred de la VPC y los grupos de seguridad.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su Amazon FSx cuenta (ONTAP). NetApp El secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "username": "user@corp.example.com",
      "password": "password"
  }
  ```

  Si utiliza el protocolo NFS para su sistema de archivos Amazon FSx (NetApp ONTAP), el secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "leftId": "left ID",
      "rightId": "right ID",
      "preSharedKey": "pre-shared key"
  }
  ```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector Amazon FSx (NetApp ONTAP) y. Amazon Kendra Para obtener más información, consulte las IAM funciones de las fuentes de datos Amazon FSx (NetApp ONTAP).

También puede añadir las siguientes características opcionales:

- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar

una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:

- `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Lista de control de acceso (ACL): especifique si desea rastrear la información de la ACL de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

  > **Note**
  >
  > Para probar el filtrado de contexto de usuario con un usuario, debe incluir el nombre de dominio DNS como parte del nombre de usuario al realizar la consulta. Debe disponer de permisos administrativos del dominio de Active Directory. También puede probar el filtrado de contexto de usuario con el nombre de un grupo.

- Asignaciones de campos: elija asignar los campos de la fuente de datos Amazon FSx (NetApp ONTAP) a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice `_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el esquema de plantillas Amazon FSx (NetApp ONTAP).

Amazon RDS/Aurora

Puede indexar los documentos que están almacenados en una base de datos mediante un origen de datos de base de datos. Después de proporcionar la información de conexión a la base de datos, Amazon Kendra conecta e indexa los documentos.

Amazon Kendra admite las siguientes bases de datos:

- Amazon Aurora MySQL
- Amazon Aurora PostgreSQL
- Amazon RDS para MySQL
- Amazon RDS para PostgreSQL

> **Note**
>
> No se admiten las bases de datos Aurora sin servidor.

> **Important**
>
> Está previsto que este conector Amazon RDS/Aurora deje de estar disponible a finales de 2023.
> Amazon Kendra ahora admite nuevos conectores de fuentes de datos de bases de datos. Para mejorar la experiencia, le recomendamos que elija uno de los siguientes nuevos conectores para su caso de uso:

- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon RDS (MySQL)
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- IBM DB2
- Microsoft SQL Server
- MySQL
- Oracle Database
- PostgreSQL

Puede conectarse Amazon Kendra a la fuente de datos de su base de datos mediante la Amazon Kendra consola y la DatabaseConfigurationAPI.

Para solucionar problemas del conector de fuentes Amazon Kendra de datos de su base de datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión

Características admitidas

Amazon Kendra el conector de fuente de datos de base de datos admite las siguientes funciones:

- Asignaciones de campo
- Filtrado de contexto de usuario
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder usarlo Amazon Kendra para indexar la fuente de datos de la base de datos, realice estos cambios en la base de datos y en AWS las cuentas.

En la base de datos, asegúrese de que:

- Ha apuntado las credenciales de autenticación básica del nombre de usuario y la contraseña de la base de datos.
- Ha copiado el nombre de host, el número de puerto, la dirección de host, el nombre de la base de datos y el nombre de la tabla de datos que contiene los datos del documento. En el caso de PostgreSQL, la tabla de datos debe ser una tabla pública o un esquema público.

  > **Note**
  >
  > El host y el puerto indican Amazon Kendra dónde encontrar el servidor de base de datos en Internet. El nombre de la base de datos y el nombre de la tabla indican Amazon Kendra dónde encontrar los datos del documento en el servidor de la base de datos.

- Ha copiado los nombres de las columnas de la tabla de datos que contienen los datos del documento. Debe incluir el ID del documento, el cuerpo del documento, las columnas para detectar si un documento ha cambiado (por ejemplo, la columna actualizada por última vez) y las columnas opcionales de la tabla de datos que se asignan a campos de índice personalizados. También puede asignar cualquiera de los nombres de campo reservados de Amazon Kendra a una columna de la tabla.
- Se ha copiado la información del tipo de motor de base de datos, por ejemplo, si se utiliza Amazon RDS para MySQL u otro tipo.
- Ha comprobado que cada documento es único en la base de datos y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de la base de datos en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar la fuente de datos de la base de datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a la fuente de datos de la base de datos, debe proporcionar los detalles necesarios de la fuente de datos de la base de datos para Amazon Kendra poder acceder a los datos. Si aún no ha configurado la base de datos para Amazon Kendra, consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a una base de datos

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector de base de datos y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el conector de base de datos con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. Punto de conexión: un nombre de host DNS, una dirección IPv4 o una dirección IPv6.

   b. Puerto: un número de puerto.

   c. Base de datos: nombre de la base de datos.

   d. Nombre de tabla: nombre de la tabla.

   e. En Tipo de autenticación, elija entre Existente y Nuevo para almacenar las credenciales de autenticación de la base de datos. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

      - Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

        A. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-database-» se añade automáticamente a tu nombre secreto.

B. En Nombre de usuario y Contraseña: introduzca los valores de las credenciales de autenticación de su cuenta de base de datos.

C. Seleccione Guardar autenticación.

f. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

> **Note**
>
> Debe utilizar una subred privada. Si la instancia de RDS está en una subred pública en la VPC, puede crear una subred privada que tenga acceso saliente a una puerta de enlace NAT en la subred pública. Las subredes proporcionadas en la configuración de VPC deben estar en Oeste de EE. UU. (Oregón), Este de EE. UU. (Norte de Virginia), Europa (Irlanda).

g. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

h. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. Seleccione entre Aurora MySQL, MySQL, Aurora PostgreSQL y PostgreSQL según el caso de uso.

b. Incluir los identificadores SQL entre comillas dobles: seleccione esta opción para incluir los identificadores SQL entre comillas dobles. Por ejemplo, “columnName”.

c. Columna ACL y columnas de detección de cambios: configure las columnas que se Amazon Kendra utilizan para la detección de cambios (por ejemplo, la columna actualizada por última vez) y su lista de control de acceso.

d. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que Amazon Kendra se sincronizará con la fuente de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Amazon Kendra asignaciones de campos predeterminadas: seleccione entre las fuentes de datos predeterminadas Amazon Kendra generadas los campos que desee asignar a su índice. Debe agregar los valores de la Columna de base de datos para `document_id` y `document_body`

   b. Asignaciones de campo personalizado: para agregar campos de origen de datos personalizados a fin de crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a una base de datos

Debe especificar la siguiente DatabaseConfigurationAPI:

- ColumnConfiguration—Información sobre dónde debe obtener el índice la información del documento de la base de datos. Para obtener más información, consulte ColumnConfiguration. Debe especificar los campos `DocumentDataColumnName` (cuerpo del documento o texto principal), `DocumentIdColumnName` y `ChangeDetectingColumn` (por ejemplo, la columna actualizada por última vez). La columna asignada al campo `DocumentIdColumnName` debe ser una columna de números enteros. En el siguiente ejemplo se muestra una configuración simple de columnas para un origen de datos de base de datos:

```
"ColumnConfiguration": {
    "ChangeDetectingColumns": [
        "LastUpdateDate",
        "LastUpdateTime"
    ],
    "DocumentDataColumnName": "TextColumn",
    "DocumentIdColumnName": "IdentifierColumn",
    "DocoumentTitleColumnName": "TitleColumn",
    "FieldMappings": [
        {
```

```
            "DataSourceFieldName": "AbstractColumn",
            "IndexFieldName": "Abstract"
        }
    ]
}
```

- ConnectionConfiguration—Información de configuración necesaria para conectarse a una base de datos. Para obtener más información, consulte ConnectionConfiguration.
- DatabaseEngineType—El tipo de motor de base de datos que ejecuta la base de datos. El `DatabaseHost` campo `ConnectionConfiguration` debe ser el punto final Amazon Relational Database Service (Amazon RDS) de la instancia de la base de datos. No utilice el punto de conexión del clúster.
- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta de base de datos. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "username": "user name",
    "password": "password"
}
```

En el siguiente ejemplo se muestra una configuración de base de datos que incluye el ARN secreto.

```
"DatabaseConfiguration": {
"ConnectionConfiguration": {
"DatabaseHost": "host.subdomain.domain.tld",
        "DatabaseName": "DocumentDatabase",
        "DatabasePort": 3306,
        "SecretArn": "arn:aws:secretmanager:region:account ID:secret/secret name",
        "TableName": "DocumentTable"
    }
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de base de datos y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de base de datos.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique `VpcConfiguration` como parte de la configuración del origen de datos. Consulte Configuración de Amazon Kendra para utilizar una VPC.

> **Note**
>
> Debe utilizar solo una subred privada. Si la instancia de RDS está en una subred pública en la VPC, puede crear una subred privada que tenga acceso saliente a una puerta de enlace NAT en la subred pública. Las subredes proporcionadas en la configuración de VPC deben estar en Oeste de EE. UU. (Oregón), Este de EE. UU. (Norte de Virginia), Europa (Irlanda).

- Asignaciones de campos: elija asignar los campos del origen de datos de base de datos a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

Amazon RDS (Microsoft SQL Server)

SQL Server es un sistema de administración de bases de datos desarrollado por Microsoft. Amazon RDS for SQL Server facilita la configuración, el funcionamiento y el escalado de las implementaciones de SQL Server en la nube. Si es un usuario Amazon RDS (Microsoft SQL Server), puede utilizarlo Amazon Kendra para indexar su fuente de datos Amazon RDS (Microsoft SQL Server). El conector de fuente de datos Amazon Kendra JDBC es compatible con Microsoft SQL Server 2019.

Puede conectarse Amazon Kendra a su fuente de datos Amazon RDS (Microsoft SQL Server) mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas de su conector de fuente de datos Amazon Kendra Amazon RDS (Microsoft SQL Server), consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de que pueda Amazon Kendra utilizarla para indexar su fuente de datos Amazon RDS (Microsoft SQL Server), realice estos cambios en su Amazon RDS (Microsoft SQL Server) y en sus AWS cuentas.

En Amazon RDS (Microsoft SQL Server), asegúrese de tener:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos.
- Marcó que cada documento es único en Amazon RDS (Microsoft SQL Server) y en otras fuentes de datos que planea usar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Guardó sus credenciales de autenticación Amazon RDS (Microsoft SQL Server) en AWS Secrets Manager secreto y, si utiliza la API, anotó el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y Secrets Manager secreto al conectar su fuente de datos Amazon RDS (Microsoft SQL Server) a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos Amazon RDS (Microsoft SQL Server), debe proporcionar detalles de sus credenciales Amazon RDS (Microsoft SQL Server) para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configurado Amazon RDS (Microsoft SQL Server), Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Amazon RDS (Microsoft SQL Server)

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector Amazon RDS (Microsoft SQL Server) y, a continuación, elija Agregar conector. Si usa la versión 2 (si corresponde), elija el conector Amazon RDS (Microsoft SQL Server) con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. En Origen, introduzca la siguiente información:

b. Host: introduzca el nombre del host de la base de datos.

c. Puerto: introduzca el puerto de la base de datos.

d. Instancia: introduzca la instancia de la base de datos.

e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.

f. En Autenticación, introduzca la siguiente información:

- AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de autenticación Amazon RDS (Microsoft SQL Server). Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

  A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

     I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Amazon RDS (Microsoft SQL Server) -' se añade automáticamente a su nombre secreto.

     II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

  B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

   - Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

     > **Note**
     >
     > Si el nombre de una tabla incluye caracteres especiales (no alfanuméricos), debe colocar corchetes alrededor del nombre de la tabla. Por ejemplo, `seleccione * de [] my-database-table`

   - Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
   - Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
   - Columna de cuerpo: proporcione el nombre de la columna del cuerpo del documento en la tabla de la base de datos.

   b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

   - Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.
   - Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.

- Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
- Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.
- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Amazon RDS (Microsoft SQL Server)

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `sqlserver`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

> **Note**
>
> Si el nombre de una tabla incluye caracteres especiales (no alfanuméricos), debe colocar corchetes alrededor del nombre de la tabla. Por ejemplo, *`seleccione * de [] my-database-table`*

- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no

seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:

- `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre de recurso secreto de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta (Amazon RDS Microsoft SQL Server). El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector Amazon RDS (Microsoft SQL Server) y Amazon Kendra. Para obtener más información, consulte IAM funciones para las fuentes de datos Amazon RDS (Microsoft SQL Server).

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para los documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos de la fuente de datos (de Amazon RDS Microsoft SQL Server) a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Amazon RDS Esquema de plantillas (Microsoft SQL Server).

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.
- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.

- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Amazon RDS (MySQL)

Amazon RDS (Amazon Relational Database Service) es un servicio web que facilita la configuración, el funcionamiento y el escalado de una base de datos relacional en AWS la nube. Si es un Amazon RDS usuario, puede utilizarlo Amazon Kendra para indexar su fuente de Amazon RDS (MySQL) datos. El conector de la fuente de Amazon Kendra datos admite las versiones Amazon RDS MySql 5.6, 5.7 y 8.0.

Puede conectarse Amazon Kendra a su fuente Amazon RDS (MySQL) de datos mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de la fuente de Amazon Kendra Amazon RDS (MySQL) datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de Amazon RDS (MySQL) datos, realice estos cambios en sus AWS cuentas Amazon RDS (MySQL) y.

En Amazon RDS (MySQL), asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos. Puede encontrar esta información en la Amazon RDS consola.
- Ha comprobado que cada documento es único en Amazon RDS (MySQL) y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Amazon RDS (MySQL) en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de Amazon RDS (MySQL) datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de Amazon RDS (MySQL) datos, debe proporcionar los detalles de sus Amazon RDS (MySQL) credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configuradoAmazon RDS (MySQL), Amazon Kendra consulte Requisitos previos.

Console

Para conectarse Amazon Kendra a Amazon RDS (MySQL)

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el Amazon RDS (MySQL)conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el Amazon RDS (MySQL)conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. En Origen, introduzca la siguiente información:

b. Host: ingrese la URL del host de la base de datos, por ejemplo: `http://instance URL.region.rds.amazonaws.com`.

c. Puerto: ingrese el puerto de la base de datos, por ejemplo, `5432`.

d. Instancia: ingrese la instancia de la base de datos, por ejemplo `postgres`.

e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.

f. En Autenticación, introduzca la siguiente información:

- AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de Amazon RDS (MySQL) autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

  A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

     I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Amazon RDS (MySQL) -' se añade automáticamente a tu nombre secreto.

     II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

  B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

      - Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Las consultas SQL deben tener menos de 32 KB y no contener puntos y comas (;). Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
      - Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
      - Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
      - Columna de cuerpo: proporcione el nombre de la columna del cuerpo del documento en la tabla de la base de datos.

   b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

      - Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.
      - Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
      - Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
      - Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.

- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Amazon RDS (MySQL)

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `mySql`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de

datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. Amazon RDS (MySQL) El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. Amazon RDS (MySQL) Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Amazon RDS (MySQL).

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Asignaciones de campos: elija asignar los campos del origen de datos de Amazon RDS (MySQL) a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Amazon RDS Esquema de plantillas (MySQL).

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.
- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.
- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Amazon RDS (Oracle)

Amazon RDS (Amazon Relational Database Service) es un servicio web que facilita la configuración, el funcionamiento y el escalado de una base de datos relacional en AWS la nube. Si es un Amazon RDS (Oracle) usuario, puede utilizarlo Amazon Kendra para indexar su fuente de Amazon RDS

(Oracle) datos. El conector Amazon Kendra Amazon RDS (Oracle) de fuente de datos es compatible con Amazon RDS Oracle Database 21c, Oracle Database 19c y Oracle Database 12c.

Puede conectarse Amazon Kendra a su fuente de Amazon RDS (Oracle) datos mediante la Amazon Kendra consola y la API. TemplateConfiguration

Para solucionar problemas del conector de la fuente de Amazon Kendra Amazon RDS (Oracle) datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de Amazon RDS (Oracle) datos, realice estos cambios en sus AWS cuentas Amazon RDS (Oracle) y.

En Amazon RDS (Oracle), asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos.
- Ha comprobado que cada documento es único en Amazon RDS (Oracle) y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Amazon RDS (Oracle) en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de Amazon RDS (Oracle) datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de Amazon RDS (Oracle) datos, debe proporcionar los detalles de sus Amazon RDS (Oracle) credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configuradoAmazon RDS (Oracle), Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Amazon RDS (Oracle)

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el Amazon RDS (Oracle)conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el Amazon RDS (Oracle)conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. En Origen, introduzca la siguiente información:
   b. Host: introduzca el nombre del host de la base de datos.
   c. Puerto: introduzca el puerto de la base de datos.
   d. Instancia: introduzca la instancia de la base de datos.
   e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.

f. En Autenticación, introduzca la siguiente información:

- AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de Amazon RDS (Oracle) autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

  A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

     I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Amazon RDS (Oracle) -' se añade automáticamente a tu nombre secreto.

     II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

  B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

- Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.

- Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
- Columna de cuerpo: proporcione el nombre de la columna de cuerpo del documento en la tabla de la base de datos.

b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

- Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.
- Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
- Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
- Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.
- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.

- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Amazon RDS (Oracle)

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `oracle`.

- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. Amazon RDS (Oracle) El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. Amazon RDS (Oracle) Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Amazon RDS (Oracle).

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para sus documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos del origen de datos de Amazon RDS (Oracle) a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Amazon RDS Esquema de plantillas (Oracle).

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.
- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.
- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Amazon RDS (PostgreSQL)

Amazon RDS es un servicio web que facilita la configuración, el funcionamiento y el escalado de una base de datos relacional en la AWS nube. Si es un Amazon RDS usuario, puede usarlo Amazon Kendra para indexar su fuente de Amazon RDS (PostgreSQL) datos. El conector Amazon Kendra Amazon RDS (PostgreSQL) de fuente de datos es compatible con PostgreSQL 9.6.

Puede conectarse Amazon Kendra a su fuente de Amazon RDS (PostgreSQL) datos mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de la fuente de Amazon Kendra Amazon RDS (PostgreSQL) datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo

- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de Amazon RDS (PostgreSQL) datos, realice estos cambios en sus AWS cuentas Amazon RDS (PostgreSQL) y.

En Amazon RDS (PostgreSQL), asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos. Puede encontrar esta información en la Amazon RDS consola.
- Ha comprobado que cada documento es único en Amazon RDS (PostgreSQL) y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Amazon RDS (PostgreSQL) en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de Amazon RDS (PostgreSQL) datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de Amazon RDS (PostgreSQL) datos, debe proporcionar los detalles de sus Amazon RDS (PostgreSQL) credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configuradoAmazon RDS (PostgreSQL), Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Amazon RDS (PostgreSQL)

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.

4. En la página Agregar fuente de datos, elija el Amazon RDS (PostgreSQL)conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el Amazon RDS (PostgreSQL)conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

    a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

    b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

    c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

    d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

    e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:

    a. En Origen, introduzca la siguiente información:

    b. Host: ingrese la URL del host de la base de datos, por ejemplo: `http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`.

    c. Puerto: ingrese el puerto de la base de datos, por ejemplo, `5432`.

    d. Instancia: ingrese la instancia de la base de datos, por ejemplo `postgres`.

    e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.

    f. En Autenticación, introduzca la siguiente información:

        - AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de Amazon RDS (PostgreSQL) autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

            A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

                I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Amazon RDS (PostgreSQL) -' se añade automáticamente a tu nombre secreto.

II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

- Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas SQL deben tener menos de 32 KB. Las consultas SQL deben tener menos de 32 KB y no contener puntos y comas (;). Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
- Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
- Columna de cuerpo: proporcione el nombre de la columna del cuerpo del documento en la tabla de la base de datos.

b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

- Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra

volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.

- Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
- Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
- Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.
- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

   - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
   - Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
   - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Amazon RDS (PostgreSQL)

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `postgresql`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:

  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. Amazon RDS (PostgreSQL) El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. Amazon RDS (PostgreSQL) Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Amazon RDS (PostgreSQL).

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para sus documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos del origen de datos de Amazon RDS (PostgreSQL) a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Amazon RDS Esquema de plantillas (PostgreSQL).

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.
- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.

- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Amazon S3

Amazon S3 es un servicio de almacenamiento de objetos que almacena datos como objetos dentro de cubos. Puedes usarlo Amazon Kendra para indexar el repositorio de documentos de tu Amazon S3 depósito.

> **Warning**
>
> Amazon Kendra no utiliza una política de bucket que conceda permisos a un Amazon Kendra director para interactuar con un bucket de S3. En su lugar, usa IAM roles. Asegúrate de Amazon Kendra no incluirlo como miembro de confianza en tu política de grupos para evitar problemas de seguridad de los datos al conceder permisos accidentalmente a directores arbitrarios. Sin embargo, puede añadir una política de bucket para utilizar un bucket de Amazon S3 en distintas cuentas. Para obtener más información, consulte Políticas para usar Amazon S3 en varias cuentas (en la pestaña de roles de IAM de S3, en la sección Roles de IAM para orígenes de datos). Para obtener información sobre las IAM funciones de las fuentes de datos de S3, consulte las IAM funciones.

> **Note**
>
> Amazon Kendra ahora es compatible con un Amazon S3 conector actualizado.
> La consola se ha actualizado automáticamente para usted. Todos los conectores nuevos que cree en la consola utilizarán la arquitectura actualizada. Si usa la API, ahora debe usar el TemplateConfigurationobjeto en lugar del `S3DataSourceConfiguration` objeto para configurar el conector.
> Los conectores configurados con la antigua arquitectura de consola y API seguirán funcionando tal y como estaban configurados. Sin embargo, no podrá editarlos ni actualizarlos. Si desea editar o actualizar la configuración del conector, debe crear un conector nuevo.

Se recomienda migrar el flujo de trabajo del conector a la versión actualizada. Está previsto que el soporte para los conectores configurados con la arquitectura anterior finalice en junio de 2024.

Puede conectarse a su fuente de Amazon S3 datos mediante la Amazon Kendra consola o la TemplateConfigurationAPI.

> **Note**
>
> Para generar un informe de estado de sincronización para su fuente de Amazon S3 datos, consulte Solución de problemas con las fuentes de datos.

Para solucionar problemas del conector de fuente de datos Amazon Kendra S3, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Creación de una fuente Amazon S3 de datos
- Amazon S3 metadatos del documento
- Control de acceso a las fuentes de Amazon S3 datos
- Utilización Amazon VPC con una fuente de Amazon S3 datos

Características admitidas

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de datos de S3, realice estos cambios en su S3 y en sus AWS cuentas.

En S3, asegúrese de que:

- Copiaste el nombre de tu Amazon S3 bucket.

> **Note**
>
> El depósito debe estar en la misma región que el Amazon Kendra índice y el índice debe tener permiso para acceder al depósito que contiene los documentos.

- Ha comprobado que cada documento es único en S3 y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu AWS cuenta, asegúrate de tener:

- Creó un Amazon Kendra índice y, si utiliza la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Si no tiene un IAM rol existente, puede usar la consola para crear un nuevo IAM rol al conectar su fuente de datos de S3. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol existente y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a la fuente de datos de S3, debe proporcionar los detalles necesarios de la fuente de datos de S3 para Amazon Kendra poder acceder a los datos. Si aún no ha configurado S3 para Amazon Kendra, consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Amazon S3

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.

2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector S3 y, a continuación, elija Agregar conector. Si usa la versión 2 (si corresponde), elija el conector S3 con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información opcional:

   a. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales de su repositorio e indexar el contenido.

   > **Note**
   >
   > IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

   b. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

   c. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Para la ubicación de la fuente de datos: especifique la ruta al Amazon S3 depósito donde se almacenan los datos. Seleccione Browse S3 para elegir su bucket de S3.

   b. Para obtener el tamaño máximo de archivo: especifique un límite en MB para rastrear solo los archivos que estén por debajo de este límite. El tamaño máximo de archivo permitido Amazon Kendra es de 50 MB.

   c. Para los archivos de metadatos (opcionales), prefija la ubicación de la carpeta: especifique la ruta a la carpeta en la que se almacenan los campos o atributos y otros metadatos del documento. Seleccione Examinar S3 para localizar la carpeta de metadatos.

   d. Para la ubicación del archivo de configuración de la lista de control de acceso (opcional): especifique la ruta al archivo que contiene una estructura JSON de los usuarios y su acceso a los documentos. Seleccione Examinar S3 para localizar el archivo de la ACL.

   e. (Opcional) Seleccionar clave de descifrado: seleccione esta opción para usar una clave de descifrado. Puede optar por utilizar una AWS KMS clave existente.

   f. Para una configuración adicional (opcional): añada patrones para incluir o excluir determinados archivos. Todas las rutas se expresan con relación al bucket de S3 de ubicación del origen de datos.

   g. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

      - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
      - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

   h. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

i. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información opcional:

   a. Asignaciones de campos predeterminadas: seleccione entre las fuentes de datos predeterminadas Amazon Kendra generadas los campos que desee asignar a su índice.

   b. Agregar campo: elija esta opción para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Amazon S3

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `S3` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- BucketName: el nombre del depósito que contiene los documentos.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de

la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector S3 y Amazon Kendra. Para obtener más información, consulte Roles de IAM para orígenes de datos de S3.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados nombres, tipos y rutas de archivos. Se utilizan patrones globales (patrones que pueden expandir un patrón comodín hasta convertirse en una lista de nombres de rutas que coincidan con el patrón dado). Para ver ejemplos, consulte Uso de filtros de exclusión e inclusión en la referencia de comandos de la AWS CLI.
- Configuración de metadatos de documentos y control de acceso: agregue metadatos de documentos y archivos de control de acceso que contengan información como el URI de origen, el autor del documento o los campos o atributos del documento personalizados, así como sus usuarios y los documentos a los que pueden acceder. Cada archivo de metadatos contiene metadatos sobre un solo documento.
- Asignaciones de campos: elija asignar los campos del origen de datos de S3 a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el Esquema de plantilla de S3.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de datos de S3, consulte:

- Busque respuestas con precisión mediante el conector Amazon Kendra S3 compatible con VPC

Creación de una fuente Amazon S3 de datos

Los siguientes ejemplos muestran la creación de una fuente de Amazon S3 datos. En los ejemplos se supone que ya ha creado un índice y un IAM rol con permiso para leer los datos del índice. Para obtener más información sobre el IAM rol, consulte roles de IAM acceso. Para obtener más información acerca de cómo crear un índice, consulte Creación de un índice.

CLI

```
aws kendra create-data-source \
 --index-id index ID \
 --name example-data-source \
 --type S3 \
 --configuration '{"S3Configuration":{"BucketName":"bucket name"}}'
 --role-arn 'arn:aws:iam::account id:role:/role name
```

Python

El siguiente fragmento de código Python crea una fuente de Amazon S3 datos. Para ver el ejemplo completo, consulte Introducción (AWS SDK for Python (Boto3)).

```
print("Create an Amazon S3 data source.")

    # Provide a name for the data source
    name = "getting-started-data-source"
    # Provide an optional description for the data source
    description = "Getting started data source."
    # Provide the IAM role ARN required for data sources
    role_arn = "arn:aws:iam::${accountID}:role/${roleName}"
    # Provide the data soource connection information
    s3_bucket_name = "S3-bucket-name"
    type = "S3"
    # Configure the data source
    configuration = {"S3DataSourceConfiguration":
        {
```

```
            "BucketName": s3_bucket_name
        }
    }

    data_source_response = kendra.create_data_source(
        Configuration = configuration,
        Name = name,
        Description = description,
        RoleArn = role_arn,
        Type = type,
        IndexId = index_id
    )
```

La creación del origen de datos puede tardar algún tiempo. Puede supervisar el progreso mediante la DescribeDataSourceAPI. Cuando el estado del origen de datos es `ACTIVE`, está listo para usarse.

Los siguientes ejemplos muestran cómo obtener el estado de un origen de datos.

CLI

```
aws kendra describe-data-source \
 --index-id index ID \
 --id data source ID
```

Python

El siguiente fragmento de código Python obtiene información sobre un origen de datos de S3. Para ver el ejemplo completo, consulte Introducción (AWS SDK for Python (Boto3)).

```
print("Wait for Amazon Kendra to create the data source.")

    while True:
        data_source_description = kendra.describe_data_source(
            Id = "data-source-id",
            IndexId = "index-id"
        )
        status = data_source_description["Status"]
        print(" Creating data source. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break
```

Este origen de datos no tiene una programación, por lo que no se ejecuta automáticamente. Para indexar la fuente de datos, llame StartDataSourceSyncJobpara sincronizar el índice con la fuente de datos.

Los siguientes ejemplos muestran la sincronización de un origen de datos.

CLI

```
aws kendra start-data-source-sync-job \
 --index-id index ID \
 --id data source ID
```

Python

El siguiente fragmento de código Python sincroniza un origen de datos de Amazon S3 . Para ver el ejemplo completo, consulte Introducción (AWS SDK for Python (Boto3)).

```
print("Synchronize the data source.")

    sync_response = kendra.start_data_source_sync_job(
        Id = "data-source-id",
        IndexId = "index-id"
    )
```

Amazon S3 metadatos del documento

Puede añadir metadatos (información adicional sobre un documento) a los documentos de un bucket de Amazon S3 mediante un archivo de metadatos. Cada archivo de metadatos está asociado a un documento indexado.

Los archivos de metadatos deben almacenarse en el mismo bucket que los archivos indexados. Puede especificar una ubicación dentro del depósito para sus archivos de metadatos mediante la consola o el `S3Prefix` campo del `DocumentsMetadataConfiguration` parámetro al crear una fuente de Amazon S3 datos. Si no especifica un prefijo de Amazon S3 , los archivos de metadatos deben almacenarse en la misma ubicación que los documentos indexados.

Si especifica un Amazon S3 prefijo para los archivos de metadatos, estarán en una estructura de directorios paralela a los documentos indexados. Amazon Kendra busca sus metadatos únicamente en el directorio especificado. Si no se leen los metadatos, compruebe que la ubicación del directorio coincide con la ubicación de los metadatos.

En los siguientes ejemplos se muestra cómo la ubicación del documento indexado se asigna a la ubicación del archivo de metadatos. Tenga en cuenta que la Amazon S3 clave del documento se añade al Amazon S3 prefijo de los metadatos y, a continuación, se añade el sufijo con el sufijo `.metadata.json` para formar la ruta del archivo de metadatos. Amazon S3 La Amazon S3 clave combinada, con el Amazon S3 prefijo y el `.metadata.json` sufijo de los metadatos, no debe tener más de 1024 caracteres en total. Se recomienda mantener la Amazon S3 clave por debajo de los 1000 caracteres para tener en cuenta los caracteres adicionales al combinar la clave con el prefijo y el sufijo.

```
Bucket name:
     s3://bucketName
Document path:
     documents
Metadata path:
     none
File mapping
     s3://bucketName/documents/file.txt ->
        s3://bucketName/documents/file.txt.metadata.json
```

```
Bucket name:
     s3://bucketName
Document path:
     documents/legal
Metadata path:
     metadata
File mapping
     s3://bucketName/documents/legal/file.txt ->
        s3://bucketName/metadata/documents/legal/file.txt.metadata.json
```

Los metadatos del documento se definen en un archivo JSON. El archivo debe ser un archivo de texto UTF-8 sin un marcador BOM. El nombre del archivo JSON debe ser `<document>.<extension>.metadata.json`. En este ejemplo, “document” es el nombre del documento al que se aplican los metadatos y “extension” es la extensión de archivo del documento. El ID del documento debe ser único en `<document>.<extension>.metadata.json`.

El contenido del archivo JSON sigue esta plantilla. Todos los atributos/campos son opcionales, por lo que no es necesario incluir todos los atributos. Debe proporcionar un valor para cada atributo que desee incluir; el valor no puede estar vacío. Si no especificas el`_source_uri`, los enlaces que aparecen Amazon Kendra en los resultados de la búsqueda apuntan al compartimento que contiene

Amazon S3 el documento. `DocumentId`se asigna al campo `s3_document_id` y es la ruta absoluta al documento en S3.

```
{
    "DocumentId": "S3 document ID, the S3 path to doc",
    "Attributes": {
        "_category": "document category",
        "_created_at": "ISO 8601 encoded string",
        "_last_updated_at": "ISO 8601 encoded string",
        "_source_uri": "document URI",
        "_version": "file version",
        "_view_count": number of times document has been viewed,
        "custom attribute key": "custom attribute value",
        additional custom attributes
    },
    "AccessControlList": [
         {
             "Name": "user name",
             "Type": "GROUP | USER",
             "Access": "ALLOW | DENY"
         }
    ],
    "Title": "document title",
    "ContentType": "For example HTML | PDF. For supported content types, see Types of
 documents."
}
```

Los campos de metadatos `_created_at` y `_last_updated_at` son fechas codificadas según la norma ISO 8601. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en la zona horaria de Europa Central.

Puede añadir información adicional al campo `Attributes` sobre un documento que utilice para filtrar consultas o agrupar las respuestas a las consultas. Para obtener más información, consulte Creación de campos de documento personalizados.

Puede utilizar el campo `AccessControlList` para filtrar la respuesta de una consulta. De esta forma, solo determinados usuarios y grupos tienen acceso a los documentos. Para obtener más información, consulte Filtrar por contexto de usuario.

Control de acceso a las fuentes de Amazon S3 datos

Puede controlar el acceso a los documentos de una fuente de Amazon S3 datos mediante un archivo de configuración. El archivo se especifica en la consola o como `AccessControlListConfiguration` parámetro al llamar a la UpdateDataSourceAPI CreateDataSourceo.

El archivo de configuración contiene una estructura JSON que identifica un prefijo S3 y enumera la configuración de acceso del prefijo. El prefijo puede ser una ruta o un archivo individual. Si el prefijo es una ruta, la configuración de acceso se aplica a todos los archivos de esa ruta. Hay un número máximo de prefijos S3 en el archivo de configuración JSON y un tamaño de archivo máximo predeterminado. Para obtener más información, consulte Cuotas para Amazon Kendra.

En la configuración de acceso, se pueden especificar tanto los usuarios como los grupos. Cuando se consulta el índice, se especifica la información del usuario y del grupo. Para obtener más información, consulte Filtrado por atributo de usuario.

La estructura JSON del archivo de configuración debe tener el siguiente formato:

```
[
    {
        "keyPrefix": "s3://BUCKETNAME/prefix1/",
        "aclEntries": [
            {
                "Name": "user1",
                "Type": "USER",
                "Access": "ALLOW"
            },
            {
                "Name": "group1",
                "Type": "GROUP",
                "Access": "DENY"
            }
        ]
    },
    {
        "keyPrefix": "s3://prefix2",
        "aclEntries": [
            {
                "Name": "user2",
                "Type": "USER",
                "Access": "ALLOW"
```

```
            },
            {
                "Name": "user1",
                "Type": "USER",
                "Access": "DENY"
            },
            {
                "Name": "group1",
                "Type": "GROUP",
                "Access": "DENY"
            }
        ]
    }
]
```

Utilización Amazon VPC con una fuente de Amazon S3 datos

En este tema se proporciona un step-by-step ejemplo que muestra cómo conectarse a un bucket de Amazon S3 mediante un conector de Amazon S3 a través de Amazon VPC. En el ejemplo se supone que parte de un bucket de S3 existente. Le recomendamos que cargue solo algunos documentos en su bucket de S3 para probar el ejemplo.

Puedes conectarte Amazon Kendra a tu Amazon S3 bucket a través Amazon VPC de. Para ello, debe especificar la Amazon VPC subred y los grupos de Amazon VPC seguridad al crear el conector de la fuente de Amazon S3 datos.

> **Important**
>
> Para que un Amazon Kendra Amazon S3 conector pueda acceder a su Amazon S3 depósito, asegúrese de haber asignado un Amazon S3 punto final a su nube privada virtual (VPC).

Amazon Kendra Para sincronizar los documentos de su Amazon S3 depósito Amazon VPC, debe completar los siguientes pasos:

- Configura un Amazon S3 punto final para Amazon VPC. Para obtener más información sobre cómo configurar un Amazon S3 punto final, consulte los puntos finales de puerta de enlace Amazon S3 en la AWS PrivateLink guía.
- (Opcional) Comprobó las políticas de su Amazon S3 bucket para asegurarse de que se pueda acceder al Amazon S3 bucket desde la nube privada virtual (VPC) a la que lo asignó. Amazon

Kendra Para obtener más información, consulte Controlar el acceso desde los puntos de enlace de la VPC con políticas de bucket en la Guía del usuario de Amazon S3

Pasos

- Paso 1: configurar un Amazon VPC
- (Opcional) Paso 2: configurar la política Amazon S3 de bucket
- Paso 3: Cree un conector de fuente Amazon S3 de datos de prueba

Paso 1: configurar un Amazon VPC

Cree una red de VPC que incluya una subred privada con un punto final de Amazon S3 puerta de enlace y un grupo de seguridad Amazon Kendra para utilizarla más adelante.

Para configurar una VPC con una subred privada, un punto final S3 y un grupo de seguridad

1. Inicie sesión en AWS Management Console y abra la Amazon VPC consola en. https://console.aws.amazon.com/vpc/
2. Cree una VPC con una subred privada y un punto de conexión S3 para Amazon Kendra usar:

   En el panel de navegación, elija Sus VPC y, a continuación, elija Crear VPC.

   a. En Recursos para crear, elija VPC y más.
   b. En Etiqueta de nombre, habilite Generación automática y, a continuación, introduzca **`kendra-s3-example`**.
   c. En Bloque de CIDR IPv4/IPv6, deje los valores predeterminados.
   d. Para Número de zonas de disponibilidad (AZ), elija el número 1.
   e. Seleccione Personalizar las AZ y, a continuación, seleccione una zona de disponibilidad de la lista Primera zona de disponibilidad.

      Amazon Kendra solo admite un conjunto específico de zonas de disponibilidad.
   f. Para Número de subredes públicas, elija el número 0.
   g. Para Número de subredes privadas, elija el número 1.
   h. Para NAT gateways (puertas de enlace NAT), elija None (Ninguna).
   i. Para los puntos de conexión de VPC, elija Puerta de enlace de Amazon S3 .
   j. Deje el resto de los ajustes con sus valores predeterminados.

k. Seleccione Create VPC (Crear VPC).

Espere a que finalice el flujo de trabajo de Crear VPC. A continuación, elija Ver VPC para comprobar la VPC que acaba de crear.

Ahora ha creado una red de VPC con una subred privada que no tiene acceso a la internet pública.

3. Copie el ID de punto de conexión de VPC de su punto de conexión de Amazon S3:

a. En el panel de navegación, elija Puntos de conexión.

b. En la lista Puntos de conexión, busque el punto de conexión de Amazon S3 `kendra-s3-example-vpce-s3` que acaba de crear junto con su VPC.

c. Anote el ID del punto de conexión de VPC.

Ya ha creado un punto de conexión de puerta de enlace de Amazon S3 para acceder a su bucket de Amazon S3 a través de una subred.

4. Cree un grupo de seguridad Amazon Kendra para usar:

a. En el panel de navegación, elija Grupos de seguridad y, a continuación, elija Crear un grupo de seguridad.

b. En Nombre del grupo de seguridad, introduzca **`s3-data-source-security-group`**.

c. Elija su VPC en la lista de Amazon VPC.

d. Deje las reglas de entrada y las reglas de salida como predeterminadas.

e. Elija Crear grupo de seguridad.

Ya ha creado un grupo de seguridad de VPC.

Usted asigna la subred y el grupo de seguridad que creó a su conector de fuente de datos de Amazon Kendra Amazon S3 durante el proceso de configuración del conector.

(Opcional) Paso 2: configurar la política Amazon S3 de bucket

En este paso opcional, aprenda a configurar una política de bucket de Amazon S3 para que solo se pueda acceder a su bucket de Amazon S3 desde la VPC a la que lo asigne. Amazon Kendra

Amazon Kendra utiliza funciones de IAM para acceder a su bucket de Amazon S3 y no requiere que configure una política de bucket de Amazon S3. Sin embargo, puede resultarle útil crear una política de bucket si quiere configurar un Amazon S3 conector mediante un bucket de Amazon S3 que tenga políticas existentes que restrinjan el acceso a él desde la Internet pública.

Para configurar su política Amazon S3 de bucket

1. Abra la consola de Amazon S3 en https://console.aws.amazon.com/s3/.
2. En el panel de navegación, elija Buckets.
3. Elige el nombre del bucket de Amazon S3 con el que quieres sincronizarte Amazon Kendra.
4. Seleccione la pestaña Permisos, desplácese hacia abajo hasta Política de buckets y, a continuación, haga clic en Editar.
5. Agregue o modifique su política de buckets para permitir el acceso solo desde el punto de conexión de VPC que creó.

   A continuación se muestra un ejemplo de política de bucket. Sustituya *bucket-name* y *vpce-id* por el nombre de su bucket de Amazon S3 y el ID del punto de conexión de Amazon S3 que indicó anteriormente.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::bucket-name/*",
      "Condition": {
        "StringNotEquals": {
          "aws:SourceVpce": "vpce-id"
        }
      }
    }
  ]
}
```

6. Seleccione Guardar cambios.

Ahora solo se puede acceder a su bucket de S3 desde la VPC específica que creó.

Paso 3: Cree un conector de fuente Amazon S3 de datos de prueba

Para probar la Amazon VPC configuración, cree un Amazon S3 conector. A continuación, configúrelo con la VPC que creó siguiendo los pasos descritos en Amazon S3.

Para los valores de Amazon VPC configuración, elija los valores que creó en este ejemplo:

- Amazon VPC(VPC): `kendra-s3-example-vpc`
- Subredes: `kendra-s3-example-subnet-private1-[availability zone]`
- Grupos de seguridad: `s3-data-source-security-group`

Espere a que termine de crearse el conector. Una vez creado el Amazon S3 conector, elija Sincronizar ahora para iniciar una sincronización.

La sincronización puede tardar entre varios minutos y varias horas en finalizar, según el número de documentos que haya en el Amazon S3 depósito. Para probar el ejemplo, le recomendamos que cargue solo algunos documentos en su bucket de S3. Si la configuración es correcta, en algún momento debería aparecer el Estado de sincronización como Completado.

Si encuentras algún error, consulta Solución de problemas de Amazon VPC conexión.

Amazon Kendra Rastreador web

Puede usar Amazon Kendra Web Crawler para rastrear e indexar páginas web.

Solo puede rastrear sitios web de cara al público o sitios web internos de la empresa que utilicen el protocolo de comunicación segura Hypertext Transfer Protocol Secure (HTTPS). Si recibe un error al rastrear un sitio web, es posible que el sitio web esté bloqueado para que no pueda rastrearse. Para rastrear sitios web internos, puede configurar un proxy web. El proxy web debe estar orientado al público. También puede utilizar la autenticación para acceder a sitios web y rastrearlos.

Al seleccionar los sitios web que se van a indexar, se debe respetar la Política de uso aceptable de Amazon y todas las demás condiciones de Amazon. Recuerde que solo debe usar Amazon Kendra Web Crawler para indexar sus propias páginas web o páginas web para las que tenga autorización para indexar. Para obtener información sobre cómo impedir que Amazon Kendra Web Crawler indexe sus sitios web, consulte. Configuración del archivo *`robots.txt`* para el rastreador web de Amazon Kendra

> **Note**
>
> El uso indebido de Amazon Kendra Web Crawler para rastrear agresivamente sitios web o páginas web que no son de su propiedad no se considera un uso aceptable.

Amazon Kendra tiene dos versiones del conector. web crawler Las características compatibles de cada versión incluyen:

Amazon Kendra Conector Web Crawler v1.0/API WebCrawlerConfiguration

- Proxy de web
- Filtros de inclusión/exclusión

Amazon Kendra Conector Web Crawler v2.0/API TemplateConfiguration

- Asignaciones de campo
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Proxy de web
- Autenticación básica, NTLM/Kerberos, SAML y mediante formularios para sus sitios web
- Nube privada virtual (VPC)

> **Important**
>
> La creación de conectores Web Crawler v2.0 no es compatible con. AWS CloudFormation Utilice el conector Web Crawler v1.0 si necesita asistencia. AWS CloudFormation

Para solucionar problemas del conector de fuente de datos de su rastreador Amazon Kendra web, consulte. Solución de problemas con los orígenes de datos

Temas

- Amazon Kendra Conector Web Crawler v1.0
- Amazon Kendra Conector Web Crawler v2.0

- Configuración del archivo robots.txt para el rastreador web de Amazon Kendra

Amazon Kendra Conector Web Crawler v1.0

Puede utilizar Amazon Kendra Web Crawler para rastrear e indexar páginas web.

Solo puede rastrear sitios web de cara al público y sitios web que utilicen el protocolo de comunicación segura Hypertext Transfer Protocol Secure (HTTPS). Si recibe un error al rastrear un sitio web, es posible que el sitio web esté bloqueado para que no pueda rastrearse. Para rastrear sitios web internos, puede configurar un proxy web. El proxy web debe estar orientado al público.

Al seleccionar los sitios web que se van a indexar, se debe respetar la Política de uso aceptable de Amazon y todas las demás condiciones de Amazon. Recuerde que solo debe usar Amazon Kendra Web Crawler para indexar sus propias páginas web o páginas web para las que tenga autorización para indexar. Para obtener información sobre cómo impedir que Amazon Kendra Web Crawler indexe sus sitios web, consulte. Configuración del archivo *robots.txt* para el rastreador web de Amazon Kendra

> **Note**
>
> El uso indebido de Amazon Kendra Web Crawler para rastrear agresivamente sitios web o páginas web que no son de su propiedad no se considera un uso aceptable.

Para solucionar problemas del conector de fuente de datos del rastreador Amazon Kendra web, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

- Proxy de web
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder usarlo Amazon Kendra para indexar sus sitios web, compruebe los detalles de sus sitios web y AWS cuentas.

Para sus sitios web, asegúrese de que:

- Copie las URL semilla o de mapa del sitio de los sitios web que desea indexar.
- Para los sitios web que requieren una autenticación básica: Apuntó el nombre de usuario y la contraseña y copió el nombre de host del sitio web y el número de puerto.
- Opcional: copió el nombre de host del sitio web y el número de puerto si quiere usar un proxy web para conectarse a los sitios web internos que desea rastrear. El proxy web debe estar orientado al público. Amazon Kendra admite la conexión a servidores proxy web respaldados por una autenticación básica o puede conectarse sin autenticación.
- Compruebe que cada documento de página web que desea indexar es único y que se encuentra entre otros orígenes de datos que piensa utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu AWS cuenta, asegúrate de tener:

- Creó un Amazon Kendra índice y, si utiliza la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- En el caso de los sitios web que requieren autenticación, o si utilizan un proxy web con autenticación, guardan las credenciales de autenticación en AWS Secrets Manager secreto y, si utilizan la API, anotan el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de web crawler datos. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de web crawler datos, debe proporcionar los detalles necesarios de la fuente de web crawler datos para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configuradoweb crawler, Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a web crawler

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector Web Crawler y, a continuación, selecciona Añadir conector. Si utilizas la versión 2 (si corresponde), elige el conector para rastreadores web con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. Para Origen, elija entre URL de origen y Mapas de sitio de origen en función de su caso de uso e introduzca los valores para cada uno.

Puede tener hasta 10 URL de origen y tres mapas del sitio.

> **Note**
>
> Si desea rastrear un mapa del sitio, compruebe que la URL base o raíz coincide con las URL que figuran en la página de su mapa del sitio. Por ejemplo, si la URL de su mapa del sitio es https://example.com/sitemap-page.html, las URL enumeradas en esta página del mapa del sitio también deberían utilizar la URL base “https://example.com/”.

b. (Opcional) Para el Proxy web, introduzca la siguiente información:

i. Nombre de host: el nombre de host donde se requiere el proxy web.

ii. Número de puerto: puerto utilizado por el protocolo de transporte de URL del host. El número de puerto debe ser un valor numérico entre 0 y 65535.

iii. Para las credenciales del proxy web: si su conexión de proxy web requiere autenticación, elija un secreto existente o cree uno nuevo para almacenar sus credenciales de autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

iv. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager Secrets Manager :

A. Nombre del secreto: un nombre para su secreto. El prefijo “AmazonKendra-WebCrawler-” se añade automáticamente al nombre del secreto.

B. Para el nombre de usuario y la contraseña: introduzca estas credenciales de autenticación básicas para sus sitios web.

C. Seleccione Guardar.

c. (Opcional) Hosts con autenticación: seleccione esta opción para agregar hosts adicionales con autenticación.

d. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales de su repositorio e indexar el contenido.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

e. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. Rango de rastreo: elige el tipo de páginas web que desea rastrear.

b. Profundidad de rastreo: seleccione el número de niveles de la URL inicial que Amazon Kendra se deben rastrear.

c. En Configuración avanzada de rastreo y Configuración adicional, introduzca la siguiente información:

i. Tamaño máximo de archivo: tamaño máximo de página web o archivo adjunto que se deben rastrear. Mínimo 0,000001 MB (1 byte). Máximo de 50 MB.

ii. Número máximo de enlaces por página: número máximo de enlaces rastreados por página. Los enlaces se rastrean en orden de aparición. Mínimo 1 enlace/página. Máximo 1000 enlaces/página.

iii. Limitación máxima: el número de direcciones URL rastreadas por nombre de host por minuto. Mínimo 1 URL por nombre de host por minuto. Máximo 300 URL/nombre de host/minuto.

iv. Patrones regex: añada patrones de expresiones regulares para incluir o excluir determinadas URL. Puede agregar hasta 100 patrones.

d. En Sincronizar el programa de ejecución, en Frecuencia: elija la frecuencia con la que Amazon Kendra se sincronizará con la fuente de datos.

e. Elija Siguiente.

8. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra web crawler

Debe especificar lo siguiente mediante la WebCrawlerConfigurationAPI:

- URL: especifique las URL semilla o de punto de partida de los sitios web o las URL de mapa del sitio de los sitios web que desea rastrear utilizando SeedUrlConfiguration y SiteMapsConfiguration.

  > **Note**
  >
  > Si desea rastrear un mapa del sitio, compruebe que la URL base o raíz coincide con las URL que figuran en la página de su mapa del sitio. Por ejemplo, si la URL de su mapa del sitio es https://example.com/sitemap-page.html, las URL enumeradas en esta página del mapa del sitio también deberían utilizar la URL base “https://example.com/”.

- Nombre de recurso de Amazon (ARN) secreto: si un sitio web requiere autenticación básica, usted proporciona el nombre del host, el número de puerto y un secreto que almacena sus credenciales de autenticación básica de su nombre de usuario y contraseña. El ARN secreto se proporciona mediante la API AuthenticationConfiguration. El secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  También puede proporcionar credenciales de proxy web mediante un secreto de AWS Secrets Manager . Utilice la API ProxyConfiguration para proporcionar el nombre de host y el número de puerto del sitio web y, opcionalmente, el secreto que almacena sus credenciales de proxy web.

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector del rastreador web y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos del rastreador web.

También puede añadir las siguientes características opcionales:

- Modo de rastreo: elija si desea rastrear solo los nombres de host de los sitios web o los nombres de host con subdominios, o también rastrear otros dominios a los que enlazan las páginas web.
- La “profundidad” o número de niveles desde el nivel semilla hasta el nivel rastreo. Por ejemplo, la página URL semilla tiene la profundidad 1 y todos los hipervínculos de esta página que también se rastreen tienen la profundidad 2.
- El número máximo de URL de una misma página web que se rastrearán.
- El tamaño máximo en MB de una página web a rastrear.
- El número de direcciones URL rastreadas por host de sitio web por minuto.
- El host del proxy web y el número de puerto para conectarse a sitios web internos y rastrearlos. Por ejemplo, el nombre de host de https://a.example.com/page1.html es “a.example.com“ y el número de puerto es 443, el puerto estándar para HTTPS. Si se requieren credenciales de proxy web para conectarse a un host de sitio web, puede crear un AWS Secrets Manager que almacene las credenciales.
- La información de autenticación para acceder y rastrear sitios web que requieren la autenticación del usuario.
- Puede extraer las metaetiquetas HTML como campos con la herramienta de enriquecimiento de documentos personalizados. Para más información, consulte Personalización de los metadatos del documento durante el proceso de ingesta. Para ver un ejemplo de cómo extraer metaetiquetas HTML, consulte los ejemplos de CDE.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinadas URL.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un

filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de web crawler datos, consulte:

- Reimagine el descubrimiento de conocimientos con Web Amazon Kendra Crawler

Amazon Kendra Conector Web Crawler v2.0

Puede utilizar Amazon Kendra Web Crawler para rastrear e indexar páginas web.

Solo puede rastrear sitios web de cara al público o sitios web internos de la empresa que utilicen el protocolo de comunicación segura Hypertext Transfer Protocol Secure (HTTPS). Si recibe un error al rastrear un sitio web, es posible que el sitio web esté bloqueado para que no pueda rastrearse. Para rastrear sitios web internos, puede configurar un proxy web. El proxy web debe estar orientado al público. También puede utilizar la autenticación para acceder a sitios web y rastrearlos.

Amazon Kendra Web Crawler v2.0 utiliza el paquete de rastreadores web Selenium y un controlador Chromium. Amazon Kendra actualiza automáticamente la versión de Selenium y el controlador Chromium mediante la integración continua (CI).

Al seleccionar los sitios web que se van a indexar, se debe respetar la Política de uso aceptable de Amazon y todas las demás condiciones de Amazon. Recuerde que solo debe usar Amazon Kendra Web Crawler para indexar sus propias páginas web o páginas web para las que tenga autorización para indexar. Para obtener información sobre cómo impedir que Amazon Kendra Web Crawler indexe sus sitios web, consulte. Configuración del archivo *robots.txt* para el rastreador web de Amazon Kendra . El uso indebido de Amazon Kendra Web Crawler para rastrear agresivamente sitios web o páginas web que no son de su propiedad no se considera un uso aceptable.

Para solucionar problemas del conector de fuente de datos del rastreador Amazon Kendra web, consulte. Solución de problemas con los orígenes de datos

> **Note**
>
> El conector Web Crawler v2.0 no admite el rastreo de listas de sitios web desde depósitos cifrados. AWS KMS Amazon S3 Solo admite el cifrado del lado del servidor con claves administradas. Amazon S3

> **Important**
>
> La creación de conectores Web Crawler v2.0 no es compatible con. AWS CloudFormation Utilice el conector Web Crawler v1.0 si necesita asistencia. AWS CloudFormation

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión

Características admitidas

- Asignaciones de campo
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Proxy de web
- Autenticación básica, NTLM/Kerberos, SAML y mediante formularios para sus sitios web
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarlos Amazon Kendra para indexar tus sitios web, comprueba los detalles de tus sitios web y AWS cuentas.

Para sus sitios web, asegúrese de que:

- Copie las URL semilla o de mapa del sitio de los sitios web que desea indexar. Puede almacenar las URL en un archivo de texto y subirlo a un bucket Amazon S3 . Cada URL del archivo de texto

debe estar formateada en una línea independiente. Si quieres almacenar tus mapas de sitio en un Amazon S3 depósito, asegúrate de haber copiado el XML del mapa del sitio y de haberlo guardado en un archivo XML. También puede agrupar varios archivos XML de mapa del sitio en un archivo ZIP.

> **Note**
>
> (local o en el servidor) Amazon Kendra comprueba si la información de punto final incluida AWS Secrets Manager es la misma que la información de punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a evitar el problema del suplente confuso, que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, pero utiliza Amazon Kendra como proxy para acceder al secreto configurado y realizar la acción. Si más adelante cambia la información de punto de conexión, debe crear un nuevo secreto para sincronizar esta información.

- Para los sitios web que requieren autenticación básica, NTLM o Kerberos:
  - Anote las credenciales de autenticación de su sitio web, que incluyen un nombre de usuario y una contraseña.

    > **Note**
    >
    > Amazon Kendra Web Crawler v2.0 admite el protocolo de autenticación NTLM, que incluye el cifrado de contraseñas, y el protocolo de autenticación Kerberos, que incluye el cifrado de contraseñas.

- Para los sitios web que requieren autenticación mediante SAML o mediante formulario de inicio de sesión:
  - Anote las credenciales de autenticación de su sitio web, que incluyen un nombre de usuario y una contraseña.
  - Se copiaron los XPaths (lenguaje de rutas XML) del campo de nombre de usuario (y el botón de nombre de usuario si se utiliza SAML), el campo y el botón de contraseña, y se copió la URL de la página de inicio de sesión. Puede encontrar los XPaths de los elementos utilizando las herramientas para desarrolladores de su navegador web. Los XPaths suelen seguir este formato: `//tagname[@Attribute='Value']`.

> **Note**
>
> Amazon Kendra Web Crawler v2.0 utiliza un navegador Chrome sin interfaz y la información del formulario para autenticar y autorizar el acceso con una URL protegida por OAuth 2.0.

- Opcional: copie el nombre del host y el número de puerto del servidor proxy web si desea utilizar un proxy web para conectarse a los sitios web internos que desea rastrear. El proxy web debe estar orientado al público. Amazon Kendra admite la conexión a servidores proxy web respaldados por una autenticación básica o puede conectarse sin autenticación.
- Opcional: copió el ID de subred de la nube privada virtual (VPC) si quiere usar una VPC para conectarse a los sitios web internos que desea rastrear. Para obtener más información, consulte Configuración de un Amazon VPC.
- Compruebe que cada documento de página web que desea indexar es único y que se encuentra entre otros orígenes de datos que piensa utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En su AWS cuenta, asegúrese de tener:

- Creó un Amazon Kendra índice y, si utiliza la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el nombre del recurso de Amazon del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- En el caso de los sitios web que requieren autenticación, o si utilizan un proxy web con autenticación, guardan las credenciales de autenticación en AWS Secrets Manager secreto y, si utilizan la API, anotan el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de web crawler datos. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de web crawler datos, debe proporcionar los detalles necesarios de la fuente de web crawler datos para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configuradoweb crawler, Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a web crawler

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector Web Crawler y, a continuación, selecciona Añadir conector. Si utilizas la versión 2 (si corresponde), elige el conector para rastreadores web con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. Origen: Elija entre URL de origen, mapas de sitio, archivo de URL de origen, archivo de mapas de sitio de origen. Si opta por utilizar un archivo de texto que incluya una lista de hasta 100 direcciones URL iniciales, debe especificar la ruta al Amazon S3 depósito en el que está almacenado el archivo. Si opta por utilizar un archivo XML de mapa del sitio, debe especificar la ruta al bucket Amazon S3 en el que está almacenado el archivo. También puede agrupar varios archivos XML de mapa del sitio en un archivo ZIP. De lo contrario, puede introducir manualmente hasta 10 URL semilla o punto de partida y hasta tres URL de mapa del sitio.

> **Note**
>
> Si desea rastrear un mapa del sitio, compruebe que la URL base o raíz coincide con las URL que figuran en la página de su mapa del sitio. Por ejemplo, si la URL de su mapa del sitio es https://example.com/sitemap-page.html, las URL enumeradas en esta página del mapa del sitio también deberían utilizar la URL base “https://example.com/”.

Si sus sitios web requieren autenticación para acceder a ellos, puede elegir entre autenticación básica, NTLM/Kerberos, SAML o de formulario. En caso contrario, elija la opción de no autenticación.

> **Note**
>
> Si más adelante desea editar su origen de datos para cambiar las direcciones URL con autenticación a mapas del sitio, deberá crear un nuevo origen de datos. Amazon Kendra configura el origen de datos utilizando la información del punto de conexión de las URL semilla en el secreto Secrets Manager para la autenticación y, por lo tanto, no puede volver a configurar el origen de datos al cambiar a mapas de sitio.

- AWS Secrets Manager secreto: si sus sitios web requieren la misma autenticación para acceder a los sitios web, elija un secreto existente o cree uno nuevo Secrets Manager para almacenar las credenciales del sitio web. Si decides crear un secreto nuevo, se abrirá una ventana AWS Secrets Manager secreta.

  Si eligió la autenticación Básica o NTLM/Kerberos, introduzca un nombre para el secreto, además del nombre de usuario y la contraseña. El protocolo de autenticación NTLM incluye el hash de contraseñas y el protocolo de autenticación de Kerberos incluye el cifrado de contraseñas.

  Si eligió la autenticación SAML o Formularios, introduzca un nombre para el secreto, además del nombre de usuario y la contraseña. Utilice XPath para el campo del nombre de usuario (y XPath para el botón del nombre de usuario si utiliza SAML). Utilice XPaths para el campo y el botón de contraseña y para la URL de la página de inicio de sesión. Puede encontrar los XPaths (lenguaje de rutas XML) de los elementos utilizando las herramientas para desarrolladores de su navegador web. Los XPaths suelen seguir este formato: `//tagname[@Attribute='Value']`.

b. Proxy web (opcional): introduzca el nombre de host y el número de puerto del servidor proxy que desee utilizar para conectarse a sitios web internos. Por ejemplo, el nombre de host de https://a.example.com/page1.html es “a.example.com“ y el número de puerto es 443, el puerto estándar para HTTPS. Si se requieren credenciales de proxy web para conectarse a un servidor de sitios web, puede crear una AWS Secrets Manager que almacene las credenciales.

c. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

d. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales de su repositorio e indexar el contenido.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

e. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Ámbito de sincronización: establece límites para el rastreo de páginas web, incluidos sus dominios, tamaños de archivo y enlaces, y filtra las URL mediante patrones de expresiones regulares.

      i. (Opcional) Rango de dominios de rastreo: elija si desea rastrear solo los dominios del sitio web, los dominios con subdominios o rastrear también otros dominios a los que enlazan las páginas web. De forma predeterminada, Amazon Kendra solo rastrea los dominios de los sitios web que desee rastrear.

      ii. (Opcional) Configuración adicional: configure los siguientes ajustes:

      - Profundidad de rastreo: la “profundidad” o el número de niveles desde el nivel inicial hasta el de rastreo. Por ejemplo, la página URL semilla tiene la profundidad 1 y todos los hipervínculos de esta página que también se rastreen tienen la profundidad 2.
      - Tamaño máximo de archivo: tamaño máximo en MB de una página web o archivo adjunto que se deben rastrear.
      - Máximo de enlaces por página: el número de direcciones URL de una sola página web para rastrear.
      - Limitación máxima de la velocidad de rastreo: el número de direcciones URL rastreadas por host de sitio web por minuto.
      - Archivos: elija rastrear los archivos a los que enlazan las páginas web.
      - Rastrear e indexar direcciones URL: añada patrones de expresiones regulares para incluir o excluir el rastreo de determinadas direcciones URL y la indexación de cualquier hipervínculo de estas páginas web con direcciones URL.

b. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

   - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
   - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

c. Programa de ejecución de sincronización: en Frecuencia, elija la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

d. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Seleccione entre los campos predeterminados Amazon Kendra generados por las páginas web y los archivos que desee asignar a su índice.

   b. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a web crawler

Debe especificar un JSON del esquema del origen de datos mediante la API TemplateConfiguration. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `WEBCRAWLERV2` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.

- URL: especifique las URL semilla o de punto de partida de los sitios web o las URL de mapa del sitio de los sitios web que desea rastrear. Puedes especificar la ruta a un Amazon S3 bucket que almacene tu lista de URL iniciales. Cada URL en el archivo de texto para las URL de semillas debe formatearse en una línea separada. También puedes especificar la ruta a un Amazon S3 depósito que almacene los archivos XML de tu mapa del sitio. Puede agrupar varios archivos de mapa del sitio en un archivo ZIP y almacenar el archivo ZIP en su bucket de Amazon S3 .

  > **Note**
  >
  > Si desea rastrear un mapa del sitio, compruebe que la URL base o raíz coincide con las URL que figuran en la página de su mapa del sitio. Por ejemplo, si la URL de su mapa del sitio es https://example.com/sitemap-page.html, las URL enumeradas en esta página del mapa del sitio también deberían utilizar la URL base “https://example.com/”.

- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Autenticación: si sus sitios web requieren la misma autenticación, especifique autenticación `BasicAuth`, `NTLM_Kerberos`, `SAML` o `Form`. Si sus sitios web no requieren autenticación, especifique `NoAuthentication`.
- Nombre de recurso de Amazon (ARN) secreto: si sus sitios web requieren autenticación básica, NTLM o Kerberos, debe proporcionar un secreto que almacene las credenciales de autenticación de su nombre de usuario y contraseña. Debe proporcionar el nombre de recurso de Amazon (ARN) de un secreto de AWS Secrets Manager . El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "seedUrlsHash": "Hash representation of all seed URLs",
    "userName": "user name",
    "password": "password"
}
```

Si sus sitios web requieren autenticación SAML, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "seedUrlsHash": "Hash representation of all seed URLs",

    "userName": "user name",
    "password": "password",
    "userNameFieldXpath": "XPath for user name field",
    "userNameButtonXpath": "XPath for user name button",
    "passwordFieldXpath": "XPath for password field",
    "passwordButtonXpath": "XPath for password button",
    "loginPageUrl": "Full URL for website login page"
}
```

Si sus sitios web requieren autenticación de formularios, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "seedUrlsHash": "Hash representation of all seed URLs",
    "userName": "user name",
    "password": "password",
    "userNameFieldXpath": "XPath for user name field",
    "passwordFieldXpath": "XPath for password field",
    "passwordButtonXpath": "XPath for password button",
    "loginPageUrl": "Full URL for website login page"
}
```

Puede encontrar los XPaths (lenguaje de rutas XML) de los elementos utilizando las herramientas para desarrolladores de su navegador web. Los XPaths suelen seguir este formato: `//tagname[@Attribute='Value']`.

También puede proporcionar credenciales de proxy web mediante un secreto de AWS Secrets Manager .

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector del rastreador web y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos del rastreador web.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Rango de dominios: elija si desea rastrear solo los dominios web con subdominios o rastrear también otros dominios a los que enlazan las páginas web. De forma predeterminada, Amazon Kendra solo rastrea los dominios de los sitios web que deseas rastrear.
- La “profundidad” o número de niveles desde el nivel semilla hasta el nivel rastreo. Por ejemplo, la página URL semilla tiene la profundidad 1 y todos los hipervínculos de esta página que también se rastreen tienen la profundidad 2.
- El número máximo de URL de una misma página web que se rastrearán.
- Tamaño máximo (en MB) de una página web o un archivo adjunto que se van a rastrear.
- El número de direcciones URL rastreadas por host de sitio web por minuto.
- El host del proxy web y el número de puerto para conectarse a sitios web internos y rastrearlos. Por ejemplo, el nombre de host de https://a.example.com/page1.html es “a.example.com“ y el número de puerto es 443, el puerto estándar para HTTPS. Si se requieren credenciales de proxy web para conectarse a un host de sitio web, puede crear un AWS Secrets Manager que almacene las credenciales.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir el rastreo de determinadas URL y la indexación de los hipervínculos de estas páginas web con URL.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Asignaciones de campos: elija asignar los campos de las páginas web y los archivos de las páginas web a sus campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el Esquema de plantilla de Web Crawler de Amazon Kendra.

Configuración del archivo **`robots.txt`** para el rastreador web de Amazon Kendra

Amazon Kendra es un servicio de búsqueda inteligente que AWS los clientes utilizan para indexar y buscar los documentos que elijan. Para indexar documentos en la web, los clientes pueden utilizar un rastreador Amazon Kendra web, que indica qué URL deben indexarse y otros parámetros operativos. Amazon Kendra los clientes deben obtener una autorización antes de indexar cualquier sitio web en particular.

Amazon Kendra Web Crawler respeta las directivas estándar de robots.txt, como `Allow` y. `Disallow` Puede modificar el `robots.txt` archivo de su sitio web para controlar la forma en que Amazon Kendra Web Crawler lo rastrea.

Configurar el modo en que Amazon Kendra Web Crawler accede a su sitio web

Puede controlar la forma en que el Amazon Kendra Web Crawler indexa su sitio web mediante directivas y directivas. `Allow` `Disallow` También puede controlar qué páginas web se indexan y qué páginas web no se rastrean.

Para permitir que Amazon Kendra Web Crawler rastree todas las páginas web excepto las no permitidas, utilice la siguiente directiva:

```
User-agent: amazon-kendra    # Amazon Kendra Web Crawler
Disallow: /credential-pages/ # disallow access to specific pages
```

Para permitir que Amazon Kendra Web Crawler rastree solo páginas web específicas, utilice la siguiente directiva:

```
User-agent: amazon-kendra    # Amazon Kendra Web Crawler
Allow: /pages/ # allow access to specific pages
```

Para permitir que Amazon Kendra Web Crawler rastree todo el contenido del sitio web e impedir que otros robots rastreen, utilice la siguiente directiva:

```
User-agent: amazon-kendra # Amazon Kendra Web Crawler
Allow: / # allow access to all pages
User-agent: * # any (other) robot
Disallow: / # disallow access to any pages
```

Impedir que Amazon Kendra Web Crawler rastree tu sitio web

Puede impedir que Amazon Kendra Web Crawler indexe su sitio web mediante esta directiva. `Disallow` También puede controlar qué páginas web se rastrean y cuáles no.

Para evitar que Amazon Kendra Web Crawler rastree el sitio web, utilice la siguiente directiva:

```
User-agent: amazon-kendra # Amazon Kendra Web Crawler
Disallow: / # disallow access to any pages
```

Amazon Kendra Web Crawler también admite los robots `noindex` y `nofollow` las directivas de las metaetiquetas de las páginas HTML. Estas directivas impiden que el rastreador web indexe una página web y deje de seguir los enlaces de la página web. Las metaetiquetas se colocan en la sección del documento para especificar las reglas de los robots.

Por ejemplo, la siguiente página web incluye las directivas robots `noindex` y `nofollow`:

```
<html>
<head>
    <meta name="robots" content="noindex, nofollow"/>
    ...
</head>
<body>...</body>
</html>
```

Si tiene alguna pregunta o duda sobre Amazon Kendra Web Crawler, puede ponerse en contacto con el equipo de AWS soporte.

Amazon WorkDocs

Amazon WorkDocs es un servicio de colaboración de contenido seguro para crear, editar, almacenar y compartir contenido. Puede usarlo Amazon Kendra para indexar su fuente Amazon WorkDocs de datos.

Puede conectarse Amazon Kendra a su fuente Amazon WorkDocs de datos mediante la Amazon Kendra consola y la WorkDocsConfigurationAPI.

Amazon WorkDocs está disponible en las regiones de Oregón, Virginia del Norte, Sídney, Singapur e Irlanda.

Para solucionar problemas del conector de la fuente de Amazon Kendra WorkDocs datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra WorkDocs el conector de fuente de datos admite las siguientes funciones:

- Asignaciones de campo
- control de acceso de usuarios
- Filtros de inclusión/exclusión
- Registro de cambios

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de WorkDocs datos, realice estos cambios en sus AWS cuentas WorkDocs y.

En WorkDocs, asegúrate de tener:

- Apuntó el ID de Amazon WorkDocs directorio (ID de organización) de su Amazon WorkDocs repositorio.
- Comprobó que cada documento es único en WorkDocs y entre las demás fuentes de datos que planea usar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En su AWS cuenta, asegúrese de tener:

- Creó un Amazon Kendra índice y, si utiliza la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Si no tiene un IAM rol existente, puede usar la consola para crear un nuevo IAM rol cuando conecte su fuente de WorkDocs datos a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol existente y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de WorkDocs datos, debe proporcionar los detalles necesarios de la fuente de WorkDocs datos para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configurado WorkDocs Amazon Kendra, consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Amazon WorkDocs

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el WorkDocs conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el WorkDocs conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. ID de organización específico de su Amazon WorkDocs sitio: seleccione el ID del Amazon WorkDocs sitio que desea indexar. Ya debe haber creado un sitio.

b. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

c. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. Rastrear los comentarios de los documentos: las entidades Amazon WorkDocs o los tipos de contenido que quiere rastrear.

b. Utilizar registros de cambios: seleccione esta opción para actualizar el índice solo con contenido nuevo o modificado en lugar de sincronizar todos los archivos.

c. Patrones regex: patrones de expresiones regulares para incluir o excluir determinados archivos.

d. Calendario de ejecución sincronizado para la frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Amazon WorkDocs

Debe especificar lo siguiente mediante la WorkDocsConfigurationAPI:

- Amazon WorkDocs ID de directorio: especifique el ID de organización de su Amazon WorkDocs directorio. Puede encontrar el ID de la organización en el AWS Directory Service yendo a Activar directorio y luego a Directorios.
- Función de IAM: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar una IAM función con permisos para acceder al WorkDocs directorio y para llamar a las API públicas necesarias para el conector y. WorkDocs Amazon Kendra Para obtener más información, consulte Funciones de IAM para las fuentes de datos. WorkDocs

También puede añadir las siguientes características opcionales:

- Registro de cambios: si se Amazon Kendra debe utilizar el mecanismo de registro de cambios de la fuente de WorkDocs datos para determinar si un documento debe actualizarse en el índice.

> **Note**
>
> Utilice el registro de cambios si no quiere que Amazon Kendra digitalice todos los documentos. Si el registro de cambios es grande, es posible que se Amazon Kendra tarde menos en digitalizar los documentos de la fuente de WorkDocs datos que en procesar el registro de cambios. Si sincroniza la fuente de WorkDocs datos con el índice por primera vez, se digitalizarán todos los documentos.

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados documentos y comentarios de documentos. Cada comentario se indexa como un documento independiente.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos de la fuente de WorkDocs datos a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de WorkDocs datos, consulte:

- Comience con el WorkDocs conector Amazon Kendra Amazon

Box (Cuadro)

Box es un servicio de almacenamiento en la nube que ofrece capacidades de alojamiento de archivos. Puedes usarlo Amazon Kendra para indexar el contenido de tu Box, incluidos los comentarios, las tareas y los enlaces web.

Puedes conectarte Amazon Kendra a tu fuente de datos de Box mediante la Amazon Kendra consola y la BoxConfigurationAPI.

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Box, consultaSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra El conector de fuente de datos de Box admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- El registro de cambios y las sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Box, realiza estos cambios en Box y en tus AWS cuentas.

En Box, asegúrese de que:

- Tiene una cuenta Box Enterprise o Box Enterprise Plus.

- Configuró una aplicación personalizada de Box en la consola de desarrolladores de Box, con autenticación del lado del servidor mediante JSON Web Tokens (JWT). Consulte la documentación de Box sobre la creación de una aplicación personalizada y la documentación de Box sobre la configuración de JWT Auth para obtener más información.
- Ha establecido el nivel de acceso de la aplicación en App + Enterprise Access y le ha permitido Realizar llamadas a la API utilizando el encabezado como usuario.
- Ha usado el usuario administrador para agregar los siguientes Ámbitos de aplicación en su aplicación Box:
  - Escribir todos los archivos y carpetas almacenados en un Box
  - Administración de usuarios
  - Administrar grupos
  - Administrar propiedades empresariales
- Par de claves pública/privada configurado que incluye un ID de cliente, un secreto de cliente, un ID de clave pública, un ID de clave privada, una contraseña y un ID empresarial para usar como credenciales de autenticación. Consulte Par de claves públicas y privadas para obtener más información.

  > **Note**
  >
  > Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha copiado el ID empresarial de Box de la configuración de la consola para desarrolladores de Box o de la aplicación Box. Por ejemplo: `801234567`.
- Ha comprobado que cada documento es único en Box y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Box en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de datos de Box Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Box, debe proporcionar los detalles necesarios de su fuente de datos de Box para que Amazon Kendra pueda acceder a sus datos. Si todavía no has configurado Box for Amazon Kendra, consultaRequisitos previos.

Console

Para conectarse Amazon Kendra a Box

1. Inicia sesión en la Amazon Kendra consola AWS Management Console y ábrela.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector Box y, a continuación, selecciona Añadir conector. Si utiliza la versión 2 (si corresponde), elija el conector Box con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
    a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
    b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
    c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
    d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
    e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
    a. ID empresarial de Box: introduzca el ID empresarial de Box. Por ejemplo: *801234567*.
    b. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
    c. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de autenticación de Box. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .
        i. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Box» se añade automáticamente a su nombre secreto.
        ii. Para el identificador de cliente, el secreto de cliente, el identificador de clave pública, el identificador de clave privada y la contraseña, introduzca los valores de la clave pública o privada que configuró en Box.
        iii. Añada y guarde su secreto.

d. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

e. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

f. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

g. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. ID de carpeta de Box: introduzca algunos ID de carpeta de Box que desee rastrear; de lo contrario, se rastreará el contenido de todas las carpetas.

b. Archivos de Box: elige si deseas rastrear los enlaces web, los comentarios y las tareas.

c. Para una configuración adicional: añada patrones de expresiones regulares para incluir o excluir cierto contenido.

d. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

e. Calendario de ejecución sincronizado para la frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

f. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Box

Debe especificar lo siguiente mediante la BoxConfigurationAPI:

ID empresarial de Box: proporcione el ID empresarial de Box. Puede encontrar el ID empresarial en la configuración de Box Developer Console o al configurar una aplicación en Box.

- Nombre secreto del recurso de Amazon (ARN): proporciona el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de tu cuenta de Box. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "clientID": "client-id",
    "clientSecret": "client-secret",
    "publicKeyID": "public-key-id",
    "privateKey": "private-key",
    "passphrase": "pass-phrase"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Box y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Box.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique `VpcConfiguration` como parte de la configuración del origen de datos. Consulte Configuración de Amazon Kendra para utilizar una VPC.
- Registro de cambios: si se Amazon Kendra debe utilizar el mecanismo de registro de cambios de la fuente de datos de Box para determinar si un documento debe actualizarse en el índice.

> **Note**
>
> Utilice el registro de cambios si no quiere que Amazon Kendra digitalice todos los documentos. Si el registro de cambios es grande, es posible que se tarde Amazon Kendra menos en digitalizar los documentos de la fuente de datos de Box que en procesar el registro de cambios. Si está sincronizando el origen de datos de Box con su índice por primera vez, se escanean todos los documentos.

- Comentarios, tareas y enlaces web: especifique si desea rastrear este tipo de contenido.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los

documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos y carpetas de Box.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos del origen de datos de Box a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de datos de Box, consulte:

- Cómo empezar con el conector Amazon Kendra Box

Confluence

Confluence es una herramienta colaborativa de gestión del trabajo diseñada para compartir, almacenar y trabajar en la planificación de proyectos, el desarrollo de software y la gestión de productos. Puedes usarlo Amazon Kendra para indexar tus espacios, páginas (incluidas las páginas anidadas), blogs, comentarios y archivos adjuntos de páginas y blogs indexados de Confluence.

Amazon Kendra es compatible con Confluence Server/Data Center y Confluence Cloud.

> **Note**
>
> De forma predeterminada, Amazon Kendra no indexa los archivos ni los espacios personales de Confluence. Puede elegir indexarlos al crear el origen de datos. Si no quieres Amazon Kendra indexar un espacio, márcalo como privado en Confluence.

Puedes conectarte Amazon Kendra a tu fuente de datos de Confluence mediante la Amazon Kendra consola, la TemplateConfigurationAPI o la ConfluenceConfigurationAPI.

Amazon Kendra tiene dos versiones del conector de Confluence. Las características compatibles de cada versión incluyen:

Conector Confluence V1.0/API ConfluenceConfiguration

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- (Solo para Confluence Server) Nube privada virtual (VPC)

Conector Confluence V2.0/ API TemplateConfiguration

- Asignaciones de campo
- Control de acceso de usuarios
- Patrones de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

> **Note**
>
> Está previsto que el soporte para el conector ConfluenceConfiguration V1.0/API de Confluence finalice en 2023. Recomendamos migrar al conector V2.0/API de Confluence o usarlo. TemplateConfiguration

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Confluence, consulte. Solución de problemas con los orígenes de datos

Temas

- Confluence Connector V1.0
- Confluence Connector V2.0

Confluence Connector V1.0

Confluence es una herramienta colaborativa de gestión del trabajo diseñada para compartir, almacenar y trabajar en la planificación de proyectos, el desarrollo de software y la gestión de productos. Puede utilizar Amazon Kendra para indexar espacios, páginas (incluidas las páginas anidadas), blogs, comentarios y archivos adjuntos de páginas y blogs indexados de Confluence.

> **Note**
>
> Está previsto que el soporte para el conector ConfluenceConfiguration V1.0/API de Confluence finalice en 2023. Recomendamos migrar al conector V2.0/API de Confluence o usarlo. TemplateConfiguration

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Confluence, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra El conector de fuente de datos de Confluence admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- (Solo para Confluence Server) Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Confluence, realiza estos cambios en tu Confluence y AWS en tus cuentas.

En Confluence, asegúrese de que:

- Has concedido Amazon Kendra permisos para ver todo el contenido de tu instancia de Confluence de la siguiente manera:
  - Convirtiéndose Amazon Kendra en miembro de un `confluence-administrators` grupo.
  - Ha otorgado permisos de administrador del sitio para todos los espacios, blogs y páginas existentes.
- Ha copiado la dirección URL de la instancia de Confluence.
- Para los usuarios de SSO (inicio de sesión único): ha activado la página Mostrar al iniciar sesión para el nombre de usuario y la contraseña al configurar los Métodos de autenticación de Confluence en el centro de datos de Confluence.
- Para Confluence Server
  - Ha apuntado sus credenciales de autenticación básica, que incluyen el nombre de usuario y la contraseña de su cuenta administrativa de Confluence para conectarse a Amazon Kendra.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

  - Opcional: ha generado un token de acceso personal en la cuenta de Confluence para conectarse a Amazon Kendra. Para obtener más información, consulte la Documentación de Confluence sobre la generación de tokens de acceso personal.
- Para Confluence Cloud
  - Ha apuntado sus credenciales de autenticación básica, que incluyen el nombre de usuario y la contraseña de su cuenta administrativa de Confluence para conectarse a Amazon Kendra.
- Ha comprobado que cada documento es único en Confluence y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Confluence en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes un IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar tu fuente de datos de Confluence. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarte Amazon Kendra a tu fuente de datos de Confluence, debes proporcionar los detalles de tus credenciales de Confluence para que Amazon Kendra puedas acceder a tus datos. Si aún no has configurado Confluence para consultarlo. Amazon Kendra Requisitos previos

Console

Para conectarse a Amazon Kendra Confluence

1. Inicia sesión en la consola AWS de administración y abre la Amazon Kendra consola.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar origen de datos, elija el Confluence Connector V1.0 y, a continuación, elija Agregar origen de datos.
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. Elige entre Confluence Cloud y Confluence Server.
   b. Si eliges Confluence Cloud, introduce la siguiente información:

i. URL de Confluence: la URL de Confluence.

ii. AWS Secrets Manager secreto: elige un secreto existente o crea uno nuevo para almacenar tus Secrets Manager credenciales de autenticación de Confluence. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

- Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

I. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Confluence» se añade automáticamente a tu nombre secreto.

II. Para el nombre de usuario y la contraseña: introduce tu nombre de usuario y contraseña de Confluence.

III. Seleccione Guardar autenticación.

c. Si eliges Confluence Server, introduce la siguiente información:

i. URL de Confluence: su nombre de usuario y contraseña de Confluence.

ii. (Opcional) Para el Proxy web, introduzca la siguiente información:

A. Nombre de host: nombre de host de su cuenta de Confluence.

B. Número de puerto: puerto utilizado por el protocolo de transporte de URL del host.

iii. Para la autenticación, elige la autenticación básica o el token de acceso personal (solo en el servidor de Confluence).

iv. AWS Secrets Manager secreto: elige un secreto existente o crea uno nuevo para almacenar tus Secrets Manager credenciales de autenticación de Confluence. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

- Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

I. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Confluence» se añade automáticamente a tu nombre secreto.

II. Para el nombre de usuario y la contraseña: introduce los valores de las credenciales de autenticación que configuraste en Confluence. Si utilizas la autenticación básica, utiliza tu nombre de usuario (ID de correo electrónico) y contraseña (token de API) de Confluence. Si utilizas un token de acceso personal, introduce los detalles del token de acceso personal que configuraste en la cuenta de Confluence.

III. Guarda y añade tu secreto.

d. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

e. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Incluir espacios personales e Incluir espacios archivados: elija los tipos de espacio opcionales que desee incluir en este origen de datos.

   b. En Configuración adicional: especifique los patrones de expresión regular para incluir o excluir cierto contenido. Puede agregar hasta 100 patrones.

   c. También puede elegir Rastrear archivos adjuntos dentro de los espacios elegidos.

   d. Calendario de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que Amazon Kendra se sincronizará con la fuente de datos.

   e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Para espacio, página o blog: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados o las asignaciones de campos sugeridas adicionales para agregar campos de índice.

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Confluence Amazon Kendra

Debes especificar lo siguiente mediante la ConfluenceConfigurationAPI:

- Versión de Confluence: especifique la versión de la instancia de Confluence que está utilizando como `CLOUD` o `SERVER`.
- Nombre secreto de recurso de Amazon (ARN): proporciona el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga tus credenciales de autenticación de Confluence.

  Si utilizas Confluence Server, puedes usar tu nombre de usuario y contraseña de Confluence o tu token de acceso personal como credenciales de autenticación.

  Si utilizas tu nombre de usuario y contraseña de Confluence como credenciales de autenticación, guardas las siguientes credenciales como una estructura JSON en tu secreto: Secrets Manager

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  Si utilizas un token de acceso personal para conectarte a Confluence Server Amazon Kendra, guardas las siguientes credenciales como una estructura JSON en tu Secrets Manager secreto:

  ```
  {
      "patToken": "personal access token"
  }
  ```

  Si utilizas Confluence Cloud, utilizas tu nombre de usuario de Confluence y un token de API, configurado en Confluence, como contraseña. Guardas las siguientes credenciales como una estructura JSON en tu secreto: Secrets Manager

```
{
    "username": "user name",
    "password": "API token"
}
```

- IAM rol: especifica `RoleArn` cuándo llamas `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a tu Secrets Manager secreto y para llamar a las API públicas necesarias para el conector de Confluence y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Confluence.

También puede añadir las siguientes características opcionales:

- Proxy web: si conectarse a la instancia de la URL de Confluence a través de un proxy web. Puede utilizar esta opción para Confluence Server.
- (Solo para Confluence Server) Nube privada virtual (VPC): especifique `VpcConfiguration` como parte de la configuración del origen de datos. Consulte Configuración Amazon Kendra para usar una VPC.
- Filtros de inclusión y exclusión: especifique patrones de expresiones regulares para incluir o excluir determinados espacios, publicaciones de blog, páginas y archivos adjuntos. Si decide indexar los archivos adjuntos, solo se indexarán los adjuntos de las páginas y blogs indexados.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Asignaciones de campos: elija asignar los campos del origen de datos de Confluence a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

Más información

Para obtener más información sobre la integración Amazon Kendra con tu fuente de datos de Confluence, consulta:

- Configuración del conector de Amazon Kendra Confluence Server

Confluence Connector V2.0

Confluence es una herramienta colaborativa de gestión del trabajo diseñada para compartir, almacenar y trabajar en la planificación de proyectos, el desarrollo de software y la gestión de productos. Puede utilizar Amazon Kendra para indexar espacios, páginas (incluidas las páginas anidadas), blogs, comentarios y archivos adjuntos de páginas y blogs indexados de Confluence.

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Confluence, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión

Características admitidas

Amazon Kendra El conector de fuente de datos de Confluence admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Patrones de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Confluence, realiza estos cambios en tu Confluence y en tus cuentas. AWS

En Confluence, asegúrese de que:

- Ha copiado la URL de la instancia de Confluence. Por ejemplo: *https://example.confluence.com*, *https://www.example.confluence.com/* o *https://atlassian.net/*. La URL de la instancia de Confluence se debe conectar a Amazon Kendra.

  Si utilizas Confluence Cloud, la URL de tu host debe terminar en atlassian.net/.

> **Note**
>
> No se admiten los siguientes formatos de URL:
>
> - *https://example.confluence.com/xyz*
> - *https://www.example.confluence.com//wiki/spacekey/xxx*
> - *https://atlassian.net/xyz*

> **Note**
>
> (local o en el servidor) Amazon Kendra comprueba si la información del punto final incluida AWS Secrets Manager es la misma que la información del punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a evitar el problema del suplente confuso, que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, pero utiliza Amazon Kendra como proxy para acceder al secreto

configurado y realizar la acción. Si más adelante cambia la información de punto de conexión, debe crear un nuevo secreto para sincronizar esta información.

- Credenciales de autenticación básicas configuradas que contienen un nombre de usuario (el ID de correo electrónico utilizado para iniciar sesión en Confluence) y una contraseña (el token de la API de Confluence es la contraseña). Consulta Administrar los tokens de API para tu cuenta de Atlassian.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Opcional: las credenciales de OAuth 2.0 configuradas contienen una clave de aplicación de Confluence, un secreto de aplicación de Confluence, un token de acceso de Confluence y un token de actualización de Confluence para poder conectarte a tu instancia de Confluence. Amazon Kendra Si el token de acceso caduca, puede usar el token de actualización para regenerar el token de acceso y actualizar el par de tokens. También puede repetir el proceso de autorización. Para más información sobre los tokens de acceso, consulte Administrar los tokens de acceso de OAuth.
- (Solo para el servidor o el centro de datos de Confluence) Opcional: configuraste un token de acceso personal (PAT) en Confluence. Consulte Uso de tokens de acceso personal.

En su Cuenta de AWS interior, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Confluence en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes un IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar tu fuente de datos de Confluence. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarte Amazon Kendra a tu fuente de datos de Confluence, debes proporcionar los detalles necesarios de tu fuente de datos de Confluence para que Amazon Kendra puedas acceder a tus datos. Si aún no ha configurado Confluence para Amazon Kendra , consulte Requisitos previos.

Console

Para conectarte a Confluence Amazon Kendra

1. Inicia sesión en la Amazon Kendra consola AWS Management Console y ábrela.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector de Confluence y, a continuación, selecciona Añadir conector. Si utilizas la versión 2 (si corresponde), elige el conector de Confluence con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elige un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. En Fuente, selecciona Confluence Cloud o Confluence Server/Data Center.

b. URL de Confluence: introduzca la URL del host de Confluence. Por ejemplo, *`https://example.confluence.com`*.

c. (Solo para Confluence Server/Data Center) Ubicación del certificado SSL: opcional: introduce la Amazon S3 ruta del archivo de certificado SSL para Confluence Server.

d. (Solo para Confluence Server/Data Center) Proxy web: opcional: introduce el nombre de host del proxy web (sin el `https://` protocolo `http://` OR) y el número de puerto (puerto utilizado por el protocolo de transporte de URL del host). El número de puerto debe ser un valor numérico entre 0 y 65535.

e. Autorización: activa o desactiva la información de la lista de control de acceso (ACL) en tus documentos, si tienes una ACL y quieres usarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

f. Autenticación: elige entre la autenticación básica, la autenticación Oauth 2.0 o la autenticación mediante token de acceso personal (solo para el servidor o centro de datos de Confluence).

g. Secreto de AWS Secrets Manager : elija un secreto existente o cree un nuevo secreto de Secrets Manager para almacenar sus credenciales de autenticación de Confluence. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager . En la ventana, introduzca la siguiente información:

i. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Confluence» se añade automáticamente a tu nombre secreto.

ii. Si utilizas la autenticación básica: introduce el nombre secreto, el nombre de usuario y la contraseña (el token de la API de Confluence es la contraseña) que configuraste en Confluence.

Si utilizas la autenticación OAuth2.0: introduce el nombre secreto, la clave de la aplicación, el secreto de la aplicación, el token de acceso y el token de actualización que configuraste en Confluence.

(Solo en Confluence Server/Data Center) Si utilizas la autenticación con token de acceso personal: introduce el nombre secreto y el token de Confluence que configuraste en tu Confluence.

iii. Guarda y añade tu secreto.

h. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

i. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

j. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

k. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. En el ámbito de la sincronización, para sincronizar contenido: elija sincronizar entre los siguientes tipos de contenido: páginas, comentarios de página, archivos adjuntos de página, blogs, comentarios de blog, archivos adjuntos de blog, espacios personales y espacios archivados.

> **Note**
>
> Los comentarios de página y los archivos adjuntos de página solo se pueden seleccionar si eliges sincronizar las páginas. Los comentarios y los archivos adjuntos del blog solo se pueden seleccionar si eliges sincronizar los blogs.

> **Important**
>
> Si no especificas un patrón de expresiones regulares con teclas espaciadoras en la configuración adicional, se rastrearán todas las páginas y blogs de forma predeterminada.

b. En Configuración adicional, en Tamaño máximo de archivo: especifique el límite de tamaño del archivo en MB que se rastreará. Amazon Kendra Amazon Kendra rastreará solo los archivos que se encuentren dentro del límite de tamaño que usted defina. El tamaño predeterminado del archivo es de 50 MB. El tamaño máximo del archivo debe ser superior a 0 MB e inferior o igual a 50 MB.

Para los patrones de expresiones regulares de Spaces: especifique si desea incluir o excluir espacios específicos del índice mediante:

- *Tecla espaciadora (por ejemplo, my-space-123)*

> **Note**
>
> Si no especificas un patrón de expresiones regulares con la tecla espaciadora, se rastrearán todas las páginas y blogs de forma predeterminada.

- *URL (por ejemplo,. *//MySiteMyDocuments/*)

- Tipo de archivo (por ejemplo, `.*\ .pdf, .*\ .txt)`

Para los patrones de expresiones regulares de títulos de entidades: especifique patrones de expresiones regulares para incluir o excluir determinados blogs, páginas, comentarios y archivos adjuntos por título.

> **Note**
>
> Si quieres incluir o excluir el rastreo de una página o subpágina específica, puedes usar los patrones de expresiones regulares de los títulos de las páginas.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione uno de los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice. Para agregar campos de origen de datos personalizados, cree un nombre de campo de índice para asignarlos y el tipo de datos del campo.

b. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarte Amazon Kendra a Confluence

Debes especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `CONFLUENCEV2` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- URL del host: especifica la instancia de URL del host de Confluence. Por ejemplo, *`https://example.confluence.com`*.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Tipo de autenticación: especifique el tipo de autenticación, ya sea `BasicOAuth2`, (solo en Confluence Server). `Personal-token`
- (Opcional, solo para Confluence Server) Ubicación del certificado SSL: especifique el `S3bucketName` y `s3certificateName` que utilizó para almacenar su certificado SSL.
- Nombre de recurso secreto de Amazon (ARN): proporciona el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que

configuraste en Confluence. Si utiliza la autenticación básica, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "username": "email ID or user name",
    "password": "Confluence API token"
}
```

Si utiliza la autenticación OAuth 2.0, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "confluenceAppKey": "app key",
    "confluenceAppSecret": "app secret",
    "confluenceAccessToken": "access token",
    "confluenceRefreshToken": "refresh token"
}
```

(Solo para el servidor Confluence) Si utiliza la autenticación básica, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "hostUrl": "Confluence Server host URL",
    "username": "Confluence Server user name",
    "password": "Confluence Server password"
}
```

(Solo para el servidor Confluence) Si utiliza la autenticación con token de acceso personal, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "hostUrl": "Confluence Server host URL",
    "patToken": "personal access token"
}
```

- IAM rol: especifica `RoleArn` cuándo llamas `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a tu Secrets Manager secreto y para llamar a las API públicas necesarias para el conector de Confluence y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Confluence.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Tamaño de archivo: especifica el tamaño máximo de archivo que se va a rastrear.
- Tipos de documento o contenido: especifique si desea rastrear las páginas, los comentarios de las páginas, los archivos adjuntos de las páginas, los blogs, los comentarios de los blogs, los archivos adjuntos de los blogs, los espacios y los espacios archivados.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados espacios, páginas, blogs y sus comentarios y archivos adjuntos.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Proxy web: especifica la información de tu proxy web si quieres conectarte a tu instancia de URL de Confluence a través de un proxy web. Puede utilizar esta opción para Confluence Server.
- Lista de control de acceso (ACL): especifique si desea rastrear la información de la ACL de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en

todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

- Asignaciones de campos: elija asignar los campos del origen de datos de Confluence a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para ver una lista de otras claves JSON importantes que debes configurar, consulta el esquema de plantillas de Confluence.

Notas

- El token de acceso personal (PAT) no está disponible para Confluence Cloud.

Conector de orígenes de datos personalizados

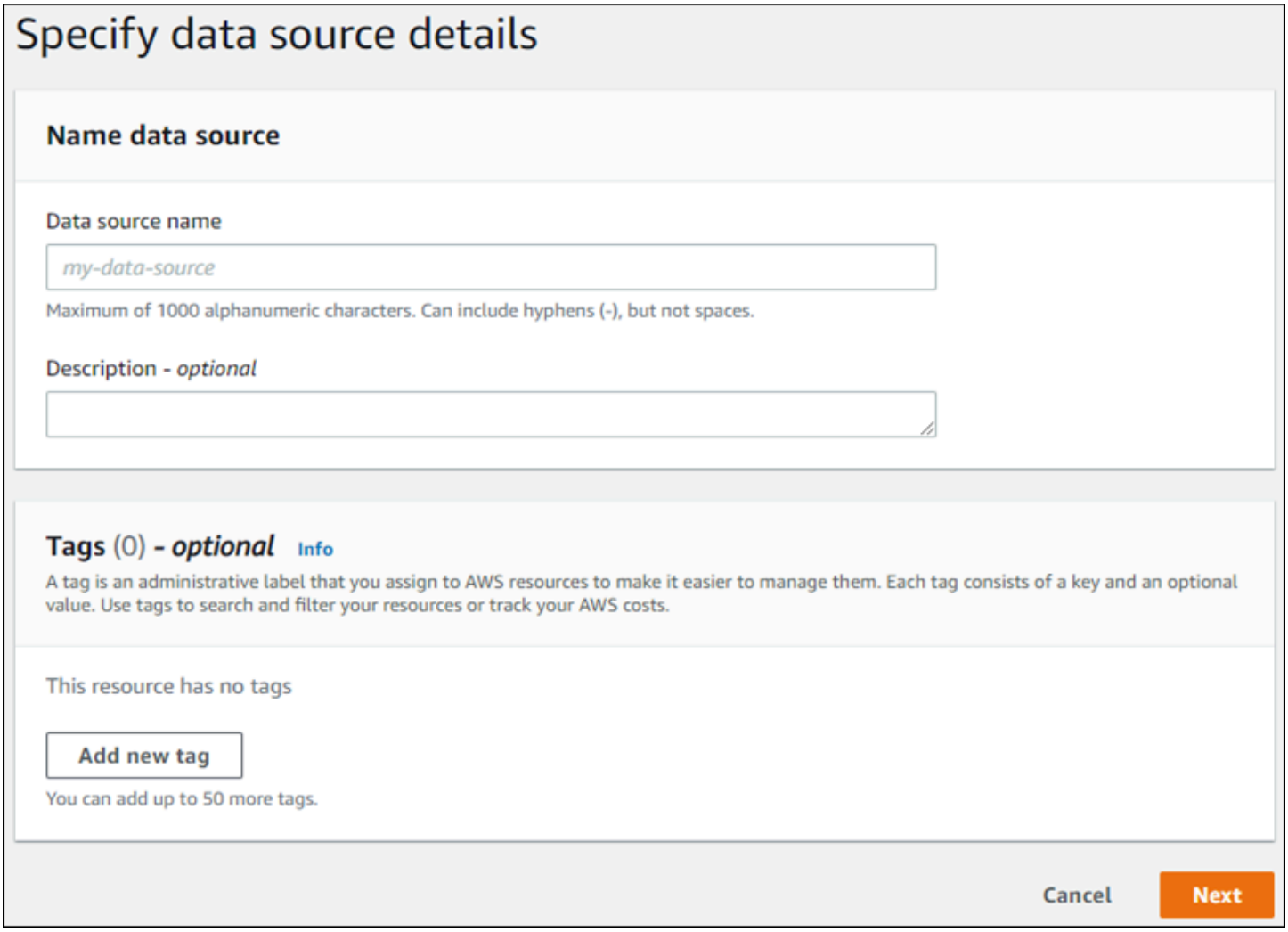
Utilice una fuente de datos personalizada cuando tenga un repositorio para el que aún Amazon Kendra no haya un conector de fuente de datos. Puedes usarlo para ver las mismas métricas del historial de ejecución que proporcionan las fuentes de Amazon Kendra datos, incluso si no puedes Amazon Kendra usarlas para sincronizar tus repositorios. Utilízala para crear una experiencia de supervisión de sincronización coherente entre las fuentes de Amazon Kendra datos y las personalizadas. En concreto, utilice una fuente de datos personalizada para ver las métricas de sincronización de un conector de fuente de datos que haya creado con las API BatchPutDocument y BatchDeleteDocument.

Para solucionar problemas del conector de origen de datos personalizado de Amazon Kendra, consulte Solución de problemas con los orígenes de datos.

Al crear una fuente de datos personalizada, tiene un control total sobre cómo se seleccionan los documentos que se van a indexar. Amazon Kendra solo proporciona información métrica que puede utilizar para supervisar los trabajos de sincronización de la fuente de datos. Debe crear y ejecutar el rastreador que determina los documentos que indexa su origen de datos.

Debe especificar el título principal de los documentos mediante el objeto Document y `_source_uri` para `DocumentTitle` `DocumentURI` incluirlo en la respuesta del `Query` resultado. DocumentAttribute

Puede crear un identificador para su fuente de datos personalizada mediante la consola o mediante la API CreateDataSource. Para usar la consola, asigne un nombre al origen de datos y, si lo desea, una descripción y etiquetas de recursos. Una vez creado el origen de datos, se muestra el ID correspondiente. Copie este ID para usarlo cuando sincronice el origen de datos con el índice.



También puede crear un origen de datos personalizada mediante la API `CreateDataSource`. La API devuelve un ID para usarlo al sincronizar el origen de datos. Cuando utiliza la API

`CreateDataSource` para crear un origen de datos personalizado, no puede configurar los parámetros `Configuration`, `RoleArn` o `Schedule`. Si estableces estos parámetros, Amazon Kendra devuelve una `ValidationException` excepción.

Para usar un origen de datos personalizado, cree una aplicación que se encargue de actualizar el índice de Amazon Kendra . La aplicación depende del rastreador que cree. El rastreador lee los documentos del repositorio y determina cuáles se deben enviar a Amazon Kendra. La aplicación debe realizar los pasos siguientes:

1. Rastrear el repositorio y hacer una lista de los documentos del repositorio que se han agregado, actualizado o eliminado.
2. Llama a la API StartDataSourceSyncJob para indicar que se está iniciando un trabajo de sincronización. Debe proporcionar un ID de fuente de datos para identificar la fuente de datos que se está sincronizando. Amazon Kendra devuelve un identificador de ejecución para identificar un trabajo de sincronización concreto.
3. Llama a la BatchDelete API de documentos para eliminar documentos del índice. Se proporciona el ID del origen de datos y el ID de ejecución para identificar el origen de datos que se está sincronizando y el trabajo al que está asociada esta actualización.
4. Llama a la API StopDataSourceSyncJob para indicar el final del trabajo de sincronización. Después de llamar a la API `StopDataSourceSyncJob`, el ID de ejecución asociado deja de ser válido.
5. Llama a la API de ListDataSourceSyncJobs con los identificadores del índice y de la fuente de datos para enumerar los trabajos de sincronización de la fuente de datos y ver las métricas de los trabajos de sincronización.

Tras finalizar un trabajo de sincronización, se puede iniciar uno nuevo. Puede transcurrir un tiempo antes de que todos los documentos enviados se añadan al índice. Use la API `ListDataSourceSyncJobs` para ver el estado del trabajo de sincronización. Si el `Status` devuelto para el trabajo de sincronización es `SYNCING_INDEXING`, algunos documentos aún se están indexando. Puede iniciar un nuevo trabajo de sincronización cuando el estado del trabajo anterior sea `FAILED` o`SUCCEEDED`.

Después de llamar a la API `StopDataSourceSyncJob`, no se puede usar un identificador de trabajo de sincronización en una llamada a las API `BatchPutDocument` o `BatchDeleteDocument`. Si lo hace, todos los documentos enviados se devolverán en el mensaje de respuesta `FailedDocuments` de la API.

Atributos obligatorios

Al enviar un documento para Amazon Kendra utilizar la `BatchPutDocument` API, cada documento requiere dos atributos para identificar la fuente de datos y la ejecución de sincronización a la que pertenece. Debe proporcionar los dos atributos siguientes para asignar correctamente los documentos del origen de datos personalizado a un índice de Amazon Kendra :

- `_data_source_id`: el identificador del origen de datos. Este se devuelve al crear el origen de datos con la consola o la API `CreateDataSource`.
- `_data_source_sync_job_execution_id`: el identificador de la ejecución de sincronización. Se devuelve al iniciar la sincronización del índice con la API `StartDataSourceSyncJob`.

El siguiente es el JSON necesario para indexar un documento mediante un origen de datos personalizado.

```
{
    "Documents": [
        {
            "Attributes": [
                {
                    "Key": "_data_source_id",
                    "Value": {
                        "StringValue": "data source identifier"
                    }
                },
                {
                    "Key": "_data_source_sync_job_execution_id",
                    "Value": {
                        "StringValue": "sync job identifier"
                    }
                }
            ],
            "Blob": "document content",
            "ContentType": "content type",
            "Id": "document identifier",
            "Title": "document title"
        }
    ],
    "IndexId": "index identifier",
    "RoleArn": "IAM role ARN"
```

```
}
```

Al eliminar un documento del índice mediante la API `BatchDeleteDocument`, se deben especificar los dos campos siguientes en el parámetro `DataSourceSyncJobMetricTarget`:

- `DataSourceId`: el identificador del origen de datos. Este se devuelve al crear el origen de datos con la consola o la API `CreateDataSource`.
- `DataSourceSyncJobId`: el identificador de la ejecución de sincronización. Se devuelve al iniciar la sincronización del índice con la API `StartDataSourceSyncJob`.

El siguiente es el JSON necesario para eliminar un documento del índice mediante la API `BatchDeleteDocument`.

```
{
    "DataSourceSyncJobMetricTarget": {
        "DataSourceId": "data source identifier",
        "DataSourceSyncJobId": "sync job identifier"
    },
    "DocumentIdList": [
        "document identifier"
    ],
    "IndexId": "index identifier"
}
```

Visualización de métricas

Una vez finalizado un trabajo de sincronización, puedes usar la API de DataSourceSyncJobmétricas para obtener las métricas asociadas al trabajo de sincronización. Utilícela para supervisar las sincronizaciones de sus orígenes de datos personalizados.

Si se envía el mismo documento varias veces, ya sea como parte de la API `BatchPutDocument`, la API `BatchDeleteDocument` o si el documento se envía tanto para su adición como para su eliminación, el documento solo se cuenta una vez en las métricas.

- `DocumentsAdded`: la cantidad de documentos enviados mediante la API `BatchPutDocument` asociada a este trabajo de sincronización agregados al índice por primera vez. Si un documento se envía para agregarlo más de una vez en una sincronización, el documento solo se cuenta una vez en las métricas.

- `DocumentsDeleted`: la cantidad de documentos enviados mediante la API `BatchDeleteDocument` asociada a este trabajo de sincronización eliminados del índice. Si un documento se envía para eliminarlo más de una vez en una sincronización, el documento solo se cuenta una vez en las métricas.
- `DocumentsFailed`: el número de documentos asociados a este trabajo de sincronización que no se pudieron indexar. Se trata de documentos que fueron aceptados por Amazon Kendra para su indexación, pero que no se pudieron indexar ni eliminar. Si un documento no es aceptado por Amazon Kendra, el identificador del documento se devuelve en la propiedad de `FailedDocuments` respuesta de las `BatchDeleteDocument` API `BatchPutDocument` y.
- `DocumentsModified`—El número de documentos modificados enviados mediante la `BatchPutDocument` API asociada a este trabajo de sincronización y que se modificaron en el Amazon Kendra índice.

Amazon Kendra también emite Amazon CloudWatch métricas al indexar los documentos. Para obtener más información, consulte Amazon Kendra Monitorear con. Amazon CloudWatch

Amazon Kendra no devuelve la `DocumentsScanned` métrica de las fuentes de datos personalizadas. También emite CloudWatch las métricas que figuran en el documento Métricas de las fuentes de Amazon Kendra datos.

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos personalizada, consulte:

- Añadir fuentes de datos personalizadas a Amazon Kendra

Origen de datos personalizado (Java)

El código siguiente proporciona un ejemplo de implementación de un origen de datos personalizado mediante Java. El programa crea primero un origen de datos personalizado y, a continuación, sincroniza los documentos recién agregados al índice con dicho origen.

El código siguiente muestra la creación y el uso de un origen de datos personalizado. Al utilizar un origen de datos personalizado en la aplicación, no es necesario crear un nuevo origen de datos (un proceso único) cada vez que se sincronice el índice con él. Se utiliza el ID de índice y el ID del origen de datos para sincronizar los datos.

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import csoftware.amazon.awssdk.services.kendra.model.BatchPutDocumentRequest;
import csoftware.amazon.awssdk.services.kendra.model.BatchPutDocumentResponse;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.Document;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;
import software.amazon.awssdk.services.kendra.model.StopDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StopDataSourceSyncJobResponse;

public class SampleSyncForCustomDataSource {
  public static void main(String[] args) {
    KendraClient kendra = KendraClient.builder().build();

    String myIndexId = "yourIndexId";
    String dataSourceName = "custom data source";
    String dataSourceDescription = "Amazon Kendra custom data source connector"

    // Create custom data source
    CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
        .builder()
        .indexId(myIndexId)
        .name(dataSourceName)
        .description(dataSourceDescription)
        .type(DataSourceType.CUSTOM)
        .build();

    CreateDataSourceResponse createDataSourceResponse =
 kendra.createDataSource(createDataSourceRequest);
    System.out.println(String.format("Response of creating data source: %s",
 createDataSourceResponse));

    // Get the data source ID from createDataSourceResponse
    String dataSourceId = createDataSourceResponse.Id();

    // Wait for the custom data source to become active
```

```
   System.out.println(String.format("Waiting for Amazon Kendra to create the data
source %s", dataSourceId));
   // You can use the DescribeDataSource API to check the status
   DescribeDataSourceRequest describeDataSourceRequest = DescribeDataSourceRequest
       .builder()
       .indexId(myIndexId)
       .id(dataSourceId)
       .build();

   while (true) {
       DescribeDataSourceResponse describeDataSourceResponse =
kendra.describeDataSource(describeDataSourceRequest);

       DataSourceStatus status = describeDataSourceResponse.status();
       System.out.println(String.format("Creating data source. Status: %s", status));
       if (status != DataSourceStatus.CREATING) {
           break;
       }

       TimeUnit.SECONDS.sleep(60);
   }

   // Start syncing yor data source by calling StartDataSourceSyncJob and providing
your index ID
   // and your custom data source ID
   System.out.println(String.format("Synchronize the data source %s", dataSourceId));
   StartDataSourceSyncJobRequest startDataSourceSyncJobRequest =
StartDataSourceSyncJobRequest
       .builder()
       .indexId(myIndexId)
       .id(dataSourceId)
       .build();
   StartDataSourceSyncJobResponse startDataSourceSyncJobResponse =
kendra.startDataSourceSyncJob(startDataSourceSyncJobRequest);

   // Get the  sync job execution ID from startDataSourceSyncJobResponse
   String executionId = startDataSourceSyncJobResponse.ExecutionId();
System.out.println(String.format("Waiting for the data source to sync with the index
%s for execution ID %s", indexId, startDataSourceSyncJobResponse.executionId()));

   // Add 2 documents uploaded to S3 bucket to your index using the BatchPutDocument
API
   // The added documents should sync with your custom data source
   Document pollyDoc = Document
```

```
        .builder()
        .s3Path(
            S3Path.builder()
            .bucket("s3-test-bucket")
            .key("what_is_Amazon_Polly.docx")
            .build())
        .title("What is Amazon Polly?")
        .id("polly_doc_1")
        .build();

    Document rekognitionDoc = Document
        .builder()
        .s3Path(
            S3Path.builder()
            .bucket("s3-test-bucket")
            .key("what_is_amazon_rekognition.docx")
            .build())
        .title("What is Amazon rekognition?")
        .id("rekognition_doc_1")
        .build();

    BatchPutDocumentRequest batchPutDocumentRequest = BatchPutDocumentRequest
        .builder()
        .indexId(myIndexId)
        .documents(pollyDoc, rekognitionDoc)
        .build();

    BatchPutDocumentResponse result = kendra.batchPutDocument(batchPutDocumentRequest);
    System.out.println(String.format("BatchPutDocument result: %s", result));

    // Once custom data source synced, stop the sync job using the
StopDataSourceSyncJob API
    StopDataSourceSyncJobResponse stopDataSourceSyncJobResponse =
kendra.stopDataSourceSyncJob(
        StopDataSourceSyncJobRequest()
            .indexId(myIndexId)
            .id(dataSourceId)
    );

// List your sync jobs
    ListDataSourceSyncJobsRequest listDataSourceSyncJobsRequest =
ListDataSourceSyncJobsRequest
        .builder()
        .indexId(myIndexId)
```

```
        .id(dataSourceId)
        .build();

    while (true) {
        ListDataSourceSyncJobsResponse listDataSourceSyncJobsResponse =
 kendra.listDataSourceSyncJobs(listDataSourceSyncJobsRequest);
        DataSourceSyncJob job = listDataSourceSyncJobsResponse.history().get(0);
        System.out.println(String.format("Status: %s", job.status()));
    }
  }
}
```

Dropbox

Dropbox es un servicio de alojamiento de archivos que ofrece servicios de almacenamiento en la nube, organización de documentos y creación de plantillas de documentos. Si eres usuario de Dropbox, puedes usarlo Amazon Kendra para indexar tus archivos de Dropbox, Dropbox Paper, las plantillas de Dropbox Paper y los accesos directos a páginas web almacenados. También puedes configurarlos Amazon Kendra para indexar archivos específicos de Dropbox, Dropbox Paper, plantillas de Dropbox Paper y accesos directos a páginas web almacenados.

Amazon Kendra es compatible con Dropbox y Dropbox Advanced para Dropbox Business.

Puedes conectarte Amazon Kendra a tu fuente de datos de Dropbox mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuentes de datos de Amazon Kendra Dropbox, consultaSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra El conector de fuentes de datos de Dropbox admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de los usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Dropbox, realiza estos cambios en tu Dropbox y en tus AWS cuentas.

En Dropbox, asegúrese de que:

- Ha creado una cuenta de Dropbox Advanced y configurado un usuario administrador.
- Configuraste una aplicación de Dropbox con un nombre de aplicación único y activaste el acceso limitado. Consulte la Documentación de Dropbox sobre la creación de una aplicación.
- Ha activado los permisos Full Dropbox en la consola de Dropbox y agregado los siguientes permisos:
  - files.content.read
  - files.metadata.read
  - sharing.read
  - file_requests.read
  - groups.read
  - team_info.read
  - team_data.content.read
- Ha apuntado la clave de la aplicación de Dropbox, el secreto de la aplicación de Dropbox y el token de acceso a Dropbox como credenciales de autenticación básica.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Has configurado y copiado un token de acceso temporal de Oauth 2.0 para tu aplicación de Dropbox. Este token es temporal y caduca a las 4 horas. Consulte la Documentación de Dropbox sobre la autenticación OAuth.

> **Note**
>
> Se recomienda crear un token de acceso actualizado de Dropbox que no caduque nunca, en lugar de utilizar un token de acceso único que caduca a las 4 horas. Un token de acceso actualizado es permanente y nunca caduca, por lo que se podrá seguir sincronizando el origen de datos en el futuro.

- Recomendado: has configurado un token de actualización permanente de Dropbox que nunca caduque Amazon Kendra para poder seguir sincronizando tu fuente de datos sin interrupciones. Consulte la Documentación de Dropbox sobre los tokens de actualización.
- Ha comprobado que cada documento es único en Dropbox y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el tuyo Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Dropbox en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes un IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar tu fuente de datos de Dropbox Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarte Amazon Kendra a tu fuente de datos de Dropbox, debes proporcionar los detalles necesarios de tu fuente de datos de Dropbox para que Amazon Kendra puedas acceder a tus datos. Si aún no has configurado Dropbox para Amazon Kendra, consultaRequisitos previos.

Console

Para conectarte Amazon Kendra a Dropbox

1. Inicia sesión en la Amazon Kendra consola AWS Management Console y ábrela.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector de Dropbox y, a continuación, selecciona Añadir conector. Si utilizas la versión 2 (si corresponde), elige el conector de Dropbox con la etiqueta «V2.0».
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elige un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

b. Tipo de token de autenticación: elija un token permanente (recomendado) o un token de acceso temporal.

c. AWS Secrets Manager secreto: elige un secreto existente o crea uno nuevo Secrets Manager para almacenar tus credenciales de autenticación de Dropbox. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

A. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Dropbox» se añade automáticamente a tu nombre secreto.

B. Para la información sobre la clave, el secreto de la aplicación y el token (permanente o temporal), introduce los valores de las credenciales de autenticación configurados en Dropbox.

ii. Guarda y añade tu secreto.

d. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

e. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para

sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

f. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

g. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Para seleccionar entidades o tipos de contenido: elige las entidades o los tipos de contenido de Dropbox que quieras rastrear.

   b. En Configuración adicional para Patrones regex: agregue patrones de expresiones regulares para incluir o excluir determinados archivos.

   c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Cuando sincronizas tu fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

      - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
      - Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Plantillas de archivos, Dropbox Paper y Dropbox Paper: selecciona uno de los campos de fuentes de datos predeterminados Amazon Kendra generados que quieras asignar a tu índice.

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarte Amazon Kendra a Dropbox

Debes especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `DROPBOX` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no

seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:

- `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Tipo de token de acceso: especifique si desea utilizar un token de acceso permanente o temporal para el AWS Secrets Manager secreto que almacena las credenciales de autenticación.

  > **Note**
  >
  > Se recomienda crear un token de acceso actualizado que no caduque nunca en Dropbox, en lugar de utilizar un token de acceso único que caduca a las 4 horas. Debe crear una aplicación y un token de acceso de actualización en la consola para desarrolladores de Dropbox y proporcionar el token de acceso en su secreto.

- Nombre secreto del recurso de Amazon (ARN): proporciona el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de tu cuenta de Dropbox. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "appKey": "Dropbox app key",
    "appSecret": "Dropbox app secret",
    "accesstoken": "temporary access token or refresh access token"
}
```

- Rastreador de identidad: especifica si deseas activar el rastreador de identidad. Amazon Kendra El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide

utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

- IAM rol: especifica `RoleArn` cuándo llamas `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a tu Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Dropbox y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Dropbox.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Tipos de documentos o contenido: especifica si deseas rastrear los archivos de tu Dropbox, los documentos de Dropbox Paper, las plantillas de Dropbox Paper y los atajos de páginas web almacenados en tu Dropbox.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Lista de control de acceso (ACL): especifica si deseas rastrear la información de la ACL de tus documentos, si tienes una ACL y quieres usarla para controlar el acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

- Asignaciones de campos: elija asignar los campos del origen de datos de Dropbox a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para ver una lista de otras claves JSON importantes que debes configurar, consulta el esquema de plantillas de Dropbox.

Más información

Para obtener más información acerca de la integración de Amazon Kendra con el origen de datos de Dropbox, consulte:

- Indexar el contenido de Dropbox mediante el conector de Dropbox para Amazon Kendra

Drupal

Drupal es un sistema de administración de contenidos (CMS) de código abierto que se puede utilizar para crear sitios web y aplicaciones web. Puedes usarlo Amazon Kendra para indexar lo siguiente en Drupal:

- Contenido: artículos, páginas básicas, bloques básicos, tipos de contenido definidos por el usuario, tipos de bloques definidos por el usuario, tipos de contenido personalizados, tipos de bloques personalizados
- Comentario: para cualquier tipo de contenido y tipo de bloque
- Adjuntos: para cualquier tipo de contenido y tipo de bloque

Puedes conectarte Amazon Kendra a tu fuente de datos de Drupal mediante la Amazon Kendra consola o la TemplateConfigurationAPI.

Para solucionar problemas de su conector de fuente de datos de Amazon Kendra Drupal, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

Amazon Kendra El conector de fuente de datos de Drupal admite las siguientes funciones:

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Drupal, realiza estos cambios en tu Drupal y en tus cuentas. AWS

En Drupal, asegúrese de que:

- Ha creado una cuenta de Drupal (Standard) Suite y un usuario con un rol de administrador.
- Ha copiado el nombre de su sitio de Drupal y configurado una URL de host. Por ejemplo, *https://<hostname>/<drupalsitename>*.
- Ha configurado credenciales de autenticación básicas que contienen un nombre de usuario (nombre de usuario de inicio de sesión en el sitio web de Drupal) y una contraseña (contraseña del sitio web de Drupal).
- Recomendado: ha configurado un token de credenciales OAuth 2.0. Use este token junto con la contraseña de Drupal, el ID de cliente, el secreto de cliente, el nombre de usuario (nombre de

usuario de inicio de sesión en el sitio web de Drupal) y la contraseña (contraseña del sitio web de Drupal) para conectarse a Amazon Kendra.

- Ha añadido los siguientes permisos a la cuenta de Drupal utilizando un rol de administrador:
  - administer blocks
  - administer block_content display
  - administer block_content fields
  - administer block_content form display
  - administer views
  - view user email addresses
  - view own unpublished content
  - view page revisions
  - view article revisions
  - view all revisions
  - view the administration theme
  - access content
  - access content overview
  - access comments
  - search content
  - access files overview
  - access contextual links

> **Note**
>
> Si hay tipos de contenido definidos por el usuario o tipos de bloques definidos por el usuario, o si se añaden vistas y bloques al sitio web de Drupal, se les debe proporcionar acceso de administrador.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Drupal en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes un IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar tu fuente de datos de Drupal. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Drupal, debe proporcionar los detalles de sus credenciales de Drupal para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Drupal, consulte. Amazon Kendra Requisitos previos

Console

Para conectarse a Amazon Kendra Drupal

1. Inicia sesión en la Amazon Kendra consola AWS Management Console y ábrela.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector de Drupal y, a continuación, selecciona Añadir conector. Si usa la versión 2 (si corresponde), elija el conector de Drupal con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

    a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

    b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

    c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

    d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

    e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:

    a. En Origen, en URL de host: la URL de host del sitio de Drupal. Por ejemplo, *https://<hostname>/<drupalsitename>*.

    b. En Ubicación del certificado SSL: introduzca la ruta al certificado SSL almacenado en el bucket de Amazon S3 .

    c. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

    d. En Autenticación: elija entre la Autenticación básica y la Autenticación OAuth 2.0 según el caso de uso.

    e. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo para almacenar sus Secrets Manager credenciales de autenticación de Drupal. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

        i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

A. Si ha elegido la Autenticación básica, introduzca un Nombre del secreto, el Nombre de usuario (nombre de usuario del sitio de Drupal) y la Contraseña (contraseña del sitio de Drupal) que ha copiado y seleccione Guardar y agregar secreto.

B. Si ha elegido la Autenticación OAuth 2.0, introduzca un Nombre del secreto, Nombre de usuario (nombre de usuario del sitio de Drupal), Contraseña (contraseña del sitio de Drupal), ID de cliente y Secreto de cliente generados en la cuenta de Drupal y seleccione Guardar y agregar secreto.

ii. Seleccione Guardar.

f. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

g. Rastreador de identidad: especifique si desea activar el rastreador de identidad. Amazon Kendra El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

h. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

> **Note**
>
> Si elige rastrear Artículos, Páginas básicas y Bloques básicos, sus campos predeterminados se sincronizarán automáticamente. También puede optar por sincronizar los comentarios, archivos adjuntos, campos personalizados y otras entidades personalizadas.

- En Entidades seleccionadas:
  - Artículos: elija si desea rastrear los Artículos, sus Comentarios y sus Archivos adjuntos.
  - Páginas básicas: elija si desea rastrear las Páginas básicas, sus Comentarios y sus Archivos adjuntos.
  - Bloques básicos: elija si desea rastrear los Bloques básicos, sus Comentarios y sus Archivos adjuntos.
  - También puede optar por añadir Tipos de contenido personalizados y Bloques personalizados.

b. En Configuración adicional (opcional):

- En Patrón regex: agregue patrones de expresiones regulares para incluir o excluir títulos de entidades y nombres de archivos específicos. Puede agregar hasta 100 patrones.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los

cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Para el contenido, los comentarios y los archivos adjuntos: seleccione uno de los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Drupal

Debe especificar un JSON del esquema del origen de datos mediante la API TemplateConfiguration. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `DRUPAL` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.

- `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto del recurso de Amazon (ARN): proporciona el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creaste en tu cuenta de Drupal.

  Si utiliza la autenticación básica, el secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  Si utiliza la autenticación OAuth 2.0, el secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "username": "user name",
      "password": "password",
      "clientId": "client id",
      "clientSecret": "client secret"
  }
  ```

  Note

  Note

  Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos

secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Drupal y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Drupal.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido, comentarios y archivos adjuntos. También puede especificar patrones de expresiones regulares para incluir o excluir contenido, comentarios y archivos adjuntos.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Rastreador de identidad: especifique si se debe activar el rastreador de identidad. Amazon Kendra El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

- Asignaciones de campos: elija asignar los campos del origen de datos de Drupal a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Esquema de plantilla de Drupal.

Notas

- Las API de Drupal no tienen límites de limitación oficiales.
- Los SDK de Java no están disponibles para Drupal.
- Los datos de Drupal solo se pueden obtener mediante API JSON nativas.
- No se pueden rastrear los tipos de contenido que no estén asociados a ninguna Vista de Drupal.
- Necesita acceso de administrador para rastrear los datos de los Bloques de Drupal.
- No hay ninguna API JSON disponible para crear el tipo de contenido definido por el usuario mediante verbos HTTP.
- El cuerpo del documento y los comentarios de los Artículos, las Páginas básicas, los Bloques básicos, el tipo de contenido definido por el usuario y el tipo de bloque definido por el usuario se muestran en formato HTML. Si el contenido HTML no está bien formado, las etiquetas relacionadas con el HTML aparecerán en el cuerpo del documento y en los comentarios y estarán visibles en los resultados de búsqueda de Amazon Kendra .
- No se incorporarán los tipos de contenido ni los tipos de bloques sin descripción ni cuerpo. Amazon Kendra Solo los comentarios y archivos adjuntos de este tipo de contenido o tipo de bloque se incorporarán a tu Amazon Kendra índice.

GitHub

GitHub es un servicio de alojamiento web para el desarrollo de software que proporciona servicios de almacenamiento y administración de códigos con control de versiones. Puede utilizarlos Amazon Kendra para indexar los archivos de repositorio de GitHub Enterprise Cloud (SaaS) y GitHub Enterprise Server (On Prem), las solicitudes de emisión y extracción, los comentarios de las solicitudes de emisión y extracción y los archivos adjuntos de comentarios de las solicitudes de emisión y extracción. También se puede optar por incluir o excluir determinados archivos.

> **Note**
>
> Amazon Kendra ahora es compatible con un conector actualizado GitHub .
> La consola se ha actualizado automáticamente. Todos los conectores nuevos que cree en la consola utilizarán la arquitectura actualizada. Si usa la API, ahora debe usar el TemplateConfigurationobjeto en lugar del `GitHubConfiguration` objeto para configurar el conector.
> Los conectores configurados con la antigua arquitectura de consola y API seguirán funcionando tal y como estaban configurados. Sin embargo, no podrá editarlos ni actualizarlos. Si desea editar o actualizar la configuración del conector, debe crear uno nuevo.
> Se recomienda migrar el flujo de trabajo del conector a la versión actualizada. Está previsto que el soporte para los conectores configurados con la arquitectura anterior finalice en junio de 2024.

Puede conectarse Amazon Kendra a su fuente GitHub de datos mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de la fuente de Amazon Kendra GitHub datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra GitHub el conector de fuente de datos admite las siguientes funciones:

- Asignaciones de campo
- control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de GitHub datos, realice estos cambios en sus AWS cuentas GitHub y.

En GitHub, asegúrate de tener:

- Creó un GitHub usuario con permisos administrativos para la GitHub organización.
- Configuraste un token de acceso personal en Git Hub para usarlo como credenciales de autenticación. Consulta GitHub la documentación sobre cómo crear un token de acceso personal.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Recomendado: configuré un token de OAuth para las credenciales de autenticación. Use el token de OAuth para mejorar los límites de limitación de la API y el rendimiento del conector. Consulta la GitHub documentación sobre la autorización de OAuth.
- Apuntaste la URL del GitHub host del tipo de GitHub servicio que utilizas. Por ejemplo, la URL del host de la GitHub nube podría ser *`https://api.github.com`* y la URL del host GitHub del servidor podría ser *`https://on-prem-host-url/api/v3/`*.
- Apuntó el nombre de su organización para GitHub la cuenta de GitHub Enterprise Cloud (SaaS) o la cuenta de GitHub Enterprise Server (local) a la que desea conectarse. Para encontrar el nombre

de su organización, inicie sesión en el GitHub escritorio y seleccione Sus organizaciones en el menú desplegable de su imagen de perfil.

- Opcional (solo para servidores): generó un certificado SSL y copió la ruta al certificado almacenado en un Amazon S3 depósito. Utilízalo para conectarte GitHub si necesitas una conexión SSL segura. Puede generar simplemente un certificado autofirmado X509 en cualquier ordenador mediante OpenSSL. Para ver un ejemplo del uso de OpenSSL para crear un certificado X509, consulte Crear y firmar un certificado X509.
- Se han añadido los siguientes permisos:

  Para la nube GitHub empresarial (SaaS)

  - `repo:status`— Otorga acceso de lectura y escritura a los estados de confirmación en repositorios públicos y privados. Este ámbito solo es necesario para conceder a otros usuarios o servicios el acceso a los estados de confirmación de los repositorios privados sin conceder acceso al código.
  - `repo_deployment`— Otorga acceso a los estados de despliegue de los repositorios públicos y privados. Este ámbito solo es necesario para conceder a otros usuarios o servicios el acceso a los estados de despliegue, sin conceder acceso al código.
  - `public_repo`— Limita el acceso a los repositorios públicos. Esto incluye el acceso de lectura y escritura al código, los estados de confirmación, los proyectos de repositorios, los colaboradores y los estados de despliegue de repositorios y organizaciones públicos. También es obligatorio para los repositorios públicos destacados.
  - `repo:invite`— Otorga la capacidad de aceptar o rechazar invitaciones a colaborar en un repositorio. Este alcance solo es necesario para permitir que otros usuarios o servicios accedan a las invitaciones sin conceder acceso al código.
  - `security_events`— Otorga: acceso de lectura y escritura a los eventos de seguridad en la API de escaneo de código. Este alcance solo es necesario para conceder a otros usuarios o servicios el acceso a los eventos de seguridad sin conceder acceso al código.
  - `read:org`— Acceso de solo lectura a la membresía de la organización, los proyectos de la organización y la membresía del equipo.
  - `user:email`— Otorga acceso de lectura a las direcciones de correo electrónico de los usuarios. Amazon Kendra lo requiere para rastrear las ACL.
  - `user:follow`— Otorga acceso para seguir o dejar de seguir a otros usuarios. Amazon Kendra lo requiere para rastrear las ACL.
  - `read:user`— Otorga acceso para leer los datos del perfil de un usuario. Amazon Kendra lo requiere para rastrear las ACL.

- `workflow`— Otorga la posibilidad de añadir y actualizar los archivos de flujo de trabajo de GitHub Actions. Los archivos de flujo de trabajo se pueden archivar sin este ámbito si el mismo archivo (con la misma ruta y el mismo contenido) existe en otra rama del mismo repositorio.

Para obtener más información, consulta los ámbitos de las aplicaciones OAuth en Docs. GitHub

Para GitHub Enterprise Server (local)

- `repo:status`— Otorga acceso de lectura y escritura a los estados de confirmación en repositorios públicos y privados. Este ámbito solo es necesario para conceder a otros usuarios o servicios el acceso a los estados de confirmación de los repositorios privados sin conceder acceso al código.
- `repo_deployment`— Otorga acceso a los estados de despliegue de los repositorios públicos y privados. Este ámbito solo es necesario para conceder a otros usuarios o servicios el acceso a los estados de despliegue, sin conceder acceso al código.
- `public_repo`— Limita el acceso a los repositorios públicos. Esto incluye el acceso de lectura y escritura al código, los estados de confirmación, los proyectos de repositorios, los colaboradores y los estados de despliegue de repositorios y organizaciones públicos. También es obligatorio para los repositorios públicos destacados.
- `repo:invite`— Otorga la capacidad de aceptar o rechazar invitaciones a colaborar en un repositorio. Este alcance solo es necesario para permitir que otros usuarios o servicios accedan a las invitaciones sin conceder acceso al código.
- `security_events`— Otorga: acceso de lectura y escritura a los eventos de seguridad en la API de escaneo de código. Este alcance solo es necesario para conceder a otros usuarios o servicios el acceso a los eventos de seguridad sin conceder acceso al código.
- `read:user`— Otorga acceso para leer los datos del perfil de un usuario. Amazon Q Business lo exige para rastrear las ACL.
- `user:email`— Otorga acceso de lectura a las direcciones de correo electrónico de los usuarios. Amazon Q Business lo exige para rastrear las ACL.
- `user:follow`— Otorga acceso para seguir o dejar de seguir a otros usuarios. Amazon Q Business lo exige para rastrear las ACL.
- `site_admin`— Otorga a los administradores del sitio acceso a los puntos finales de la API de administración de servidores GitHub empresariales.
- `workflow`— Otorga la posibilidad de añadir y actualizar los archivos de flujo de trabajo de GitHub Actions. Los archivos de flujo de trabajo se pueden archivar sin este ámbito si el mismo archivo (con la misma ruta y el mismo contenido) existe en otra rama del mismo repositorio.

Para obtener más información, consulta los ámbitos de las aplicaciones OAuth en GitHub Docs y Cómo entender los ámbitos de las aplicaciones OAuth en Developer. GitHub

- Has comprobado que cada documento es único en las demás fuentes de datos que vayas a utilizar para el mismo índice GitHub y entre ellas. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Guardó sus credenciales de GitHub autenticación en un AWS Secrets Manager secreto y, si usa la API, anotó el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de GitHub datos. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de GitHub datos, debe proporcionar los detalles necesarios de la fuente de GitHub datos para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configurado GitHub Amazon Kendra, consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a GitHub

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el GitHub conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el GitHub conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. GitHubfuente: elija entre GitHub Enterprise Cloud y GitHubEnterprise Server.
   b. GitHub URL del host: por ejemplo, la URL del host de la GitHub nube podría ser *https://api.github.com* y la URL del host del GitHub servidor podría ser *https://on-prem-host-url/api/v3/*.
   c. GitHub nombre de la organización: introduzca el nombre de su organización. GitHub Puedes encontrar la información de tu organización en tu GitHub cuenta.

d. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

e. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de GitHub autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

   i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

      A. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-GitHub -' se añade automáticamente a tu nombre secreto.

      B. Para el GitHubtoken: introduzca el valor de la credencial de autenticación configurado en. GitHub

   ii. Guarde y añada su secreto.

f. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

g. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

h. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Seleccionar repositorios: elija rastrear todos los repositorios o selecciónelos.

      Si opta por rastrear los repositorios seleccionados, añada los nombres de los repositorios y, si lo desea, el nombre de cualquier rama específica.

   b. Tipos de contenido: elige los tipos de contenido que quieres rastrear, entre los archivos, las publicaciones, las solicitudes de incorporación de datos y mucho más.

   c. Patrones regex: añada patrones de expresiones regulares para incluir o excluir determinados archivos.

   d. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Cuando sincronizas tu fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no eliges la sincronización completa como opción de modo de sincronización.

      - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
      - Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
      - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

   e. Calendario de ejecución sincronizado para la frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

f. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra GitHub

Debe especificar un JSON del esquema del origen de datos mediante la API TemplateConfiguration. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `GITHUB` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- GitHubtipo: especifique el tipo como `SAAS` o`ON_PREMISE`.
- URL del host: especifique la URL del GitHub host o la URL del punto final de la API. Por ejemplo, si utiliza GitHub SaaS/Enterprise Cloud, la URL del host podría ser, y en el caso de los servidores GitHub locales o empresarialeshttps://api.github.com, la URL del host podría ser. `https://on-prem-host-url/api/v3/`
- Nombre de la organización: especifique el nombre de la organización de la cuenta. GitHub Para encontrar el nombre de su organización, inicie sesión en la GitHub computadora y seleccione Sus organizaciones en el menú desplegable de su imagen de perfil.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no

eliges la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:

- `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.
- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta. GitHub El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "personalToken": "token"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. GitHub Amazon Kendra Para obtener más información, consulte las IAM funciones de las fuentes GitHub de datos.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.

  > **Note**
  >
  > Si usa un GitHub servidor, debe usar un Amazon VPC para conectarse a su GitHub servidor.

- Filtro de repositorios: filtra los repositorios por su nombre y nombre de rama.
- Tipos de documentos o contenido: especifique si desea rastrear los documentos del repositorio, las ediciones, los comentarios de las publicaciones, los archivos adjuntos a los comentarios de las publicaciones, las solicitudes de extracción, los comentarios de las solicitudes de extracción o los archivos adjuntos de los comentarios de las solicitudes de extracción.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos y carpetas.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Lista de control de acceso (ACL): especifique si desea rastrear la información de la ACL de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos de la fuente de GitHub datos a los campos de índice. Amazon Kendra Puede incluir campos de documentos, confirmaciones, emisiones, archivos adjuntos de publicaciones, comentarios de publicaciones, solicitudes de extracción, archivos adjuntos de solicitudes de extracción y comentarios de solicitudes de extracción. Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio para que Amazon Kendra pueda buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre `_document_body` del campo de índice. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el Esquema de plantilla de GitHub.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de GitHub datos, consulte:

- Reimagine la búsqueda en GitHub los repositorios con la potencia del conector Amazon Kendra GitHub

Gmail

Gmail es un cliente de correo desarrollado por Google a través del cual se pueden enviar mensajes de correo electrónico con archivos adjuntos. Los mensajes de Gmail se pueden ordenar y almacenar en la bandeja de entrada del correo electrónico mediante carpetas y etiquetas. Puedes usarlo Amazon Kendra para indexar tus mensajes de correo electrónico y sus archivos adjuntos. También puede configurarlo Amazon Kendra para incluir o excluir mensajes de correo electrónico, archivos adjuntos de mensajes y etiquetas específicos para su indexación.

Puedes conectarte Amazon Kendra a tu fuente de datos de Gmail mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuentes de datos de Amazon Kendra Gmail, consultaSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos

- Instrucciones de conexión
- Más información
- Notas

Características admitidas

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Gmail, realiza estos cambios en Gmail y en tus AWS cuentas.

En Gmail, asegúrese de que:

- Ha creado una cuenta de administrador de Google Cloud Platform y un proyecto de Google Cloud.
- Ha activado la API de Gmail y la API del SDK de administración en su cuenta de administrador.
- Ha creado una cuenta de servicio y descargado una clave privada JSON para la cuenta de Gmail. Para obtener información sobre cómo crear una clave privada y acceder a ella, consulte la documentación de Google Cloud sobre cómo Crear una clave de cuenta de servicio y las Credenciales de una cuenta de servicio.
- Copiaste el correo electrónico de tu cuenta de administrador, el correo de tu cuenta de servicio y tu clave privada para utilizarlos como credenciales de autenticación.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha añadido los siguientes ámbitos de OAuth (con un rol de administrador) para su usuario y los directorios compartidos que quiere indexar:
  - https://www.googleapis.com/auth/admin.directory.user.readonly
  - https://www.googleapis.com/auth/gmail.readonly
- Ha comprobado que cada documento es único en Gmail y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Gmail en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes ningún IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar tu fuente de datos de Gmail Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarte Amazon Kendra a tu fuente de datos de Gmail, debes proporcionar los detalles de tus credenciales de Gmail para que Amazon Kendra puedas acceder a tus datos. Si aún no has configurado Gmail para Amazon Kendra, consultaRequisitos previos.

Console

Para conectarte Amazon Kendra a Gmail

1. Inicia sesión en la Amazon Kendra consola AWS Management Console y ábrela.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector de Gmail y, a continuación, selecciona Añadir conector. Si utilizas la versión 2 (si corresponde), selecciona el conector de Gmail con la etiqueta «V2.0».
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elige un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

b. En Autenticación por AWS Secrets Manager secreto: elige un secreto existente o crea uno nuevo Secrets Manager para almacenar tus credenciales de autenticación de Gmail. Si decides crear un secreto nuevo, se abrirá una ventana AWS Secrets Manager secreta.

   - Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

     A. Nombre del secreto: un nombre para su secreto.

     B. Correo electrónico del cliente: el correo electrónico del cliente que copió de su cuenta de servicio de Google.

     C. Correo electrónico de la cuenta de administrador: el correo electrónico de la cuenta de administrador que quiere usar.

     D. Clave privada: la clave privada que copió de su cuenta de servicio de Google.

     E. Guarda y añade tu secreto.

c. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

d. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

e. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Para los tipos de entidad: elija sincronizar los archivos adjuntos de los mensajes.

b. (Opcional) En Configuración adicional, introduzca la siguiente información:

   i. Intervalo de fechas: introduce un intervalo de fechas para especificar la fecha de inicio y finalización de los correos electrónicos que deseas rastrear.

   ii. Dominios de correo electrónico: incluye o excluye determinados correos electrónicos según los dominios de correo «para», «desde», «cc» y «bcc».

   iii. Palabras clave en los asuntos: incluya o excluya los correos electrónicos en función de las palabras clave de sus asuntos de correo electrónico.

> **Note**
>
> También puede optar por incluir cualquier documento que coincida con todas las palabras clave del asunto que haya introducido.

   iv. Etiquetas: añada patrones de expresiones regulares para incluir o excluir determinadas etiquetas de correo electrónico.

   v. Archivos adjuntos: añada patrones de expresiones regulares para incluir o excluir determinados archivos adjuntos de correo electrónico.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no eliges la sincronización completa como opción de modo de sincronización.

   - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
   - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

> **Important**
>
> Como no hay una API para actualizar los mensajes de Gmail eliminados permanentemente, el contenido nuevo, modificado o eliminado se sincroniza:

- No eliminará de tu Amazon Kendra índice los mensajes que se hayan eliminado permanentemente de Gmail
- No sincronizará los cambios en las etiquetas de correo de Gmail

Para sincronizar los cambios en las etiquetas del origen de datos de Gmail y los mensajes de correo electrónico eliminados permanentemente con el índice de Amazon Kendra , debe realizar rastreos completos de forma periódica.

d. Calendario de ejecución sincronizado, para Frecuencia: elige la frecuencia con la que deseas sincronizar el contenido de la fuente de datos y actualizar el índice.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

> **Note**
>
> Amazon Kendra El conector de fuentes de datos de Gmail no admite la creación de campos de índice personalizados debido a las limitaciones de la API.

b. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Gmail

Debe especificar un JSON del esquema del origen de datos mediante la API TemplateConfiguration. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `GMAIL` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no eliges la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

  > **Important**
  >
  > Como no hay una API para actualizar los mensajes de Gmail eliminados permanentemente, el contenido nuevo, modificado o eliminado se sincroniza:
  >
  > - No eliminará de tu Amazon Kendra índice los mensajes que se hayan eliminado permanentemente de Gmail
  > - No sincronizará los cambios en las etiquetas de correo de Gmail
  >
  > Para sincronizar los cambios en la etiqueta de la fuente de datos de Gmail y los mensajes de correo electrónico eliminados permanentemente con tu Amazon Kendra índice, debes realizar rastreos completos de forma periódica.

- Nombre secreto del recurso de Amazon (ARN): proporciona el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de tu cuenta de Gmail. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "adminAccountEmailId": "service account email",
    "clientEmailId": "user account email",
    "privateKey": "private key"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Gmail y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Gmail.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: especifica si deseas incluir o excluir determinados correos electrónicos «para», «de», «cc» o «bcc».

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Filtrado contextual de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos del origen de datos de Gmail a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

> **Note**
>
> Amazon Kendra El conector de fuentes de datos de Gmail no admite la creación de campos de índice personalizados debido a las limitaciones de la API.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el Esquema de plantilla de Gmail.

Más información

Para obtener más información sobre la integración Amazon Kendra con tu fuente de datos de Gmail, consulta:

- Realizar una búsqueda inteligente en los correos electrónicos de Google Workspace mediante el conector de Gmail para Amazon Kendra.

Notas

- Como no hay una API para actualizar los mensajes de Gmail eliminados permanentemente, una `FULL_CRAWL`/Sincronización de contenido nuevo, modificado o eliminado:
  - No eliminará de tu Amazon Kendra índice los mensajes que se hayan eliminado permanentemente de Gmail
  - No sincronizarán los cambios en las etiquetas de correo de Gmail

  Para sincronizar los cambios en la etiqueta de la fuente de datos de Gmail y los mensajes de correo electrónico eliminados permanentemente con tu Amazon Kendra índice, debes realizar un rastreo completo de forma periódica.
- Amazon Kendra El conector de fuentes de datos de Gmail no admite la creación de campos de índice personalizados debido a las limitaciones de la API.

Google Drive

Google Drive es un servicio de almacenamiento de archivos basado en la nube. Amazon Kendra se puede utilizar para indexar los documentos almacenados en las carpetas de unidades compartidas,

Mis unidades y Compartido conmigo del origen de datos de Google Drive. Se pueden indexar tanto los documentos de Google Workspace como los documentos que aparecen en Tipos de documentación. También se pueden usar filtros de inclusión y exclusión para indexar el contenido por nombre de archivo, tipo de archivo y ruta de archivo.

Puedes conectarte Amazon Kendra a tu fuente de datos de Google Drive mediante la Amazon Kendra consola, la TemplateConfigurationAPI o la GoogleDriveConfigurationAPI.

Amazon Kendra tiene dos versiones del conector de Google Drive. Las características compatibles de cada versión incluyen:

Conector de Google Drive V1.0/API GoogleDriveConfiguration

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión

Conector TemplateConfigurationV2.0/API para Google Drive

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

> **Note**
>
> Está previsto que el soporte para el conector V1.0 y la DriveConfiguration API de Google Drive finalice en 2023. Recomendamos migrar o utilizar el conector V2.0 o la API de Google Drive. TemplateConfiguration

Para solucionar problemas del conector de fuentes de datos de Amazon Kendra Google Drive, consulta. Solución de problemas con los orígenes de datos

Temas

- Google Drive Connector V1.0
- Google Drive Connector V2.0

Google Drive Connector V1.0

Google Drive es un servicio de almacenamiento de archivos basado en la nube. Puedes usarlo Amazon Kendra para indexar documentos y comentarios almacenados en las carpetas de unidades compartidas, Mis unidades de disco y Compartidas conmigo de tu fuente de datos de Google Drive. Se pueden indexar tanto los documentos de Google Workspace como los documentos que aparecen en Tipos de documentación. También se pueden usar filtros de inclusión y exclusión para indexar el contenido por nombre de archivo, tipo de archivo y ruta de archivo.

> **Note**
>
> Está previsto que el soporte para el conector V1.0 y la DriveConfiguration API de Google Drive finalice en 2023. Recomendamos migrar o utilizar el conector V2.0 o la API de Google Drive. TemplateConfiguration

Para solucionar problemas del conector de fuentes de datos de Amazon Kendra Google Drive, consulta. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Google Drive, realiza estos cambios en tu Google Drive y en tus AWS cuentas.

En Google Drive, asegúrese de que:

- Bien se le ha concedido el acceso mediante un rol de superadministrador o es un usuario con privilegios administrativos. No necesita un rol de superadministrador si este le ha otorgado el acceso.
- Ha creado una cuenta de servicio con la opción Habilitar la delegación en todo el dominio de G Suite activada y una clave JSON como clave privada que utiliza la cuenta.
- Ha copiado el correo electrónico de su cuenta de usuario y el correo electrónico de su cuenta de servicio. Cuando te conectes, introduce el correo electrónico de tu cuenta de usuario como correo electrónico de cuenta de administrador y el correo de tu cuenta de servicio como correo electrónico de cliente en tu AWS Secrets Manager secreto. Amazon Kendra

  > **Note**
  >
  > Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha agregado la API del SDK de administración y la API de Google Drive a su cuenta.
- Ha agregado (o has pedido a un usuario con un rol de superadministrador que agregue) los siguientes permisos a su cuenta de servicio mediante un rol de superadministrador:
  - https://www.googleapis.com/auth/drive.readonly
  - https://www.googleapis.com/auth/drive.metadata.readonly
  - https://www.googleapis.com/auth/admin.directory.user.readonly
  - https://www.googleapis.com/auth/admin.directory.group.readonly
- Ha comprobado que cada documento es único en Google Drive y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Google Drive en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes ningún IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar tu fuente de datos de Google Drive Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarte Amazon Kendra a tu fuente de datos de Google Drive, debes proporcionar los detalles necesarios de tu fuente de datos de Google Drive para que Amazon Kendra puedas acceder a tus datos. Si aún no has configurado Google Drive para Amazon Kendra verRequisitos previos.

Console

Para conectarse Amazon Kendra a Google Drive

1. Inicie sesión en la consola AWS de administración y abra la Amazon Kendra consola.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar origen de datos, seleccione Google Drive Connector V1.0 y, a continuación, seleccione Añadir conector.
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. En Tipo de autenticación: elija entre Existente y Nuevo. Si elige usar un secreto existente, use Seleccionar secreto para elegir el secreto.

   b. Si decide crear un secreto nuevo, se abrirá una opción de secreto de AWS Secrets Manager .

      - Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

        A. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Google Drive» se añade automáticamente a tu nombre secreto.

        B. En Correo electrónico de la cuenta de administrador, Correo electrónico del cliente y Clave privada: introduzca los valores de las credenciales de autenticación que generó y descargó de su cuenta de Google Drive.

        C. Seleccione Guardar autenticación.

c. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales de tu repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

d. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Excluir cuentas de usuario: los usuarios de Google Drive que quiere excluir del índice. Puede agregar hasta 100 cuentas de usuario.

   b. Excluir unidades compartidas: las unidades compartidas de Google Drive que quiere excluir del índice. Puede agregar hasta 100 unidades compartidas.

   c. Excluir tipos de archivos: los tipos de archivos de Google Drive que quiere excluir del índice. También puede optar por editar las selecciones de tipo MIME.

   d. Configuración adicional: patrones de expresión regular para incluir o excluir determinado contenido. Puede agregar hasta 100 patrones.

   e. Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

   f. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Para GoogleDrive el nombre de campo y otras asignaciones de campos sugeridas: seleccione entre los campos de fuentes de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Google Drive Amazon Kendra

Debe especificar lo siguiente mediante la GoogleDriveConfigurationAPI:

- Nombre secreto del recurso de Amazon (ARN): proporciona el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de tu cuenta de Google Drive. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "clientAccount": "service account email",
    "adminAccount": "user account email"",
    "privateKey": "private key"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Google Drive y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Google Drive.

También puede añadir las siguientes características opcionales:

- Filtros de inclusión y exclusión: de forma predeterminada, Amazon Kendra indexa todos los documentos de Google Drive. Puede especificar si desea incluir o excluir determinado contenido en las unidades compartidas, las cuentas de usuario, los tipos MIME de documentos y los archivos. Si decide excluir las cuentas de usuario, no se indexará ninguno de los archivos de Mi unidad correspondiente a la cuenta. Los archivos compartidos con el usuario se indexan, a menos que también se excluya al propietario del archivo.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Asignaciones de campos: elija asignar los campos del origen de datos de Google Drive a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

Más información

Para obtener más información sobre la integración Amazon Kendra con tu fuente de datos de Google Drive, consulta:

- Cómo empezar a utilizar el conector de Amazon Kendra Google Drive

Google Drive Connector V2.0

Google Drive es un servicio de almacenamiento de archivos basado en la nube. Puedes usarlo Amazon Kendra para indexar los documentos y comentarios almacenados en las carpetas de unidades compartidas, Mis unidades de disco y Compartidas conmigo de tu fuente de datos de Google Drive. Se pueden indexar tanto los documentos de Google Workspace como los documentos que aparecen en Tipos de documentación. También se pueden usar filtros de inclusión y exclusión para indexar el contenido por nombre de archivo, tipo de archivo y ruta de archivo.

> **Note**
>
> Está previsto que el soporte para el conector V1.0 y la DriveConfiguration API de Google Drive finalice en 2023. Recomendamos migrar o utilizar el conector V2.0 o la API de Google Drive. TemplateConfiguration

Para solucionar problemas del conector de fuentes de datos de Amazon Kendra Google Drive, consulta. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Google Drive, realiza estos cambios en tu Google Drive y en tus AWS cuentas.

En Google Drive, asegúrese de que:

- Bien se le ha concedido el acceso mediante un rol de superadministrador o es un usuario con privilegios administrativos. No necesita un rol de superadministrador si este le ha otorgado el acceso.
- Ha configurado las credenciales de conexión de la cuenta de servicio de Google Drive, que incluyen el correo electrónico de la cuenta de administrador, el correo electrónico del cliente (correo electrónico de la cuenta de servicio) y la clave privada. Consulte la Documentación de Google Cloud sobre cómo crear y eliminar las claves de las cuentas de servicio.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Creó una cuenta de servicio de Google Cloud (una cuenta con autoridad delegada para asumir una identidad de usuario) con la opción Habilitar la delegación en todo el dominio de G Suite para la server-to-server autenticación y, a continuación, generó una clave privada JSON con la cuenta.

> **Note**
>
> La clave privada se debe generar después de crear la cuenta de servicio.

- Ha agregado la API del SDK de administración y la API de Google Drive a su cuenta de usuario.
- Opcional: ha configurado las credenciales de conexión OAuth 2.0 de Google Drive que contienen el ID de cliente, el secreto del cliente y el token de actualización como credenciales de conexión para un usuario específico. Las necesita para rastrear los datos de las cuentas individuales. Consulte la Documentación de Google sobre el uso de OAuth 2.0 para acceder a las API.
- Ha agregado (o has pedido a un usuario con un rol de superadministrador que agregue) los siguientes ámbitos de OAuth a su cuenta de servicio mediante un rol de superadministrador. Estos ámbitos de API son necesarios para rastrear todos los documentos y la información de control de acceso (ACL) de todos los usuarios de un dominio de Google Workspace:
  - https://www.googleapis.com/auth/drive.readonly: visualice y descargue todos los archivos de Google Drive
  - https://www.googleapis.com/auth/drive.metadata.readonly: visualice los metadatos de los archivos de Google Drive
  - https://www.googleapis.com/auth/admin.directory.group.readonly: ámbito para recuperar únicamente la información del grupo, el alias del grupo y los miembros. Esto es necesario para el Identity Crawler Amazon Kendra .
  - https://www.googleapis.com/auth/admin.directory.user.readonly: ámbito para recuperar únicamente usuarios o alias de usuario. Esto es necesario para incluir a los usuarios en el Amazon Kendra Identity Crawler y establecer las ACL.
  - https://www.googleapis.com/auth/cloud-platform: ámbito para generar un token de acceso que permita recuperar el contenido de archivos grandes de Google Drive.
  - https://www.googleapis.com/auth/forms.body.readonly: ámbito para obtener datos de Google Forms.

  Para que sea compatible con la API Forms, añada el siguiente ámbito adicional:

  - https://www.googleapis.com/auth/forms.body.readonly
- Ha comprobado que cada documento es único en Google Drive y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el tuyo Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Google Drive en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes ningún IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar tu fuente de datos de Google Drive Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarte Amazon Kendra a tu fuente de datos de Google Drive, debes proporcionar los detalles necesarios de tu fuente de datos de Google Drive para que Amazon Kendra puedas acceder a tus datos. Si aún no has configurado Google Drive para Amazon Kendra verRequisitos previos.

Console

Para conectarse Amazon Kendra a Google Drive

1. Inicia sesión en la Amazon Kendra consola AWS Management Console y ábrela.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector de Google Drive y, a continuación, selecciona Añadir conector. Si utilizas la versión 2 (si corresponde), elige el conector de Google Drive con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elige un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

b. En Autenticación: elija entre Cuenta de servicio de Google y Autenticación OAuth 2.0 según el caso de uso.

c. AWS Secrets Manager secreto: elige un secreto existente o crea uno nuevo Secrets Manager para almacenar tus credenciales de autenticación de Google Drive. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

   i. Si has elegido una cuenta de servicio de Google, introduce un nombre para tu secreto, el ID de correo electrónico del usuario administrador o «usuario de la cuenta de servicio» en la configuración de la cuenta de servicio (correo electrónico de administrador), el ID de correo electrónico de la cuenta de servicio (correo electrónico del cliente) y la clave privada que creaste en tu cuenta de servicio.

      Guarda y añade tu secreto

   ii. Si has elegido la autenticación de OAuth 2.0, introduce un nombre para el secreto, el ID de cliente, el secreto de cliente y el token de actualización que creaste en tu cuenta de OAuth.

      Guarda y añade tu secreto.

d. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

e. (Solo para usuarios de autenticación de cuentas de servicio de Google)

   Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

f. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

g. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Sincronizar contenido: seleccione las opciones o el contenido que desee rastrear. Puedes elegir entre rastrear Mi disco duro (carpetas personales), el disco compartido (carpetas compartidas contigo) o ambos. También puedes incluir comentarios en los archivos.

   b. En Configuración adicional (opcional) También puede introducir la siguiente información opcional:

      i. Audiencias objetivo: añada audiencias objetivo específicas para los documentos que desee rastrear.

      ii. Tamaño máximo de archivo: establece el límite de tamaño máximo en MB de los archivos que se van a rastrear.

      iii. Correo electrónico del usuario: añada los correos electrónicos de los usuarios que desee incluir o excluir.

      iv. Unidades compartidas: añada los nombres de las unidades compartidas que desee incluir o excluir.

      v. Tipos de MIME: añada los tipos de MIME que desee incluir o excluir.

      vi. Patrones de expresiones regulares de entidades: añada patrones de expresiones regulares para incluir o excluir determinados archivos adjuntos de todas las entidades compatibles. Puede agregar hasta 100 patrones.

   c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

> **Important**
>
> La API de Google Drive no admite la recuperación de comentarios de un archivo eliminado permanentemente. Los comentarios de los archivos colocados en la papelera se pueden recuperar. Cuando un archivo quede en la papelera, el conector eliminará los comentarios del Amazon Kendra índice.

d. En Sincronizar programación de ejecución, en Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

e. En Sincronizar el historial de ejecuciones, elija almacenar los informes generados automáticamente en una y Amazon S3 al sincronizar la fuente de datos. Esto resulta útil para realizar un seguimiento de los problemas al sincronizar la fuente de datos.

f. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Para archivos: seleccione entre los campos de la fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

> **Note**
>
> La API de Google Drive no admite la creación de campos personalizados. La asignación de campos personalizados no está disponible para el conector de Google Drive.

b. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Google Drive

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `GOOGLEDRIVEV2` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de autenticación: especifique si desea utilizar la autenticación de la cuenta de servicio o la autenticación de OAuth 2.0.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

> **Important**
>
> La API de Google Drive no admite la recuperación de comentarios de un archivo eliminado permanentemente. Los comentarios de los archivos colocados en la papelera se pueden recuperar. Cuando un archivo quede en la papelera, el conector eliminará los comentarios del Amazon Kendra índice.

- Nombre secreto del recurso de Amazon (ARN): proporciona el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creaste en tu cuenta de Google Drive. Si utiliza la autenticación de cuenta de servicio de Google, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "clientEmail": "user account email",
    "adminAccountEmail": "service account email",
    "privateKey": "private key"
}
```

  Si utiliza la autenticación OAuth 2.0, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "clientID": "OAuth client ID",
    "clientSecret": "client secret",
    "refreshToken": "refresh token"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Google Drive y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Google Drive.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.

- Mis unidades, unidades compartidas y comentarios: puedes especificar si deseas rastrear este tipo de contenido.
- Filtros de inclusión y exclusión: puede especificar si desea incluir o excluir determinadas cuentas de usuario, unidades compartidas y tipos de MIME.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Lista de control de acceso (ACL): especifique si desea rastrear la información de la ACL de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.
- Asignaciones de campos: elija asignar los campos del origen de datos de Google Drive a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el

nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice_`document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el Esquema de plantilla de Google Drive.

Notas

- La asignación de campos personalizados no está disponible para el conector de Google Drive, ya que la interfaz de usuario de Google Drive no admite la creación de este tipo de campos.
- La API de Google Drive no admite la recuperación de comentarios de un archivo eliminado permanentemente. Sin embargo, los comentarios de los archivos colocados en la papelera se pueden recuperar. Cuando un archivo quede en la papelera, el Amazon Kendra conector eliminará los comentarios del Amazon Kendra índice.
- La API de Google Drive no devuelve los comentarios presentes en un archivo .docx.

IBM DB2

IBM DB2 es un sistema de gestión de bases de datos relacionales desarrollado por IBM. Un usuario de IBM DB2 puede usar Amazon Kendra para indexar su origen de datos de IBM DB2. El conector Amazon Kendra IBM DB2 de fuente de datos es compatible con DB2 11.5.7.

Puede conectarse Amazon Kendra a su fuente de IBM DB2 datos mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de la fuente de Amazon Kendra IBM DB2 datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de IBM DB2 datos, realice estos cambios en sus AWS cuentas IBM DB2 y.

En IBM DB2, asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos.
- Ha comprobado que cada documento es único en IBM DB2 y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de IBM DB2 en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de IBM DB2 datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de IBM DB2 datos, debe proporcionar los detalles de sus IBM DB2 credenciales para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configuradoIBM DB2, Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a IBM DB2

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el IBM DB2conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el IBM DB2conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

    a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

    b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

    c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

    d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

    e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:

    a. En Origen, introduzca la siguiente información:

    b. Host: introduzca el nombre del host de la base de datos.

    c. Puerto: introduzca el puerto de la base de datos.

    d. Instancia: introduzca la instancia de la base de datos.

    e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.

    f. En Autenticación, introduzca la siguiente información:

        - AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de IBM DB2 autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

            A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

                I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-IBM DB2 -' se añade automáticamente a tu nombre secreto.

II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

- Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
- Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
- Columna de cuerpo: proporcione el nombre de la columna de cuerpo del documento en la tabla de la base de datos.

b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

- Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.

- Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
- Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
- Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.
- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

   - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
   - Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
   - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra IBM DB2

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `db2`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de

la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. IBM DB2 El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. IBM DB2 Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de IBM DB2.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.

- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para sus documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos del origen de datos de IBM DB2 a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Esquema de plantilla de IBM DB2.

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.
- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.
- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Jira

Jira es una herramienta de gestión de proyectos para el desarrollo de software, la gestión de productos y el seguimiento de errores. Puedes usarlo Amazon Kendra para indexar tus proyectos, incidencias, comentarios, archivos adjuntos, registros de trabajo y estados de Jira.

Amazon Kendra actualmente solo es compatible con Jira Cloud.

Puedes conectarte Amazon Kendra a tu fuente de datos de Jira mediante la Amazon Kendra consola o la JiraConfiguration API. Para ver una lista de las características admitidas por cada una, consulte Características admitidas.

Para solucionar problemas del conector de fuentes de datos de Amazon Kendra Jira, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra El conector de fuente de datos de Jira admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Jira, realiza estos cambios en tu Jira y en tus cuentas. AWS

En Jira, asegúrese de que:

- Credenciales de autenticación mediante token de API configuradas, que incluyen un ID de Jira (nombre de usuario o correo electrónico) y una credencial de Jira (token de API de Jira). Consulte la Documentación de Atlassian sobre la administración de los tokens de API.

  > **Note**
  >
  > Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha apuntado la URL de la cuenta de Jira en la configuración de su cuenta de Jira. Por ejemplo, *`https://company.atlassian.net/`*.
- Ha comprobado que cada documento es único en Jira y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el tuyo Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

  > **Note**
  >
  > Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Jira en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

  > **Note**
  >
  > Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes un IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar tu fuente de datos de Jira. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarte Amazon Kendra a tu fuente de datos de Jira, debes proporcionar los detalles necesarios de tu fuente de datos de Jira para que Amazon Kendra puedas acceder a tus datos. Si aún no has configurado Jira para Amazon Kendra, consulta. Requisitos previos

Console

Para conectarse Amazon Kendra a Jira

1. Inicia sesión en la Amazon Kendra consola AWS Management Console y ábrela.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector de Jira y, a continuación, selecciona Añadir conector. Si utilizas la versión 2 (si corresponde), elige el conector de Jira con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elige un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. URL de la cuenta de Jira: introduce la URL de tu cuenta de Jira. Por ejemplo, *https://company.atlassian.net/*.

   b. Autorización: activa o desactiva la información de la lista de control de acceso (ACL) de tus documentos si tienes una ACL y quieres usarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

   c. AWS Secrets Manager secreto: elige un secreto existente o crea uno nuevo Secrets Manager para almacenar tus credenciales de autenticación de Jira. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

      i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

         A. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Jira» se añade automáticamente a tu nombre secreto.

         B. En ID de Jira: introduzca el nombre de usuario o el correo electrónico de Jira.

         C. Para la contraseña/token: introduce el token de la API de Jira configurado en Jira.

      ii. Guarda y añade tu secreto.

   d. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

   e. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

f. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

g. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. Selecciona qué proyectos de Jira deseas indexar: elige rastrear todos los proyectos o proyectos específicos.

b. Configuración adicional: especifique determinados estados y tipos de problemas. Elija rastrear los comentarios, los archivos adjuntos y los registros de trabajo. Utilice patrones de expresiones regulares para incluir o excluir cierto contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Jira

Debes especificar lo siguiente mediante la JiraConfiguration API:

- URL del origen de datos: especifique la URL de su cuenta de Jira. Por ejemplo, *company.atlassian.net*.
- Nombre secreto de recurso de Amazon (ARN): proporciona el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de tu cuenta de Jira. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "jiraId": "Jira user name or email",
    "jiraCredential": "Jira API token"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Jira y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Jira.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique `VpcConfiguration` como parte de la configuración del origen de datos. Consulte Configuración de Amazon Kendra para utilizar una VPC.
- Registro de cambios: si se Amazon Kendra debe utilizar el mecanismo de registro de cambios de la fuente de datos de Jira para determinar si un documento debe actualizarse en el índice.

> **Note**
>
> Utilice el registro de cambios si no quiere que Amazon Kendra digitalice todos los documentos. Si el registro de cambios es grande, es posible que se tarde Amazon Kendra menos en escanear los documentos de la fuente de datos de Jira que en procesar el registro de cambios. Si está sincronizando el origen de datos de Jira con su índice por primera vez, se escanean todos los documentos.

- Filtros de inclusión y exclusión: puedes especificar si deseas incluir o excluir determinados archivos.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Comentarios, archivos adjuntos y registros de trabajo: puede especificar si desea rastrear determinados comentarios, archivos adjuntos y registros de trabajo relacionados con los problemas.
- Proyectos, problemas y estados: puedes especificar si deseas rastrear determinados identificadores de proyectos, tipos de problemas y estados.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

- Asignaciones de campos: elija asignar los campos del origen de datos de Jira a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Más información

Para obtener más información sobre la integración Amazon Kendra con tu fuente de datos de Jira, consulta:

- Busca tus proyectos de Jira de forma inteligente con el conector de Jira Cloud Amazon Kendra

Microsoft Exchange

Microsoft Exchange es una herramienta de colaboración empresarial para mensajería, reuniones e intercambio de archivos. Si es usuario de Microsoft Exchange, puede utilizarlo Amazon Kendra para indexar su fuente de datos de Microsoft Exchange.

Puede conectarse Amazon Kendra a su fuente de datos de Microsoft Exchange mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Microsoft Exchange, consulteSolución de problemas con los orígenes de datos.

Características admitidas

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de datos de Microsoft Exchange, realice estos cambios en su Microsoft Exchange y en sus AWS cuentas.

En Microsoft Exchange, asegúrese de que:

- Ha creado una cuenta de Microsoft Exchange en Office 365.
- Ha apuntado su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
- Configuró una aplicación OAuth en el portal de Azure y anotó el ID y el secreto del cliente o las credenciales del cliente. Consulte el tutorial de Microsoft y el ejemplo de aplicación registrada para obtener más información.

> **Note**
>
> Al crear o registrar una aplicación en el portal de Azure, el ID secreto representa el valor secreto real. Debe anotar o guardar el valor secreto real inmediatamente al crear el secreto y la aplicación. Para acceder a su secreto, seleccione el nombre de la aplicación en el portal de Azure y, a continuación, vaya a la opción de menú relativa a los certificados y secretos.
> Para acceder a su ID de cliente, seleccione el nombre de su aplicación en el portal de Azure y, a continuación, vaya a la página de información general. El ID de la aplicación (cliente) es el ID del cliente.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha agregado los siguientes permisos para la aplicación del conector:

Microsoft Graph	Office 365 Exchange Online
• Mail.Read (Aplicación)	full_access_as_app (Aplicación)

Microsoft Graph	Office 365 Exchange Online
• Correo. ReadBasic (Solicitud) • Correo. ReadBasic.All (Aplicación) • Calendars.Read (Aplicación) • User.Read.All (Aplicación) • Contacts.Read (Aplicación) • Notes.Read.All (aplicación) • Directory.Read.All (Aplicación) • NOTICIAS. AccessAsUser.Todos (delegados)	

- Ha comprobado que cada documento es único en Microsoft Exchange y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En su interior Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Microsoft Exchange en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y Secrets Manager secreto al conectar su fuente de datos de Microsoft Exchange a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a la fuente de datos de Microsoft Exchange, debe proporcionar los detalles necesarios de la fuente de datos de Microsoft Exchange para que Amazon Kendra pueda acceder a los datos. Si aún no ha configurado Microsoft Exchange para Amazon Kendra, consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Microsoft Exchange

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector de Microsoft Exchange y, a continuación, elija Agregar conector. Si usa la versión 2 (si corresponde), elija el conector Microsoft Exchange con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. ID de inquilino: introduzca su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.

b. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

c. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de autenticación de Microsoft Exchange. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

A. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Microsoft Exchange

B. Para ID de cliente, secreto de cliente: introduzca las credenciales de autenticación configuradas en Microsoft Exchange en el portal de Azure.

ii. Guarde y añada su secreto.

d. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

e. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

f. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. ID de usuario: proporcione los correos electrónicos de los usuarios si desea filtrar el contenido por determinados correos electrónicos.

   b. Configuración adicional: especifique los tipos de contenido que desea rastrear.

      - Tipos de entidad: puedes elegir rastrear el contenido del calendario o de los contactos OneNotes.
      - Rastreo del calendario: introduce la fecha de inicio y finalización para rastrear el contenido entre determinadas fechas.
      - Incluir correo electrónico: introduce las líneas «para», «de» y asunto del correo electrónico para filtrar determinados correos electrónicos que quieras rastrear.
      - Acceso a carpetas compartidas: elija habilitar el rastreo de la lista de control de acceso para controlar el acceso a su fuente de datos de Microsoft Exchange.
      - Regex para dominios: añada patrones de expresiones regulares para incluir o excluir determinados dominios de correo electrónico.
      - Patrones regex: añada patrones de expresiones regulares para incluir o excluir determinados archivos.

   c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

      - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.

- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

> **Note**
>
> El conector de fuentes de datos de Amazon Kendra Microsoft Exchange no admite asignaciones de campos personalizadas.

b. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Microsoft Exchange

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `MSEXCHANGE` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- ID de inquilino: puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Nombre de recurso secreto de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta de Microsoft Exchange. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Microsoft Exchange y Amazon Kendra. Para obtener más información, consulte Roles de IAM para orígenes de datos de Microsoft Exchange.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinado contenido.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Lista de control de acceso (ACL): especifique si desea rastrear la información de la ACL de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos del origen de datos de Microsoft Exchange a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el esquema de plantillas de Microsoft Exchange.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de datos de Microsoft Exchange, consulte:

- Indexar el contenido de Microsoft Exchange mediante el conector de Exchange para Amazon Kendra

Microsoft OneDrive

Microsoft OneDrive es un servicio de almacenamiento basado en la nube que puedes usar para almacenar, compartir y alojar tu contenido. Puede usarlo Amazon Kendra para indexar su fuente OneDrive de datos.

Puede conectarse Amazon Kendra a su fuente OneDrive de datos mediante la Amazon Kendra consola y la OneDriveConfigurationAPI.

Amazon Kendra tiene dos versiones del OneDrive conector. Las características compatibles de cada versión incluyen:

OneDrive Conector OneDriveConfigurationV1.0/API de Microsoft

- Asignaciones de campo
- Filtros de inclusión/exclusión

OneDrive Conector TemplateConfigurationV2.0/API de Microsoft

- Filtrado de contexto de usuario
- Rastreador de identidad de usuario
- Filtros de inclusión/exclusión
- Sincronización de contenido completa e incremental
- Nube privada virtual (VPC)

> **Note**
>
> Está previsto que el soporte para el OneDrive conector OneDriveConfiguration V1.0/API finalice en junio de 2023. Recomendamos utilizar el OneDrive conector TemplateConfiguration V2.0/API.

Para solucionar problemas del conector de la fuente de Amazon Kendra OneDrive datos, consulteSolución de problemas con los orígenes de datos.

Temas

- OneDrive Conector Microsoft V1.0
- OneDrive Conector Microsoft V2.0
- Más información

OneDrive Conector Microsoft V1.0

Microsoft OneDrive es un servicio de almacenamiento basado en la nube que puedes usar para almacenar, compartir y alojar tu contenido. Se puede utilizar Amazon Kendra para indexar la fuente de OneDrive datos de Microsoft.

> **Note**
>
> Está previsto que el soporte para el OneDrive conector V1.0 y la OneDrive API de Microsoft finalice en junio de 2023. Recomendamos utilizar el OneDrive conector V2.0/API. TemplateConfiguration

Para solucionar problemas del conector de la fuente de Amazon Kendra OneDrive datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión

Características admitidas

- Asignaciones de campo
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de OneDrive datos, realice estos cambios en sus AWS cuentas OneDrive y.

En Azure Active Directory (AD), asegúrese de que:

- Ha creado una aplicación de Azure Active Directory (AD).
- Ha utilizado el ID de la aplicación de AD para registrar una clave secreta para la aplicación en el sitio de AD. La clave secreta debe contener el ID de la aplicación y una clave secreta.
- Ha copiado el dominio de AD de la organización.
- Se agregaron los siguientes permisos de aplicación a su aplicación de AD en la opción Microsoft Graph:
  - Leer los archivos de todas las colecciones de sitios (File.Read-All)
  - Leer el perfil completo de todos los usuarios (User.Read.All)
  - Leer los datos del directorio (Directory.Read.All)
  - Leer todos los grupos (Group.Read.All)
  - Leer los elementos de todas las colecciones de sitios (Site.Read.All)
- Ha copiado la lista de usuarios cuyos documentos se deben indexar. Puede elegir entre proporcionar una lista de nombres de usuario o puede proporcionarlos en un archivo almacenado en un Amazon S3. Después de crear el origen de datos, puede:
  - Modificar la lista de usuarios.
  - Cambia de una lista de usuarios a una lista almacenada en un Amazon S3 bucket.
  - Cambia la ubicación del Amazon S3 depósito de una lista de usuarios. Si cambias la ubicación del depósito, también debes actualizar el IAM rol de la fuente de datos para que tenga acceso al depósito.

> **Note**
>
> Si almacena la lista de nombres de usuario en un Amazon S3 depósito, la IAM política de la fuente de datos debe proporcionar acceso al depósito y acceso a la clave con la que se cifró el depósito, si la hubiera.

- Marcó que cada documento es único en OneDrive las demás fuentes de datos que vaya a utilizar para el mismo índice y entre ellas. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Guardó sus credenciales de OneDrive autenticación en un AWS Secrets Manager secreto y, si usa la API, anotó el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de OneDrive datos. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de OneDrive datos, debe proporcionar los detalles de sus OneDrive credenciales para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configurado OneDrive , Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a OneDrive

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el OneDrive conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el OneDrive conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. OneDrive ID de inquilino: introduzca el ID de OneDrive inquilino sin el protocolo.

b. En Tipo de autenticación: elija entre Nuevo y Existente.

c. i. Si elige Existente, seleccione un secreto existente en Seleccionar secreto.

ii. Si elige Nuevo, introduzca la siguiente información en la sección Nuevo secreto de AWS Secrets Manager :

A. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-OneDrive -' se añade automáticamente a su nombre secreto.

B. Para el identificador de la aplicación y la contraseña de la aplicación: introduzca los valores de las credenciales de autenticación de su OneDrive cuenta y, a continuación, seleccione Guardar autenticación.

d. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

e. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. Elija entre Archivo de lista y Lista de nombres según el caso de uso.

i. Si elige Archivo de lista, introduzca la siguiente información:

- Seleccionar ubicación: introduzca la ruta a su bucket de Amazon S3 .

  Añadir el archivo de lista de usuarios a Amazon S3: seleccione esta opción para añadir los archivos de la lista de usuarios al bucket Amazon S3 .

  Asignaciones de grupos locales de usuarios: seleccione esta opción para utilizar la asignación de grupos locales para filtrar el contenido.

ii. Si elige Lista de nombres, introduzca la siguiente información:

- Nombre de usuario: introduzca hasta 10 unidades de usuario para indexarlas. Para añadir más de 10 usuarios, cree un archivo que contenga los nombres.

Añadir otro: elija esta opción para añadir más usuarios.

Asignaciones de grupos locales de usuarios: seleccione esta opción para utilizar la asignación de grupos locales para filtrar el contenido.

b. En Configuración adicional: añada patrones de expresión regular para incluir o excluir determinados archivos. Puede agregar hasta 100 patrones.

c. En Sincronizar el programa de ejecución, para Frecuencia: elija la frecuencia con la Amazon Kendra que se sincronizará con su fuente de datos.

d. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Para los campos de fuente de datos predeterminados y otras asignaciones de campos sugeridas: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

b. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra OneDrive

Debe especificar lo siguiente mediante la OneDriveConfigurationAPI:

- ID de inquilino: especifique el dominio de Azure Active Directory de la organización.
- OneDrive Usuarios: especifique la lista de cuentas de usuario cuyos documentos deben indexarse.
- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta. OneDrive El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "username": "OAuth client ID",
    "password": "client secret"
```

```
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. OneDrive Amazon Kendra Para obtener más información, consulte las IAM funciones de las fuentes OneDrive de datos.

También puede añadir las siguientes características opcionales:

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados documentos.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Asignaciones de campos: elija asignar los campos de la fuente de OneDrive datos a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

OneDrive Conector Microsoft V2.0

Microsoft OneDrive es un servicio de almacenamiento basado en la nube que puedes usar para almacenar, compartir y alojar tu contenido. Puede usarlo Amazon Kendra para indexar su fuente OneDrive de datos.

Puede conectarse Amazon Kendra a su fuente OneDrive de datos mediante la Amazon Kendra consola y la OneDriveConfigurationAPI.

> **Note**
>
> Está previsto que el soporte para OneDrive Connector OneDriveConfiguration V1.0/API finalice en junio de 2023. Recomendamos utilizar el OneDrive conector TemplateConfiguration V2.0/API. La versión 2.0 proporciona ACL adicionales y la funcionalidad del rastreador de identidades.

Para solucionar problemas del conector de la fuente de Amazon Kendra OneDrive datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión

Características admitidas

Amazon Kendra OneDrive el conector de fuente de datos admite las siguientes funciones:

- Asignaciones de campo
- control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de OneDrive datos, realice estos cambios en sus AWS cuentas OneDrive y.

En OneDrive, asegúrate de tener:

- Creó una OneDrive cuenta en Office 365.
- Ha apuntado su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
- Creó una aplicación OAuth en el portal de Azure y anotó el ID de cliente y el secreto del cliente o las credenciales del cliente utilizadas para la autenticación con un AWS Secrets Manager secreto. Consulte el tutorial de Microsoft y el ejemplo de aplicación registrada para obtener más información.

> **Note**
>
> Al crear o registrar una aplicación en el portal de Azure, el ID secreto representa el valor secreto real. Debe anotar o guardar el valor secreto real inmediatamente al crear el secreto y la aplicación. Para acceder a su secreto, seleccione el nombre de la aplicación en el portal de Azure y, a continuación, vaya a la opción de menú relativa a los certificados y secretos.
> Para acceder a su ID de cliente, seleccione el nombre de su aplicación en el portal de Azure y, a continuación, vaya a la página de información general. El ID de la aplicación (cliente) es el ID del cliente.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha utilizado el ID de la aplicación de AD para registrar una clave secreta para la aplicación en el sitio de AD. La clave secreta debe contener el ID de la aplicación y una clave secreta.
- Ha copiado el dominio de AD de la organización.

- Ha agregado los siguientes permisos a la aplicación de AD en la opción Microsoft Graph:
  - Leer los archivos de todas las colecciones de sitios (File.Read-All)
  - Lea los perfiles completos de todos los usuarios (User.Read.All)
  - Leer todos los grupos (Group.Read.All)
  - Lea todas las notas (Notes.Read.All)
- Ha copiado la lista de usuarios cuyos documentos se deben indexar. Puede elegir entre proporcionar una lista de nombres de usuario o puede proporcionarlos en un archivo almacenado en un Amazon S3. Después de crear el origen de datos, puede:
  - Modificar la lista de usuarios.
  - Cambie de una lista de usuarios a una lista almacenada en un Amazon S3 bucket.
  - Cambia la ubicación del Amazon S3 depósito de una lista de usuarios. Si cambias la ubicación del depósito, también debes actualizar el IAM rol de la fuente de datos para que tenga acceso al depósito.

> **Note**
>
> Si almacena la lista de nombres de usuario en un Amazon S3 depósito, la IAM política de la fuente de datos debe proporcionar acceso al depósito y acceso a la clave con la que se cifró el depósito, si la hubiera.
> El OneDrive conector utiliza el correo electrónico de la información de contacto presente en las propiedades de usuario de Onedrive. Asegúrese de que el usuario cuyos datos desea rastrear tenga el campo de correo electrónico configurado en la página Información de contacto, ya que, en el caso de los nuevos usuarios, podría estar en blanco.

En tu AWS cuenta, asegúrate de tener:

- Creó un Amazon Kendra índice y, si utiliza la API, anotó el identificador del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.
- Guardó sus credenciales de OneDrive autenticación en un AWS Secrets Manager secreto y, si usa la API, anotó el ARN del secreto.

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de OneDrive datos. Amazon Kendra Si utiliza la API,

debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un identificador de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de OneDrive datos, debe proporcionar los detalles de sus OneDrive credenciales para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configurado OneDrive Amazon Kendra, consulte Requisitos previos.

Console

Para conectarse Amazon Kendra a OneDrive

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el OneDrive conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el OneDrive conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. OneDrive ID de inquilino: introduzca el ID de OneDrive inquilino sin el protocolo.

b. Autorización: active o desactive la información de la lista de control de acceso (ACL) para sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

c. En Autenticación: elija entre Nueva y Existente.

d. i. Si elige Existente, seleccione un secreto existente en Seleccionar secreto.

   ii. Si elige Nuevo, introduzca la siguiente información en la sección Nuevo secreto de AWS Secrets Manager :

      A. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-OneDrive -' se añade automáticamente a su nombre secreto.

      B. Para el ID de cliente y el secreto del cliente: introduzca el ID y el secreto del cliente.

e. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

f. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

g. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

h. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

8. a. Para el ámbito de sincronización: elija los OneDrive datos de los usuarios que desee indexar. Puede agregar un máximo de 10 usuarios de forma manual.

   b. En Configuración adicional: añada patrones de expresión regular para incluir o excluir determinado contenido. Puede agregar hasta 100 patrones.

   c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no eliges la sincronización completa como opción de modo de sincronización.

      - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
      - Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
      - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

   d. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

   e. Elija Siguiente.

9. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

b. Elija Siguiente.

10. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra OneDrive

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `ONEDRIVEV2` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- ID de inquilino: especifique el ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no eliges la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de

datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. OneDrive

  Si utiliza la autenticación OAuth 2.0, el secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "clientId": "client ID",
      "clientSecret": "client secret"
  }
  ```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. OneDrive Amazon Kendra Para obtener más información, consulte las IAM funciones de las fuentes OneDrive de datos.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir o excluir determinados archivos, OneNote secciones y OneNote páginas.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso

(ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

- Asignaciones de campos: solo puede mapear campos de índice integrados o comunes para el conector. Amazon Kendra OneDrive La asignación de campos personalizada no está disponible para el OneDrive conector debido a las limitaciones de la API. Para obtener más información, consulte Asignación de campos de origen de datos.

Para ver una lista de otras claves JSON importantes que debes configurar, consulta el esquema OneDrive de la plantilla.

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de OneDrive datos, consulte:

- Anunciamos el OneDrive conector de Microsoft (V2) actualizado para Amazon Kendra.

Microsoft SharePoint

SharePoint es un servicio colaborativo de creación de sitios web que puede utilizar para personalizar el contenido web y crear páginas, sitios, bibliotecas de documentos y listas. Puede utilizarlo Amazon Kendra para indexar la fuente SharePoint de datos.

Amazon Kendra actualmente es compatible con SharePoint Online y SharePoint Server (versiones 2013, 2016, 2019 y Subscription Edition).

Puede conectarse Amazon Kendra a su fuente de SharePoint datos mediante la Amazon Kendra consola, la TemplateConfigurationAPI o la SharePointConfigurationAPI.

Amazon Kendra tiene dos versiones del SharePoint conector. Las características compatibles de cada versión incluyen:

SharePoint Conector V1.0/API SharePointConfiguration

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Registro de cambios
- Nube privada virtual (VPC)

SharePoint Conector V2.0/ API TemplateConfiguration

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

> **Note**
>
> Está previsto que el soporte para el SharePoint conector SharePointConfiguration V1.0/API finalice en 2023. Recomendamos migrar o utilizar el SharePoint conector V2.0/API. TemplateConfiguration

Para solucionar problemas del conector de la fuente de Amazon Kendra SharePoint datos, consulte. Solución de problemas con los orígenes de datos

Temas

- SharePoint conector V1.0
- SharePoint conector V2.0

SharePoint conector V1.0

SharePoint es un servicio colaborativo de creación de sitios web que puede utilizar para personalizar el contenido web y crear páginas, sitios, bibliotecas de documentos y listas. Si es un SharePoint usuario, puede utilizarlo Amazon Kendra para indexar su fuente SharePoint de datos.

> **Note**
>
> Está previsto que el soporte para el SharePoint conector SharePointConfiguration V1.0/API finalice en 2023. Recomendamos migrar o utilizar el SharePoint conector V2.0/API. TemplateConfiguration

Para solucionar problemas del conector de la fuente de Amazon Kendra SharePoint datos, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Registro de cambios
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de SharePoint datos, realice estos cambios en sus AWS cuentas SharePoint y.

Debe proporcionar las credenciales de autenticación, que se almacenan de forma segura en AWS Secrets Manager secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

En SharePoint, asegúrese de tener:

- Apuntó la URL de los SharePoint sitios que desea indexar.
- Para SharePoint en línea:
  - Ha apuntado sus credenciales de autenticación básica, que incluyen un nombre de usuario y una contraseña con permisos de administrador del sitio.
  - Opcional: ha generado credenciales OAuth 2.0 que contienen un nombre de usuario, una contraseña, un ID de cliente y un secreto de cliente.
  - Ha desactivado los Valores predeterminados de seguridad en su portal de Azure mediante un usuario administrativo. Para obtener más información sobre la administración de la configuración predeterminada de seguridad en el portal de Azure, consulte la Documentación de Microsoft sobre cómo habilitar o deshabilitar la configuración predeterminada de seguridad.
- Para el SharePoint servidor:
  - Apuntó el nombre de dominio de su SharePoint servidor (el nombre de NetBIOS en su Active Directory). Úselo, junto con su nombre de usuario y contraseña de autenticación SharePoint básicos, para conectarse al SharePoint Amazon Kendra servidor.

> **Note**
>
> Si utiliza SharePoint Server y necesita convertir la lista de control de acceso (ACL) al formato de correo electrónico para filtrar según el contexto del usuario, proporcione la URL del servidor LDAP y la base de búsqueda de LDAP. También puede utilizar la anulación del dominio del directorio. La URL del servidor LDAP es el nombre de dominio completo y el número de puerto (por ejemplo, ldap://example.com:389). La base de búsqueda de LDAP son los controladores de dominio “example” y “com”. Al anular el dominio del directorio, puede utilizar el dominio del correo electrónico en lugar de la URL del servidor LDAP y la base de búsqueda LDAP. Por ejemplo, el dominio de correo electrónico de username@example.com es “example.com”. Puede usar esta anulación si no le interesa validar su dominio y simplemente quiere usar su dominio de correo electrónico.

- Se agregaron los siguientes permisos a su cuenta: SharePoint

Para SharePoint listas

- Abrir elementos: vea el origen de los documentos con los controladores de archivos del servidor.
- Ver páginas de aplicaciones: vea formularios, vistas y páginas de aplicaciones. Enumere las listas.
- Ver elementos: vea los elementos de las listas y los documentos de las bibliotecas de documentos.
- Ver versiones: vea las versiones anteriores de un documento o elemento de la lista.

Para SharePoint sitios web

- Examinar directorios: enumere los archivos y carpetas de un sitio web mediante la interfaz SharePoint Designer y Web DAV.
- Examinar la información del usuario: vea información sobre los usuarios del sitio web.
- Enumerar permisos: enumere los permisos en el sitio web, la lista, la carpeta, el documento o el elemento de la lista.
- Abrir: abra un sitio web, una lista o una carpeta para acceder a los elementos del contenedor.
- Utilice las funciones de integración de clientes: utilice SOAP, WebDAV, el modelo de objetos del cliente o SharePoint las interfaces de diseñador para acceder al sitio web.
- Utilizar interfaces remotas: use características que lanzan aplicaciones cliente.
- Ver páginas: vea las páginas de un sitio web.

- Marcó que cada documento es único en las demás fuentes de datos que vaya a utilizar para el mismo índice SharePoint y entre ellas. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Guardó sus credenciales de SharePoint autenticación en un AWS Secrets Manager secreto y, si usa la API, anotó el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de SharePoint datos. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de SharePoint datos, debe proporcionar los detalles de sus SharePoint credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configurado SharePoint , Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a SharePoint

1. Inicie sesión en la consola AWS de administración y abra la Amazon Kendra consola.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el SharePoint conector v1.0 y, a continuación, elija Agregar fuente de datos.
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. Para el método de alojamiento: elija entre SharePoint en línea y SharePoint servidor.

i. Para estar SharePointen línea: introduce las URL del sitio específicas de tu repositorio. SharePoint

ii. Para el SharePointservidor: elija su SharePoint versión, introduzca las URL del sitio específicas de su SharePoint repositorio e introduzca la Amazon S3 ruta a la ubicación de su certificado SSL.

b. (Solo SharePoint servidor) Para el proxy web: introduzca el nombre de host y el número de puerto de la instancia interna. SharePoint El número de puerto debe ser un valor numérico entre 0 y 65535.

c. En Autenticación: elija entre las siguientes opciones según el caso de uso:

i. Para SharePoint Internet: elige entre la autenticación básica y la autenticación OAuth 2.0.

ii. Para el SharePoint servidor: elige entre Ninguno, LDAP y Manual.

d. Para AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo para almacenar sus credenciales de Secrets Manager autenticación. SharePoint Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager . Debe introducir un Nombre del secreto. El prefijo 'AmazonKendra- SharePoint -' se añade automáticamente a tu nombre secreto.

e. Introduzca la siguiente información adicional en la ventana Crear un secreto de AWS Secrets Manager :

i. Elija entre las siguientes opciones de autenticación SharePoint en la nube, según su caso de uso:

   A. Autenticación básica: introduzca el nombre de usuario de su SharePoint cuenta como nombre de usuario y la contraseña de la SharePoint cuenta como contraseña.

   B. Autenticación OAuth 2.0: introduce el nombre de usuario de tu SharePoint cuenta como nombre de usuario, la contraseña de la SharePoint cuenta como contraseña, tu SharePoint ID único generado automáticamente como ID de cliente y la cadena secreta compartida utilizada por ambos SharePoint y Amazon Kendra como secreto de cliente.

ii. Elige una de las siguientes opciones de autenticación SharePoint del servidor, según tu caso de uso:

   A. Ninguna: introduzca el nombre de usuario de su SharePoint cuenta como nombre de usuario, la contraseña de su SharePoint cuenta como contraseña y el nombre de dominio del servidor.

   B. LDAP ***: introduzca el nombre de usuario de su SharePoint cuenta como nombre de usuario, la contraseña de la SharePoint cuenta como contraseña, el punto de conexión del servidor LDAP (incluidos el protocolo y el número de puerto, por ejemplo, ldap: //example.com:389) y su base de búsqueda de LDAP (por ejemplo, dc=example, dc=com).***

   C. Manual: introduzca el nombre de usuario de su SharePoint cuenta como nombre de usuario, la contraseña de su SharePoint cuenta como contraseña y la anulación del dominio de correo electrónico (dominio de correo electrónico del usuario o grupo del directorio).

iii. Seleccione Guardar.

f. Nube privada virtual (VPC): también debe agregar Subredes y Grupos de seguridad de VPC.

> **Note**
>
> Debe usar una VPC si usa SharePoint Server. Amazon VPC es opcional para otras SharePoint versiones.

g. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

h. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Usar registro de cambios: seleccione esta opción para actualizar el índice en lugar de sincronizar todos los archivos.

   b. Rastrear archivos adjuntos: seleccione esta opción para rastrear los archivos adjuntos.

   c. Utilizar asignaciones de grupos locales: seleccione esta opción para asegurarse de que los documentos se filtran correctamente.

   d. Configuración adicional: añada patrones de expresión regular para incluir o excluir determinados archivos. Puede agregar hasta 100 patrones.

   e. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

   f. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Amazon Kendra asignaciones de campos predeterminadas: seleccione entre las fuentes de datos predeterminadas Amazon Kendra generadas los campos que desee asignar a su índice.

   b. En Asignaciones de campo personalizado: agregue campos de origen de datos personalizados a fin de crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra SharePoint

Debe especificar lo siguiente mediante la SharePointConfigurationAPI:

- SharePointVersión: especifique la SharePoint versión que utiliza al configurar SharePoint. Este es el caso independientemente de si utiliza SharePoint Server 2013, SharePoint Server 2016, SharePoint Server 2019 u SharePoint Online.
- Nombre de recurso secreto de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su SharePoint cuenta. El secreto se almacena en una estructura JSON.

  Para la autenticación básica SharePoint en línea, la siguiente es la estructura JSON mínima que debe estar en el secreto:

```
{
    "userName": "user name",
    "password": "password"
}
```

  Para la autenticación OAuth 2.0 SharePoint en línea, la siguiente es la estructura JSON mínima que debe estar en tu secreto:

```
{
    "userName": "SharePoint account user name'"",
    "password": "SharePoint account password",
    "clientId": "SharePoint auto-generated unique client id",
    "clientSecret": "secret string shared by Amazon Kendra and SharePoint to
 authorize communications"
}
```

  Para la autenticación básica SharePoint del servidor, la siguiente es la estructura JSON mínima que debe estar en tu secreto:

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name"
}
```

Para la autenticación LDAP SharePoint del servidor (si necesita convertir la lista de control de acceso (ACL) al formato de correo electrónico para filtrar según el contexto del usuario, puede incluir la URL del servidor LDAP y la base de búsqueda de LDAP en su secreto), la siguiente es la estructura JSON mínima que debe estar en su secreto:

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name"
    "ldapServerUrl": "ldap://example.com:389",
    "ldapSearchBase": "dc=example,dc=com"
}
```

Para la autenticación manual SharePoint del servidor, la siguiente es la estructura JSON mínima que debe estar en secreto:

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name",
    "emailDomainOverride": "example.com"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el SharePoint conector y Amazon Kendra. Para obtener más información, consulte las IAM funciones de las fuentes SharePoint de datos.
- Amazon VPC—Si usa SharePoint Server, especifíquelo `VpcConfiguration` como parte de la configuración de la fuente de datos. Consulte Configuración Amazon Kendra para usar una VPC.

También puede añadir las siguientes características opcionales:

- Proxy web: si debe conectarse a las URL de su SharePoint sitio mediante un proxy web. Puede usar esta opción solo para SharePoint el servidor.
- Listas de indexación: si se Amazon Kendra debe indexar el contenido de los archivos adjuntos a los elementos de la SharePoint lista.

- Registro de cambios: si se Amazon Kendra debe utilizar el mecanismo de registro de cambios de la fuente de SharePoint datos para determinar si un documento debe actualizarse en el índice.

  > **Note**
  >
  > Utilice el registro de cambios si no quiere que Amazon Kendra digitalice todos los documentos. Si el registro de cambios es grande, es posible que se Amazon Kendra tarde menos en digitalizar los documentos de la fuente de SharePoint datos que en procesar el registro de cambios. Si sincroniza la fuente de SharePoint datos con el índice por primera vez, se digitalizarán todos los documentos.

- Filtros de inclusión y exclusión: puede especificar si desea incluir o excluir determinado contenido.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Asignaciones de campos: elija asignar los campos de la fuente de SharePoint datos a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se

utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente SharePoint de datos, consulte:

- Cómo empezar a utilizar el conector Amazon Kendra SharePoint en línea

SharePoint conector V2.0

SharePoint es un servicio colaborativo de creación de sitios web que puede utilizar para personalizar el contenido web y crear páginas, sitios, bibliotecas de documentos y listas. Puede utilizarlo Amazon Kendra para indexar la fuente SharePoint de datos.

Amazon Kendra actualmente es compatible con SharePoint Online and SharePoint Server (2013, 2016, 2019 y Subscription Edition).

> **Note**
>
> Está previsto que el soporte para el SharePoint conector SharePointConfiguration V1.0/ API finalice en 2023. Recomendamos migrar o utilizar el SharePoint conector V2.0/API. TemplateConfiguration

Para solucionar problemas del conector de la fuente de Amazon Kendra SharePoint datos, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

Amazon Kendra SharePoint el conector de fuente de datos admite las siguientes funciones:

- Asignaciones de campo
- control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de SharePoint datos, realice estos cambios en sus AWS cuentas SharePoint y.

Debe proporcionar las credenciales de autenticación, que se almacenan de forma segura en AWS Secrets Manager secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

En SharePoint línea, asegúrese de tener:

- Ha copiado las direcciones URL de la SharePoint instancia. El formato de la URL del host que introduzca es *`https://yourdomain.sharepoint.com/sites/mysite`*. La URL debe empezar por `https` y contener `sharepoint.com`.
- Se ha copiado el nombre de dominio de la URL de la SharePoint instancia.
- Apuntó sus credenciales de autenticación básicas, que incluyen el nombre de usuario y la contraseña, además de los permisos de administrador del sitio para conectarse a SharePoint Online.
- Ha desactivado los Valores predeterminados de seguridad en su portal de Azure mediante un usuario administrativo. Para obtener más información sobre la administración de la configuración predeterminada de seguridad en el portal de Azure, consulte la Documentación de Microsoft sobre cómo habilitar o deshabilitar la configuración predeterminada de seguridad.

- Has desactivado la autenticación multifactor (MFA) en tu SharePoint cuenta para que no Amazon Kendra se bloquee el rastreo de tu contenido. SharePoint
- Si utilizas un tipo de autenticación distinto de la autenticación básica: has copiado el ID de inquilino de la instancia. SharePoint Para obtener más información sobre cómo encontrar el ID de inquilino, consulte Encontrar el ID de inquilino de Microsoft 365.
- Si necesita migrar a la autenticación de usuarios en la nube con Microsoft Entra, consulte la documentación de Microsoft sobre la autenticación en la nube.
- Para la autenticación de OAuth 2.0 y la autenticación mediante token de actualización de OAuth 2.0: anote sus credenciales de autenticación básica, que contienen el nombre de usuario y la contraseña que utiliza para conectarse a SharePoint Online, así como el ID de cliente y el secreto de cliente generados tras registrarse SharePoint en Azure AD.
  - Si no usa ACL, ha agregado los siguientes permisos:

Microsoft Graph	SharePoint
• Notes.Read.All (aplicación): lee todos los cuadernos OneNote • Sites.Read.All (aplicación): lee los elementos de todas las colecciones de sitios	• AllSites.Read (delegado): lee los elementos de todas las colecciones de sitios

> **Note**
>
> Note.Read.All y Sites.Read.All son necesarios solo si desea rastrear documentos. OneNote
> Si desea rastrear sitios específicos, el permiso puede restringirse a sitios específicos en lugar de a todos los sitios disponibles en el dominio. Puede configurar el permiso Sites.Selected (aplicación). Con este permiso de API, debe establecer el permiso de acceso en todos los sitios de forma explícita a través de la API de Microsoft Graph. Para obtener más información, consulta el blog de Microsoft sobre Sites.Permisos seleccionados.

  - Si usa ACL, ha agregado los siguientes permisos:

Microsoft Graph	SharePoint
• Group.Member.Read.All (aplicación): lee todas las pertenencias a grupos • Notes.Read.All (Aplicación): lee todos los blocs de notas OneNote • Sitios. FullControl.Todos (delegados): necesarios para recuperar las ACL de los documentos • Sites.Read.All (aplicación): lee los elementos de todas las colecciones de sitios • User.Read.All (aplicación): lee el perfil completo de todos los usuarios	• AllSites.Read (delegado): lee los elementos de todas las colecciones de sitios

> **Note**
>
> GroupMember.Read.All y User.Read.All son necesarios solo si el rastreador de identidades está activado.
> Si desea rastrear sitios específicos, puede restringir el permiso a sitios específicos en lugar de a todos los sitios disponibles en el dominio. Puede configurar el permiso Sites.Selected (aplicación). Con este permiso de API, debe establecer el permiso de acceso en todos los sitios de forma explícita a través de la API de Microsoft Graph. Para obtener más información, consulta el blog de Microsoft sobre Sites.Permisos seleccionados.

- Para la autenticación exclusiva con la aplicación Azure AD: clave privada y el ID de cliente que generaste después de registrarte en Azure AD. SharePoint Tenga en cuenta también el certificado X.509.
  - Si no usa ACL, ha agregado los siguientes permisos:

SharePoint
• Sites.Read.All (aplicación): se requiere para acceder a los elementos y listas de todas las colecciones de sitios

> **Note**
>
> Si desea rastrear sitios específicos, el permiso puede restringirse a sitios específicos en lugar de a todos los sitios disponibles en el dominio. Puede configurar el permiso Sites.Selected (aplicación). Con este permiso de API, debe establecer el permiso de acceso en todos los sitios de forma explícita a través de la API de Microsoft Graph. Para obtener más información, consulta el blog de Microsoft sobre Sites.Permisos seleccionados.

- Si usa ACL, ha agregado los siguientes permisos:

SharePoint
• Sitios. FullControl.All (solicitud): se requiere para recuperar las ACL de los documentos

> **Note**
>
> Si desea rastrear sitios específicos, el permiso puede restringirse a sitios específicos en lugar de a todos los sitios disponibles en el dominio. Puede configurar el permiso Sites.Selected (aplicación). Con este permiso de API, debe establecer el permiso de acceso en todos los sitios de forma explícita a través de la API de Microsoft Graph. Para obtener más información, consulta el blog de Microsoft sobre Sites.Permisos seleccionados.

- Para la autenticación SharePoint solo con aplicaciones: anote su ID de SharePoint cliente y el secreto de cliente generados al conceder el permiso a SharePoint App Only, y su ID de cliente y su secreto de cliente generados al registrar la aplicación en Azure AD. SharePoint

> **Note**
>
> SharePoint La autenticación solo para aplicaciones no es compatible con la versión de 2013. SharePoint

- (Opcional) Si está rastreando OneNote documentos y utilizando el rastreador de identidades, agregue los siguientes permisos:

Microsoft Graph
• GroupMember.Read.All (aplicación): lee todas las pertenencias a grupos • Notes.Read.All (Aplicación): lee todos los cuadernos OneNote • Sites.Read.All (aplicación): lee los elementos de todas las colecciones de sitios • User.Read.All (aplicación): lee el perfil completo de todos los usuarios

> **Note**
>
> No se requieren permisos de API para rastrear entidades mediante la autenticación básica y la autenticación solo mediante aplicaciones. SharePoint

En SharePoint Server, asegúrate de tener:

- Ha copiado las URL de la SharePoint instancia y el nombre de dominio de las SharePoint URL. El formato de la URL del host que introduzca es *`https://yourcompany/sites/mysite`*. La URL debe empezar por `https`.

> **Note**
>
> (local o en el servidor) Amazon Kendra comprueba si la información de punto final incluida AWS Secrets Manager es la misma que la información de punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a evitar el problema del suplente confuso, que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, pero utiliza Amazon Kendra como proxy para acceder al secreto configurado y realizar la acción. Si más adelante cambia la información de punto de conexión, debe crear un nuevo secreto para sincronizar esta información.

- Has desactivado la autenticación multifactor (MFA) en tu SharePoint cuenta para que no Amazon Kendra se bloquee el rastreo de tu contenido. SharePoint
- Si utilizas la autenticación solo por SharePoint aplicación para el control de acceso:
  - Se ha copiado el ID de SharePoint cliente generado al registrar App Only a nivel de sitio. El formato del ID de cliente es ClientId @TenantId. Por ejemplo, *ffa956f3-8f89-44e7-b0e4-49670756342c@888d0b57-69f1-4fb8-957f-e1f0bedf82fe*.
  - Se ha copiado el secreto de SharePoint cliente generado al registrar la aplicación solo a nivel de sitio.

  Nota: Dado que los ID de cliente y los secretos de cliente se generan para sitios individuales solo cuando se registra el SharePoint servidor para la autenticación solo con aplicaciones, solo se admite una URL de sitio para SharePoint la autenticación solo con aplicaciones.

> **Note**
>
> SharePoint La autenticación solo mediante aplicaciones no es compatible con la versión de SharePoint 2013.

- Si utiliza un ID de correo electrónico con dominio personalizado para el control de acceso:
  - Ha apuntado el valor del dominio de correo electrónico personalizado, por ejemplo: *"amazon.com"*.
- Si utiliza la autorización ID de correo electrónico con dominio del IDP, ha copiado:
  - El punto de conexión del servidor LDAP (punto de conexión del servidor LDAP, incluidos el protocolo y el número de puerto). Por ejemplo: *ldap://example.com:389*.

  - La base de búsqueda LDAP (base de búsqueda del usuario LDAP). Por ejemplo: *CN=Users,DC=sharepoint,DC=com*.
  - El nombre de usuario de LDAP y contraseña de LDAP.
- Credenciales de autenticación NTLM configuradas o credenciales de autenticación Kerberos configuradas que contienen un nombre de usuario (nombre de usuario de la SharePoint cuenta) y una contraseña (contraseña de la cuenta). SharePoint

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Guardó sus credenciales de SharePoint autenticación en un AWS Secrets Manager secreto y, si usa la API, anotó el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de SharePoint datos. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de SharePoint datos, debe proporcionar los detalles de sus SharePoint credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configurado SharePoint , Amazon Kendra consulteRequisitos previos.

Console: SharePoint Online

Para conectarse Amazon Kendra a SharePoint Online

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el SharePoint conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el SharePoint conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. Método de alojamiento: elija en líneaSharePoint .
   b. URL del sitio específicas de su SharePoint repositorio: introduzca las URL del SharePoint host. El formato de las URL del host que introduzca es *`https://yourdomain.sharepoint.com/sites/mysite`*. La URL debe empezar por el protocolo `https`. Separe las URL con una nueva línea. Puede añadir hasta 100 URL.

c. Dominio: introduce el dominio. SharePoint Por ejemplo, el dominio de la URL *https://yourdomain.sharepoint.com/sites/mysite* es *yourdomain*.

d. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

   También puede elegir el tipo de ID de usuario, ya sea el nombre principal del usuario o el correo electrónico del usuario obtenido del Portal de Azure. Si no lo especifica, se utiliza el correo electrónico de forma predeterminada.

e. Autenticación: elija entre la autenticación básica, la autenticación de OAuth 2.0, la autenticación de solo aplicación de Azure AD, la autenticación de solo aplicación o la autenticación con SharePoint token de actualización de OAuth 2.0. Puede elegir un AWS Secrets Manager secreto existente para almacenar sus credenciales de autenticación o crear un secreto.

   i. Si utiliza la autenticación básica, el secreto debe incluir un nombre secreto, un nombre SharePoint de usuario y una contraseña.

   ii. Si usa la autenticación OAuth 2.0, su secreto debe incluir el ID de SharePoint inquilino, el nombre secreto, el nombre de SharePoint usuario, la contraseña, el ID de cliente de Azure AD generado al registrarse SharePoint en Azure AD y el secreto de cliente de Azure AD generado al registrarse SharePoint en Azure AD.

   iii. Si utiliza la autenticación exclusiva para aplicaciones de Azure AD, su secreto debe incluir el ID de SharePoint inquilino, el certificado X.509 autofirmado de Azure AD, el nombre secreto, el ID de cliente de Azure AD generado al registrarse SharePoint en Azure AD y la clave privada para autenticar el conector de Azure AD.

   iv. Si utiliza la autenticación SharePoint solo por aplicación, su secreto debe incluir el ID de SharePoint inquilino, el nombre secreto, el ID de SharePoint cliente que generó al registrar App Only a nivel de inquilino, el secreto de SharePoint cliente generado cuando se registró en App Only a nivel de inquilino, el ID de cliente de Azure AD generado al registrarse SharePoint en Azure AD y el secreto de cliente de Azure AD generado SharePoint al registrarse en Azure AD.

El formato del ID de SharePoint cliente es ClientID@. TenantId Por ejemplo, *ffa956f3-8f89-44e7-b0e4-49670756342c@888d0b57-69f1-4fb8-957f-e1f0bedf82fe*.

v. Si utilizas la autenticación con token de actualización de OAuth 2.0, tu secreto debe incluir el ID de SharePoint inquilino, el nombre secreto, el ID de cliente único de Azure AD generado al registrarte SharePoint en Azure AD, el secreto de cliente de Azure AD generado al registrarte en Azure AD y el token de actualización generado SharePoint para conectarte. Amazon Kendra SharePoint

f. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

g. Rastreador de identidad: especifique si se debe activar el rastreador de identidades. Amazon Kendra El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

   También puede elegir rastrear el mapeo de grupos locales o el mapeo de grupos de Azure Active Directory.

   > **Note**
   >
   > El rastreo de mapas de grupos de AD solo está disponible para la autenticación OAuth 2.0, el token de actualización de OAuth 2.0 y la autenticación solo con aplicaciones. SharePoint

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

      i. Seleccionar entidades: elija las entidades que desee rastrear. Puede seleccionar rastrear Todas las entidades o cualquier combinación de Archivos, Archivos adjuntos, Enlaces, Páginas, Eventos, Comentarios y Datos de listas.

      ii. En Configuración adicional, en Patrones de regex de entidades: agregue patrones de expresiones regulares para los Enlaces, Páginas y Eventos a fin de incluir entidades específicas en lugar de sincronizar todos los documentos.

      iii. Patrones de expresiones regulares: agregue patrones de expresiones regulares para incluir o excluir archivos por ruta de archivo, nombre de archivo, tipo de archivo, nombre de OneNote sección y nombre de OneNote página en lugar de sincronizar todos los documentos. Puede añadir hasta 100.

      > **Note**
      >
      > OneNote El rastreo solo está disponible para la autenticación OAuth 2.0, el token de actualización de OAuth 2.0 y la autenticación solo con aplicaciones. SharePoint

   b. En Modo de sincronización, elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar el origen de datos con Amazon Kendra por primera vez, todo el contenido se sincroniza de forma predeterminada.

      - Sincronización completa: sincroniza todo el contenido independientemente del estado de sincronización anterior.
      - Sincronización de documentos nuevos o modificados: sincroniza solo los documentos nuevos o modificados.

- Sincronización de documentos nuevos, modificados o eliminados: sincroniza solo los documentos nuevos, modificados y eliminados.

c. Calendario de ejecución sincronizado, para Frecuencia: elige la frecuencia con la que deseas sincronizar el contenido de la fuente de datos y actualizar el índice.

d. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

Console: SharePoint Server

Para conectarse a Amazon Kendra SharePoint

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el SharePoint conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el SharePoint conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. Método de alojamiento: elija un SharePoint servidor.

b. Elija SharePoint la versión: elija entre SharePoint 2013, SharePoint 2016, SharePoint 2019 y SharePoint (edición de suscripción).

c. URL del sitio específicas de su SharePoint repositorio: introduzca las URL del SharePoint host. El formato de las URL del host que introduzca es *`https://yourcompany/sites/mysite`*. La URL debe empezar por el protocolo `https`. Separe las URL con una nueva línea. Puede añadir hasta 100 URL.

d. Dominio: introduce el dominio. SharePoint Por ejemplo, el dominio de la URL *`https://yourcompany/sites/mysite`* es *`yourcompany`*

e. Ubicación del certificado SSL: introduzca la Amazon S3 ruta al archivo de certificado SSL.

f. (Opcional) En Proxy web: introduzca el nombre de host del proxy web (sin el protocolo `http://` o `https://`) y el número de puerto utilizado por el protocolo de transporte de URL del host. El valor numérico del número de puerto debe estar entre 0 y 65535.

g. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

   Para el SharePoint servidor, puede elegir entre las siguientes opciones de ACL:

i. ID de correo electrónico con dominio del IDP: el ID de usuario se basa en los ID de correo electrónico cuyos dominios se obtienen del proveedor de identidad (IDP) subyacente. Usted proporciona los detalles de conexión del IDP en su Secrets Manager secreto como parte de la autenticación.

ii. ID de correo electrónico con dominio personalizado: el ID de usuario se basa en el valor del dominio de correo electrónico personalizado. Por ejemplo, *"amazon.com"*. El dominio de correo electrónico se utilizará para crear el ID de correo electrónico para el control de acceso. Debe introducir su dominio de correo electrónico personalizado.

iii. Dominio\ Usuario con dominio: el ID de usuario se crea con el formato Dominio \ ID de usuario. Debe proporcionar un nombre de dominio válido. Por ejemplo: *"sharepoint2019"* para crear un control de acceso.

h. Para la autenticación, elija la autenticación SharePoint solo por aplicación, la autenticación NTLM o la autenticación Kerberos. Puede elegir un AWS Secrets Manager secreto existente para almacenar sus credenciales de autenticación o crear un secreto.

i. Si utiliza la autenticación NTLM o la autenticación Kerberos, el secreto debe incluir un nombre secreto, un nombre de usuario y una contraseña.

   Si utiliza un ID de correo electrónico con dominio de IDP, introduzca también su:

   - Punto de conexión del servidor LDAP: punto de conexión del servidor LDAP, incluidos el protocolo y el número de puerto. Por ejemplo: *ldap://example.com:389*.
   - Base de búsqueda LDAP: base de búsqueda del usuario de LDAP. Por ejemplo: *CN=Users,DC=sharepoint,DC=com*.
   - Nombre de usuario de LDAP: su nombre de usuario de LDAP.
   - Contraseña de LDAP: su contraseña LDAP.

ii. Si utiliza la autenticación SharePoint solo por aplicación, su secreto debe incluir un nombre secreto, el ID de SharePoint cliente que generó al registrar App Only a nivel de sitio y el secreto de SharePoint cliente generado cuando se registró en App Only a nivel de sitio.

   El formato del ID de SharePoint cliente es ClientID@. TenantId Por ejemplo, *ffa956f3-8f89-44e7-b0e4-49670756342c@888d0b57-69f1-4fb8-957f-e1f0bedf82fe*.

Nota: Dado que los ID de cliente y los secretos de los clientes se generan para sitios individuales solo cuando se registra el SharePoint servidor para la autenticación solo con aplicaciones, solo se admite una URL de sitio para la autenticación solo con SharePoint aplicaciones.

Si utiliza un ID de correo electrónico con dominio de IDP, introduzca también su:

- Punto de conexión del servidor LDAP: punto de conexión del servidor LDAP, incluidos el protocolo y el número de puerto. Por ejemplo: *ldap://example.com:389*.
- Base de búsqueda LDAP: base de búsqueda del usuario de LDAP. Por ejemplo: *CN=Users,DC=sharepoint,DC=com*.
- Nombre de usuario de LDAP: su nombre de usuario de LDAP.
- Contraseña de LDAP: su contraseña LDAP.

i. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

j. Rastreador de identidades: especifique si se debe activar el rastreador Amazon Kendra de identidades. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

   También puede elegir rastrear el mapeo de grupos locales o el mapeo de grupos de Azure Active Directory.

   > **Note**
   >
   > El rastreo de la representación cartográfica de grupos de AD solo está disponible para SharePoint la autenticación mediante aplicaciones.

k. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

l. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

      i. Seleccionar entidades: elija las entidades que desee rastrear. Puede seleccionar rastrear Todas las entidades o cualquier combinación de Archivos, Archivos adjuntos, Enlaces, Páginas, Eventos y Datos de la lista.

      ii. En Configuración adicional, en Patrones de regex de entidades: agregue patrones de expresiones regulares para los Enlaces, Páginas y Eventos a fin de incluir entidades específicas en lugar de sincronizar todos los documentos.

      iii. Patrones de expresiones regulares: agregue patrones de expresiones regulares para incluir o excluir archivos por ruta de archivo, nombre de archivo, tipo de archivo, nombre de OneNotesección y nombre de OneNotepágina, en lugar de sincronizar todos los documentos. Puede añadir hasta 100.

      > **Note**
      >
      > OneNote El rastreo solo está disponible para la autenticación solo con aplicaciones. SharePoint

   b. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos con ella Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

c. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

d. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra SharePoint

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `SHAREPOINTV2` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Metadatos del punto de conexión del repositorio: especifique el `tenantID` `domain` final `siteUrls` de la SharePoint instancia.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

> **Note**
>
> El rastreador de identidades solo está disponible cuando lo `crawlAcl` configuras. `true`

- Propiedades adicionales del repositorio: especifique lo siguiente:
  - (Para Azure AD) `s3bucketName` y `s3certificateName` se utiliza para almacenar el certificado X.509 autofirmado de Azure AD.
  - Tipo de autenticación (`auth_Type`) que utiliza, ya sea `0Auth20Auth2App`,,`0Auth2Certificate`, `Basic` `0Auth2_RefreshTokenNTLM`, y. `Kerberos`
  - Versión (`version`) que usa, ya sea `Server` o`0nline`. Si usa `Server`, puede especificar adicionalmente la on`PremVersion` como `2013`, `2016`, `2019` o`SubscriptionEdition`.
- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. SharePoint

  Si usa la autenticación SharePoint en línea, puede elegir entre la autenticación básica, OAuth 2.0, solo para aplicaciones de Azure AD y solo para aplicaciones. SharePoint La siguiente es la estructura JSON mínima que debe contener el secreto para cada opción de autenticación:
  - Autenticación básica

    ```
    {
        "userName": "SharePoint account user name",
        "password": "SharePoint account password"
    }
    ```
  - Autenticación OAuth 2.0

    ```
    {
        "clientId": "client id generated when registering SharePoint with Azure AD",
        "clientSecret": "client secret generated when registering SharePoint with
     Azure AD",
        "userName": "SharePoint account user name",
        "password": "SharePoint account password"
    }
    ```
  - Autenticación App-Only de Azure AD

```
{
    "clientId": "client id generated when registering SharePoint with Azure AD",
    "privateKey": "private key to authorize connection with Azure AD"
}
```

- SharePoint Autenticación solo mediante aplicación

```
{
    "clientId": "client id generated when registering SharePoint for App Only at
 Tenant Level",
    "clientSecret": "client secret generated when registering SharePoint for App
 Only at Tenant Level",
    "adClientId": "client id generated while registering SharePoint with Azure
 AD",
    "adClientSecret": "client secret generated while registering SharePoint with
 Azure AD"
}
```

- Autenticación de token de actualización de OAuth 2.0

```
{
    "clientId": "client id generated when registering SharePoint with Azure AD",
    "clientSecret": "client secret generated when registering SharePoint with
 Azure AD",
    "refreshToken": "refresh token generated to connect to SharePoint"
}
```

Si usa SharePoint Server, puede elegir entre la autenticación SharePoint solo por aplicación, la autenticación NTLM y la autenticación Kerberos. La siguiente es la estructura JSON mínima que debe contener el secreto para cada opción de autenticación:

- SharePoint Autenticación solo por aplicación

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "clientId": "client id generated when registering SharePoint for App Only at
 Site Level",
    "clientSecret": "client secret generated when registering SharePoint for App
 Only at Site Level"
}
```

- SharePoint Autenticación solo para aplicaciones con el dominio de la autorización del IDP

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "clientId": "client id generated when registering SharePoint for App Only at
 Site Level",
    "clientSecret": "client secret generated when registering SharePoint for App
 Only at Site Level",
    "ldapUrl": "LDAP Account url eg. ldap://example.com:389",
    "baseDn": "LDAP Account base dn eg. CN=Users,DC=sharepoint,DC=com",
    "ldapUser": "LDAP account user name",
    "ldapPassword": "LDAP account password"
}
```

- (Solo para servidor) Autenticación NTLM o de Kerberos

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "userName": "SharePoint account user name",
    "password": "SharePoint account password"
}
```

- (Solo para servidor) Autenticación NTLM o de Kerberos con autorización de dominio de IDP

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "userName": "SharePoint account user name",
    "password": "SharePoint account password",
    "ldapUrl": "ldap://example.com:389",
    "baseDn": "CN=Users,DC=sharepoint,DC=com",
    "ldapUser": "LDAP account user name",
    "ldapPassword": "LDAP account password"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a su Secrets Manager secreto y llamar a las API públicas requeridas para el conector y. SharePoint Amazon Kendra Para obtener más información, consulte las IAM funciones de las fuentes SharePoint de datos.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir o excluir determinados archivos y otro contenido. OneNotes

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Asignaciones de campos: elija asignar los campos de la fuente de SharePoint datos a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para ver una lista de otras claves JSON importantes que debes configurar, consulta el esquema SharePoint de la plantilla.

Notas

- El conector admite asignaciones de campos personalizados solo para la entidad Archivos.
- Para todas las versiones SharePoint del servidor, el token ACL debe estar en minúsculas. Para la ACL Correo electrónico con dominio de IDP e ID de correo electrónico con dominio personalizado, por ejemplo: *`user@sharepoint2019.com`*. Para la ACL Dominio\Usuario con dominio, por ejemplo: *`sharepoint2013\user`*.

- El conector no admite el modo de registro de cambios ni la sincronización de contenido nuevo o modificado para SharePoint 2013.
- Si el nombre de una entidad contiene un carácter `%`, el conector omitirá estos archivos debido a las limitaciones de la API.
- OneNote El conector solo puede rastrearlo con un ID de inquilino y con OAuth 2.0, el token de actualización de OAuth 2.0 o la autenticación solo por SharePoint aplicación activada para Internet. SharePoint
- El conector rastrea la primera sección de un OneNote documento utilizando únicamente su nombre predeterminado, incluso si se cambia el nombre del documento.
- El conector rastrea los enlaces en la edición SharePoint 2019, SharePoint en línea y en la edición de suscripción, solo si se seleccionan páginas y archivos como entidades que se rastrearán además de los enlaces.
- El conector rastrea los enlaces en SharePoint 2013 y SharePoint 2016 si se selecciona Links como entidad para rastrearlos.
- El conector rastrea los archivos adjuntos y los comentarios solo cuando Datos de la lista también se selecciona como entidad que se va a rastrear.
- El conector rastrea los archivos adjuntos de eventos solo cuando Eventos también se selecciona como entidad que se va a rastrear.
- Para la versión SharePoint en línea, el token ACL estará en minúsculas. Por ejemplo, si el nombre principal del usuario es *`MaryMajor@domain .com`* en Azure Portal, el token ACL del SharePoint conector será *`marymajor@domain.com`*.
- En Identity Crawler for SharePoint Online and Server, si quieres rastrear grupos anidados, tienes que activar el rastreo local y el rastreo de grupos de AD.
- Si utilizas SharePoint Internet y el nombre principal del usuario de Azure Portal es una combinación de mayúsculas y minúsculas, la SharePoint API lo convierte internamente a minúsculas. Por este motivo, el Amazon Kendra SharePoint conector pone la ACL en minúsculas.

Microsoft SQL Server

Microsoft SQL Server es un sistema de administración de bases de datos relacionales (RDBMS) desarrollado por Microsoft. Si es un Microsoft SQL Server usuario, puede usarlo Amazon Kendra para indexar su fuente Microsoft SQL Server de datos. El conector Amazon Kendra Microsoft SQL Server de fuente de datos es compatible con MS SQL Server 2019.

Puede conectarse Amazon Kendra a su fuente Microsoft SQL Server de datos mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de la fuente de Amazon Kendra Microsoft SQL Server datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de Microsoft SQL Server datos, realice estos cambios en sus AWS cuentas Microsoft SQL Server y.

En Microsoft SQL Server, asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos.
- Ha comprobado que cada documento es único en Microsoft SQL Server y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un

índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Microsoft SQL Server en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de Microsoft SQL Server datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de Microsoft SQL Server datos, debe proporcionar los detalles de sus Microsoft SQL Server credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configuradoMicrosoft SQL Server, Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Microsoft SQL Server

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el Microsoft SQL Serverconector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el Microsoft SQL Serverconector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. En Origen, introduzca la siguiente información:
   b. Host: introduzca el nombre del host de la base de datos.
   c. Puerto: introduzca el puerto de la base de datos.
   d. Instancia: introduzca la instancia de la base de datos.
   e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.
   f. En Autenticación, introduzca la siguiente información:

- AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de Microsoft SQL Server autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

  A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

     I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Microsoft SQL Server -' se añade automáticamente a tu nombre secreto.

     II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

  B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

   - Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

> **Note**
>
> Si el nombre de una tabla incluye caracteres especiales (no alfanuméricos), debe colocar corchetes alrededor del nombre de la tabla. Por ejemplo, *`seleccione * de [] my-database-table`*

- Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
- Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
- Columna de cuerpo: proporcione el nombre de la columna del cuerpo del documento en la tabla de la base de datos.

b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

- Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.
- Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
- Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
- Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.
- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

   - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
   - Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
   - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Microsoft SQL Server

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `sqlserver`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.

  > **Note**
  >
  > Si el nombre de una tabla incluye caracteres especiales (no alfanuméricos), debe colocar corchetes alrededor del nombre de la tabla. Por ejemplo, *`seleccione * de [] my-database-table`*

- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. Microsoft SQL Server El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. Microsoft SQL Server Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Microsoft SQL Server.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para sus documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

- Asignaciones de campos: elija asignar los campos del origen de datos de Microsoft SQL Server a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice_`document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Esquema de plantilla de Microsoft SQL Server.

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.
- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.
- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Microsoft Teams

Microsoft Teams es una herramienta de colaboración empresarial para mensajería, reuniones e intercambio de archivos. Si es usuario de Microsoft Teams, puede usarlo Amazon Kendra para indexar su fuente de datos de Microsoft Teams.

Puedes conectarte Amazon Kendra a tu fuente de datos de Microsoft Teams mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Microsoft Teams, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de datos de Microsoft Teams, realice estos cambios en sus AWS cuentas y equipos de Microsoft.

En Microsoft Teams, asegúrese de que:

- Ha creado una cuenta de Microsoft Teams en Office 365.
- Ha apuntado su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
- Configuró una aplicación OAuth en el portal de Azure y anotó el ID y el secreto del cliente o las credenciales del cliente. Consulte el tutorial de Microsoft y el ejemplo de aplicación registrada para obtener más información.

> **Note**
>
> Al crear o registrar una aplicación en el portal de Azure, el ID secreto representa el valor secreto real. Debe anotar o guardar el valor secreto real inmediatamente al crear el secreto y la aplicación. Para acceder a su secreto, seleccione el nombre de la aplicación en el portal de Azure y, a continuación, vaya a la opción de menú relativa a los certificados y secretos.
> Para acceder a su ID de cliente, seleccione el nombre de su aplicación en el portal de Azure y, a continuación, vaya a la página de información general. El ID de la aplicación (cliente) es el ID del cliente.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha añadido los permisos necesarios. Puedes añadir todos los permisos o limitar el alcance seleccionando menos permisos en función de las entidades que quieras rastrear. En la siguiente tabla, se muestran los permisos a nivel de aplicación por entidad correspondiente:

Entidad	Permisos necesarios para la sincronización de datos	Permisos necesarios para la sincronización de identidades
Publicación de Canal	• ChannelMessage.Leer.Todo • Group.Read.All • User.Read • User.Read.All	TeamMember.Leer todo
Adjunto de Canal	• ChannelMessage.Leer todo • Group.Read.All • User.Read	TeamMember.Leer todo

Entidad	Permisos necesarios para la sincronización de datos	Permisos necesarios para la sincronización de identidades
	• User.Read.All	
Wiki de Canal	• Group.Read.All • User.Read • User.Read.All	TeamMember.Leer todo
Mensaje de chat	• Chat.Read.All • ChatMessage.Leer todo • ChatMember.Leer todo • User.Read • User.Read.All • Group.Read.All	TeamMember.Leer todo
Chat de reuniones	• Chat.Read.All • ChatMessage.Leer • ChatMember.Leer todo • User.Read • User.Read.All • Group.Read.All	TeamMember.Leer todo
Adjunto de chat	• Chat.Read.All • ChatMessage.Leer • ChatMember.Leer todo • User.Read • User.Read.All • Group.Read.All	TeamMember.Leer todo

Entidad	Permisos necesarios para la sincronización de datos	Permisos necesarios para la sincronización de identidades
Archivo de la reunión	• Chat.Read.All • ChatMessage.Leer todo • ChatMember.Leer todo • User.Read • User.Read.All • Group.Read.All • Files.Read.All	TeamMember.Leer todo
Reuniones del calendario	• Chat.Read.All • ChatMessage.Leer todo • ChatMember.Leer todo • User.Read • User.Read.All • Group.Read.All • Files.Read.All	TeamMember.Leer todo
Notas de la reunión	• User.Read • User.Read.All • Group.Read.All • Files.Read.All	TeamMember.Leer todo

- Comprobó que cada documento es único en Microsoft Teams y en otros orígenes que planea usar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Microsoft Teams en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes un IAM rol o secreto existente, puedes usar la consola para crear un nuevo IAM rol y Secrets Manager secreto al conectar tu fuente de datos de Microsoft Teams a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Microsoft Teams, debe proporcionar los detalles necesarios de su fuente de datos de Microsoft Teams para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Microsoft Teams para Amazon Kendra, consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Microsoft Teams

1. Inicia sesión en la Amazon Kendra consola AWS Management Console y ábrela.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector de Microsoft Teams y, a continuación, elija Agregar conector. Si usa la versión 2 (si corresponde), elija el conector Microsoft Teams con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

   c. En el idioma predeterminado: elige un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. ID de inquilino: introduzca su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.

   b. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

   c. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de autenticación de Microsoft Teams. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

A. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Microsoft Teams-' se añade automáticamente a tu nombre secreto.

B. Para el ID de cliente y el secreto del cliente: introduzca las credenciales de autenticación configuradas en Microsoft Teams en el portal de Azure.

ii. Guarde y añada su secreto.

d. Modelo de pago: puede elegir un modelo de licencia y pago para su cuenta de Microsoft Teams. Los modelos de pago del modelo A están restringidos a los modelos de licencia y pago que requieren el cumplimiento de las normas de seguridad. Los modelos de pago del modelo B son adecuados para los modelos de licencia y pago que no requieren el cumplimiento de las normas de seguridad.

e. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

f. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

g. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

h. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Sincronizar contenido: seleccione los tipos de contenido que desee rastrear. Puedes elegir rastrear el contenido del chat, los equipos y el calendario.

   b. Configuración adicional: especifique determinadas fechas de inicio y finalización del calendario, los correos electrónicos de los usuarios, los nombres de los equipos y los nombres de los canales, los archivos adjuntos y. OneNotes

   c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Cuando sincronizas tu fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

      - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
      - Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
      - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

   d. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

   e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Microsoft Teams

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `MSTEAMS` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- ID de inquilino: puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Nombre secreto del recurso de Amazon (ARN): proporcione el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su

cuenta de Microsoft Teams. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Microsoft Teams y Amazon Kendra. Para obtener más información, consulte los roles de IAM para los orígenes de datos de Microsoft Teams.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Tipos de documentos o contenido: especifique si desea rastrear los mensajes y archivos adjuntos del chat, las publicaciones y los archivos adjuntos de los canales, las wikis de los canales, el contenido del calendario, los chats de las reuniones, los archivos y las notas.
- Contenido del calendario: especifique una fecha y hora de inicio y finalización para rastrear el contenido del calendario.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos en Microsoft Teams. Puedes incluir o excluir los nombres de los equipos, los nombres de los canales, los nombres y tipos de archivos, el correo electrónico de los usuarios, OneNote las secciones y las páginas. OneNote

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Rastreador de identidades: especifique si se debe activar el rastreador Amazon Kendra de identidades. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.
- Asignaciones de campos: elija asignar los campos del origen de datos de Microsoft Teams a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que se deben configurar, consulte el esquema de plantillas de Microsoft Teams.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de datos de Microsoft Teams, consulte:

- Busque de forma inteligente en la fuente de datos de Microsoft Teams de su organización con el Amazon Kendra conector para Microsoft Teams

Microsoft Yammer

Microsoft Yammer es una herramienta de colaboración empresarial para mensajería, reuniones e intercambio de archivos. Si es usuario de Microsoft Yammer, puede usarlo Amazon Kendra para indexar su fuente de datos de Microsoft Yammer.

Puede conectarse Amazon Kendra a la fuente de datos de Microsoft Yammer mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Microsoft Yammer, consulteSolución de problemas con los orígenes de datos.

Características admitidas

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de datos de Microsoft Yammer, realice estos cambios en su cuenta y AWS en su cuenta de Microsoft Yammer.

En Microsoft Yammer, asegúrese de que:

- Creó una cuenta administrativa de Microsoft Yammer en Office 365.
- Apuntó su nombre de usuario y contraseña de Microsoft Yammer.
- Ha apuntado su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en su aplicación OAuth.
- Configuré una aplicación OAuth en el portal de Azure y anoté el ID del cliente y el secreto del cliente o las credenciales del cliente. Consulte el tutorial de Microsoft y el ejemplo de aplicación registrada para obtener más información.

> **Note**
>
> Al crear o registrar una aplicación en el portal de Azure, el ID secreto representa el valor secreto real. Debe anotar o guardar el valor secreto real inmediatamente al crear el secreto y la aplicación. Para acceder a su secreto, seleccione el nombre de la aplicación en el portal de Azure y, a continuación, vaya a la opción de menú relativa a los certificados y secretos.
> Para acceder a su ID de cliente, seleccione el nombre de su aplicación en el portal de Azure y, a continuación, vaya a la página de información general. El ID de la aplicación (cliente) es el ID del cliente.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Comprobó que cada documento es único en Microsoft Yammer y en otros orígenes que planea usar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Microsoft Yammer en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y Secrets Manager secreto al conectar la fuente de datos de Microsoft Yammer a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a la fuente de datos de Microsoft Yammer, debe proporcionar los detalles necesarios de la fuente de datos de Microsoft Yammer para que Amazon Kendra pueda acceder a los datos. Si aún no ha configurado Microsoft Yammer para Amazon Kendra, consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Microsoft Yammer

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector de Microsoft Yammer y, a continuación, elija Agregar conector. Si usa la versión 2 (si corresponde), elija el conector Microsoft Yammer con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

b. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de autenticación de Microsoft Yammer. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

   i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

      A. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Microsoft Yammer-' se añade automáticamente a tu nombre secreto.

      B. Para Nombre de usuario y Contraseña: introduzca su nombre de usuario y contraseña de Microsoft Yammer.

      C. Para ID de cliente, secreto de cliente: introduzca las credenciales de autenticación configuradas en Microsoft Yammer en el portal de Azure.

   ii. Guarde y añada su secreto.

c. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

d. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

e. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

f. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. Desde la fecha: especifique la fecha para empezar a rastrear los datos en Microsoft Yammer.

b. Sincronizar contenido: seleccione el tipo de contenido que desee rastrear. Por ejemplo, mensajes públicos, mensajes privados y archivos adjuntos.

c. Configuración adicional: especifique los nombres de las comunidades que desee rastrear y, además, utilice patrones de expresiones regulares para incluir o excluir cierto contenido.

d. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no eliges la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

e. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

f. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Microsoft Yammer

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `YAMMER` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no eliges la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Nombre de recurso secreto de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta de Microsoft Yammer. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "username": "user name",
    "password": "password",
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Microsoft Yammer y. Amazon Kendra Para obtener más información, consulte los roles de IAM para los orígenes de datos de Microsoft Yammer.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Tipos de documentos o contenido: especifique si desea rastrear el contenido de la comunidad, los mensajes y archivos adjuntos y los mensajes privados.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinado contenido.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Rastreador de identidad: especifique si se debe activar el rastreador de identidad. Amazon Kendra El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden buscar públicamente todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.
- Asignaciones de campos: elija asignar los campos del origen de datos de Microsoft Yammer a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el

nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice_`document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que se deben configurar, consulte el esquema de plantillas de Microsoft Yammer.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de datos de Microsoft Yammer, consulte:

- Presentamos el conector Yammer para Amazon Kendra

MySQL

MySQL es un sistema de administración de bases de datos relacionales de código abierto. Si es un MySQL usuario, puede usarlo Amazon Kendra para indexar su fuente MySQL de datos. El conector Amazon Kendra MySQL de fuente de datos es compatible con MySQL 8.0. 21.

Puede conectarse Amazon Kendra a su fuente MySQL de datos mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de la fuente de Amazon Kendra MySQL datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario

- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de MySQL datos, realice estos cambios en sus AWS cuentas MySQL y.

En MySQL, asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos.
- Ha comprobado que cada documento es único en MySQL y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de MySQL en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de MySQL datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de MySQL datos, debe proporcionar los detalles de sus MySQL credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configuradoMySQL, Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a MySQL

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el MySQLconector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el MySQLconector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. En Origen, introduzca la siguiente información:

b. Host: introduzca el nombre del host de la base de datos.

c. Puerto: introduzca el puerto de la base de datos.

d. Instancia: introduzca la instancia de la base de datos.

e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.

f. En Autenticación, introduzca la siguiente información:

- AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de MySQL autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

  A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

     I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-MySQL -' se añade automáticamente a tu nombre secreto.

     II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

  B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

   - Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
   - Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
   - Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
   - Columna de cuerpo: proporcione el nombre de la columna de cuerpo del documento en la tabla de la base de datos.

   b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

   - Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.
   - Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
   - Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
   - Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.

- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para rastrear los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra MySQL

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `mySql`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de

datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. MySQL El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. MySQL Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de MySQL.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para sus documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

- Asignaciones de campos: elija asignar los campos del origen de datos de MySQL a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.
- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.
- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Oracle Database

Oracle Database es un sistema de administración de bases de datos. Si es un Oracle Database usuario, puede usarlo Amazon Kendra para indexar su fuente Oracle Database de datos. El conector Amazon Kendra Oracle Database de fuente de datos es compatible con Oracle Database 18c, 19c y 21c.

Puede conectarse Amazon Kendra a su fuente de Oracle Database datos mediante la Amazon Kendra consola y la API. TemplateConfiguration

Para solucionar problemas del conector de la fuente de Amazon Kendra Oracle Database datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de Oracle Database datos, realice estos cambios en sus AWS cuentas Oracle Database y.

En Oracle Database, asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos.
- Ha comprobado que cada documento es único en Oracle Database y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Oracle Database en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de Oracle Database datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de Oracle Database datos, debe proporcionar los detalles de sus Oracle Database credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configuradoOracle Database, Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a Oracle Database

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el Oracle Databaseconector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el Oracle Databaseconector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. En Origen, introduzca la siguiente información:
   b. Host: introduzca el nombre del host de la base de datos.
   c. Puerto: introduzca el puerto de la base de datos.
   d. Instancia: introduzca la instancia de la base de datos.
   e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.
   f. En Autenticación, introduzca la siguiente información:
      - AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de Oracle Database autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Oracle Database -' se añade automáticamente a tu nombre secreto.

II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

- Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
- Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
- Columna de cuerpo: proporcione el nombre de la columna de cuerpo del documento en la tabla de la base de datos.

b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

- Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.
- Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
- Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
- Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.
- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra Oracle Database

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `oracle`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra

por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:

- `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. Oracle Database El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. Oracle Database Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Oracle Database.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para sus documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos del origen de datos de Oracle Database a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Esquema de plantilla de Oracle Database.

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.

- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.
- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

PostgreSQL

PostgreSQL es un sistema de administración de bases de datos de código abierto. Si es un PostgreSQL usuario, puede usarlo Amazon Kendra para indexar su fuente PostgreSQL de datos. El conector Amazon Kendra PostgreSQL de fuente de datos es compatible con PostgreSQL 9.6.

Puede conectarse Amazon Kendra a su fuente de PostgreSQL datos mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de la fuente de Amazon Kendra PostgreSQL datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Notas

Características admitidas

- Asignaciones de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de PostgreSQL datos, realice estos cambios en sus AWS cuentas PostgreSQL y.

En PostgreSQL, asegúrese de que:

- Ha anotado el nombre de usuario y contraseña de la base de datos.

> **Important**
>
> Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.

- Ha copiado la URL, el puerto y la instancia del host de la base de datos.
- Ha comprobado que cada documento es único en PostgreSQL y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de PostgreSQL en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de PostgreSQL datos Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de PostgreSQL datos, debe proporcionar los detalles de sus PostgreSQL credenciales para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configuradoPostgreSQL, Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a PostgreSQL

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el PostgreSQLconector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el PostgreSQLconector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. En Origen, introduzca la siguiente información:

b. Host: introduzca el nombre del host de la base de datos.

c. Puerto: introduzca el puerto de la base de datos.

d. Instancia: introduzca la instancia de la base de datos.

e. Habilitar la ubicación del certificado SSL: elija introducir la Amazon S3 ruta al archivo de certificado SSL.

f. En Autenticación, introduzca la siguiente información:

- AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de PostgreSQL autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

  A. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

     I. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-PostgreSQL -' se añade automáticamente a tu nombre secreto.

     II. Para el nombre de usuario y la contraseña de la base de datos: introduzca los valores de las credenciales de autenticación que ha copiado de la base de datos.

  B. Seleccione Guardar.

g. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

h. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Ámbito de sincronización, seleccione de entre las siguientes opciones:

      - Consulta SQL: introduzca instrucciones de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
      - Columna de clave principal: proporcione la clave principal de la tabla de la base de datos. Esto identifica una tabla dentro de la base de datos.
      - Columna de título: proporcione el nombre de la columna del título del documento en la tabla de la base de datos.
      - Columna de cuerpo: proporcione el nombre de la columna de cuerpo del documento en la tabla de la base de datos.

   b. En Configuración adicional (opcional), elija una de las siguientes opciones para sincronizar contenido específico en lugar de sincronizar todos los archivos:

      - Columnas de detección de cambios: introduzca los nombres de las columnas que se Amazon Kendra utilizarán para detectar cambios en el contenido. Amazon Kendra volverá a indexar el contenido cuando se produzca un cambio en alguna de estas columnas.
      - Columna de ID de usuario: introduzca el nombre de la columna que contiene los ID de usuario a los que se dará acceso al contenido.
      - Columna de grupos: introduzca el nombre de la columna que contiene los grupos a los que se dará acceso al contenido.
      - Columna de URL de origen: introduzca el nombre de la columna que contiene las URL de origen que se van a indexar.

- Columna de marcas de tiempo: introduzca el nombre de la columna que contiene las marcas de tiempo. Amazon Kendra utiliza la información de las marcas de tiempo para detectar cambios en el contenido y sincronizar solo el contenido modificado.
- Columna de zonas horarias: introduzca el nombre de la columna que contiene las zonas horarias del contenido que se va a rastrear.
- Formato de marcas temporales: introduzca el nombre de la columna que contiene los formatos de marcas temporales que se deben utilizar para detectar cambios en el contenido y volver a sincronizar su contenido.

c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

d. En Programa de ejecución de sincronización, en Frecuencia: la frecuencia con la que Amazon Kendra se sincronizará con el origen de datos.

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione uno de los campos de fuente de datos predeterminados generados (ID de documento, títulos de documentos y URL de origen) que desee mapear para indexar Amazon Kendra .

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra PostgreSQL

Debe especificar lo siguiente mediante la TemplateConfigurationAPI:

- Fuente de datos: especifique el tipo de fuente de datos como `JDBC` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- Tipo de base de datos: debe especificar el tipo de base de datos como `postgresql`.
- Consulta SQL: especifique las sentencias de consulta SQL, como las operaciones SELECT y JOIN. Las consultas de SQL deben ser inferiores a 32 KB. Amazon Kendra rastreará todo el contenido de la base de datos que coincida con su consulta.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de

datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. PostgreSQL El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. PostgreSQL Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de PostgreSQL.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir contenido específico mediante los identificadores de usuario, los grupos, las direcciones URL de origen, las marcas temporales y las zonas horarias.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para sus documentos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

- Asignaciones de campos: elija asignar los campos del origen de datos de PostgreSQL a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte Esquema de plantilla de PostgreSQL.

Notas

- No se realizará un seguimiento de las filas de la base de datos eliminadas cuando se Amazon Kendra compruebe si hay contenido actualizado.
- El tamaño de los nombres y valores de los campos de una fila de la base de datos no puede superar los 400 KB.
- Si tiene una gran cantidad de datos en la fuente de datos de la base de datos y no desea Amazon Kendra indexar todo el contenido de la base de datos después de la primera sincronización, puede optar por sincronizar solo los documentos nuevos, modificados o eliminados.
- Como práctica recomendada, proporcione credenciales de base Amazon Kendra de datos de solo lectura.
- Como práctica recomendada, evite añadir tablas con datos confidenciales o información de identificación personal (PII).

Quip

Quip es un software de productividad colaborativa que ofrece capacidades de creación de documentos en tiempo real. Puede usarlo Amazon Kendra para indexar sus carpetas, archivos, comentarios de archivos, salas de chat y archivos adjuntos de Quip.

Puede conectarse Amazon Kendra a su fuente de datos de Quip mediante la Amazon Kendra consola y la API. QuipConfiguration

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Quip, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra El conector de fuente de datos Quip admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar su fuente de datos de Quip, realice estos cambios en Quip y AWS en sus cuentas.

En Quip, asegúrese de que:

- Tiene una cuenta de Quip con permisos administrativos.
- Ha creado credenciales de autenticación de Quip que incluyen un token de acceso personal. El token se utiliza como su credencial de autenticación almacenada en un secreto. AWS Secrets Manager Consulte la Documentación de Quip sobre la autenticación para obtener más información.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha copiado el dominio de su sitio de Quip. Por ejemplo, *`https://quip-company.quipdomain.com/browse`*, donde *`quipdomain`* es el dominio.
- Ha comprobado que cada documento es único en Quip y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Quip en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de datos de Quip. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Quip, debe proporcionar los detalles necesarios de su fuente de datos de Quip para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Quip para Amazon Kendra, consulte. Requisitos previos

Console

Para conectarse Amazon Kendra a Quip

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector Quip y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el conector Quip con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. Nombre de dominio de Quip: introduzca el Quip que ha copiado de la cuenta de Quip.

   b. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo para almacenar sus Secrets Manager credenciales de autenticación de Quip. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

      i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

         A. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Quip» se añade automáticamente a su nombre secreto.

         B. Token de Quip: introduzca el Quip configurado para el acceso personal de Quip.

      ii. Añada y guarde su secreto.

   c. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

   d. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

   > **Note**
   >
   > IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

   e. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Agregar los ID de carpeta de Quip para rastrear: los ID de las carpetas de Quip que desea rastrear.

   > **Note**
   >
   > Para rastrear una carpeta raíz, incluidas todas las subcarpetas y documentos que contiene, añada el ID de la carpeta raíz. Para rastrear subcarpetas específicas, añade los ID de las subcarpetas específicas.

b. Configuración adicional (tipos de contenido): introduzca los tipos de contenido que desee rastrear.

c. Patrones regex: patrones de expresiones regulares para incluir o excluir determinados archivos. Puede agregar hasta 100 patrones.

d. Calendario de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice

e. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

a. Seleccione uno de los campos de fuente de datos predeterminados generados que desee mapear para Amazon Kendra indexar.

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Quip

Debe especificar lo siguiente mediante la QuipConfiguration API:

- Dominio del sitio de Quip: por ejemplo, *https://quip-company.quipdomain.com/browse*, donde *quipdomain* es el dominio.
- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta de Quip. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "accessToken": "token"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas

requeridas para el conector de Quip y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Quip.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique `VpcConfiguration` como parte de la configuración del origen de datos. Consulte Configuración de Amazon Kendra para utilizar una VPC.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Carpetas: especifique las carpetas y subcarpetas de Quip que desee indexar

> **Note**
>
> Para rastrear una carpeta raíz, incluidas todas las subcarpetas y documentos que contiene, introduzca el ID de la carpeta raíz. Para rastrear subcarpetas específicas, añade los ID de las subcarpetas específicas.

- Archivos adjuntos, salas de chat, comentarios de archivos: elija si desea incluir el rastreo de los archivos adjuntos, el contenido de las salas de chat y los comentarios de los archivos.
- Filtrado del contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
- Asignaciones de campos: elija asignar los campos del origen de datos de Quip a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice `_document_body`. Todos los demás campos son opcionales.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de datos de Quip, consulte:

- Busque información en los documentos de Quip mediante la búsqueda inteligente mediante el conector Quip para Amazon Kendra

Salesforce

Salesforce es una herramienta de gestión de relaciones con los clientes (CRM) para administrar los equipos de soporte, ventas y marketing. Puede usarlo Amazon Kendra para indexar sus objetos estándar de Salesforce e incluso objetos personalizados.

Puede conectarse Amazon Kendra a su fuente de datos de Salesforce mediante la Amazon Kendra consola, la TemplateConfigurationAPI o la SalesforceConfigurationAPI.

Amazon Kendra tiene dos versiones del conector de Salesforce. Las características compatibles de cada versión incluyen:

Conector de Salesforce V1.0/API SalesforceConfiguration

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión

Conector Salesforce V2.0/ API TemplateConfiguration

- Asignaciones de campo

- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

> **Note**
>
> Está previsto que el soporte para el conector SalesforceConfiguration V1.0/API de Salesforce finalice en 2023. Recomendamos migrar o utilizar el conector V2.0/API de Salesforce. TemplateConfiguration

Para solucionar problemas de su conector de fuente de datos de Amazon Kendra Salesforce, consulte. Solución de problemas con los orígenes de datos

Temas

- Salesforce Connector V1.0
- Salesforce Connector V2.0

Salesforce Connector V1.0

Salesforce es una herramienta de gestión de relaciones con los clientes (CRM) para administrar los equipos de soporte, ventas y marketing. Puede utilizarlo Amazon Kendra para indexar sus objetos estándar de Salesforce e incluso objetos personalizados.

> **Important**
>
> Amazon Kendra utiliza la versión 48 de la API de Salesforce. La API de Salesforce limita la cantidad de solicitudes que se pueden realizar por día. Si Salesforce supera esas solicitudes, lo volverá a intentar hasta que pueda continuar.

> **Note**
>
> Está previsto que el soporte para el conector SalesforceConfiguration V1.0/API de Salesforce finalice en 2023. Recomendamos migrar o utilizar el conector V2.0/API de Salesforce. TemplateConfiguration

Para solucionar problemas de su conector de fuente de datos de Amazon Kendra Salesforce, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión

Características admitidas

Amazon Kendra El conector de fuente de datos de Salesforce admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar su fuente de datos de Salesforce, realice estos cambios en Salesforce y AWS en sus cuentas.

En Salesforce, asegúrese de que:

- Ha creado una cuenta de Salesforce y ha anotado el nombre de usuario y la contraseña que utiliza para conectarse a Salesforce.
- Ha creado una cuenta de la aplicación Salesforce Connected con OAuth activada y ha copiado la clave de consumidor (ID de cliente) y el secreto de consumidor (secreto de cliente) asignados a la aplicación Salesforce Connected. El ID de cliente y el secreto del cliente se utilizan como credenciales de autenticación almacenadas en un AWS Secrets Manager secreto. Consulte la Documentación de Salesforce sobre aplicaciones conectadas para obtener más información.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha copiado el token de seguridad de Salesforce asociado a la cuenta utilizada para conectarse a Salesforce.
- Ha copiado la URL de la instancia de Salesforce que desea indexar. Normalmente, es *https://<company>.salesforce.com/*. El servidor debe ejecutar una aplicación conectada de Salesforce.
- Se agregaron credenciales a su servidor de Salesforce para un usuario con acceso de solo lectura a Salesforce. Para ello, clonó el ReadOnly perfil y, a continuación, agregó los permisos Ver todos los datos y Administrar artículos. Estas credenciales identifican al usuario que realiza la conexión y a la aplicación conectada de Salesforce a la que se conecta. Amazon Kendra
- Ha comprobado que cada documento es único en Salesforce y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En la suya Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Salesforce en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de datos de Salesforce. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Salesforce, debe proporcionar los detalles necesarios de su fuente de datos de Salesforce para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Salesforce para consultarlo. Amazon Kendra Requisitos previos

Console

Para conectarse a Amazon Kendra Salesforce

1. Inicie sesión en la consola AWS de administración y abra la Amazon Kendra consola.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar origen de datos, seleccione Salesforce Connector V1.0 y, a continuación, seleccione Añadir conector.
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

a. Nombre de origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. Idioma predeterminado: un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos anula el idioma seleccionado.

d. Agregar nueva etiqueta: etiquetas para buscar y filtrar los recursos o hacer un seguimiento de los costos compartidos.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. URL de Salesforce: introduzca la URL de la instancia para el sitio de Salesforce que desea indexar.

b. En Tipo de autenticación, elija entre Existente y Nuevo para almacenar las credenciales de autenticación de Salesforce. Si decide crear un secreto nuevo, se abrirá una ventana AWS Secrets Manager secreta.

- Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

A. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Salesforce-» se añade automáticamente a su nombre secreto.

B. En Nombre de usuario, Contraseña, Token de seguridad, Clave de consumidor, Secreto del consumidor y URL de autenticación, introduzca los valores de las credenciales de autenticación que creó en la cuenta de Salesforce.

C. Seleccione Guardar autenticación.

c. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

d. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Rastrear archivos adjuntos: seleccione esta opción para rastrear todos los objetos, artículos y feeds adjuntos.

   b. En Objetos estándar, Artículos de conocimiento y Fuente de chat, seleccione las entidades o los tipos de contenido de Salesforce que desee rastrear.

   > **Note**
   >
   > Debe proporcionar información de configuración para indexar al menos uno de los objetos estándar, artículos de conocimiento o fuentes de chat. Si decide rastrear los Artículos de conocimiento, debe especificar los tipos de artículos de conocimiento que desea indexar, el nombre de los artículos y si desea indexar los campos estándar de todos los artículos de conocimiento o solo los campos de un tipo de artículo personalizado. Si decide indexar artículos personalizados, debe especificar el nombre interno del tipo de artículo. Puede especificar hasta 10 tipos de artículos.

   c. Frecuencia: la frecuencia con la Amazon Kendra que se sincronizará con la fuente de datos.

   d. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Para ver el artículo de conocimiento estándar, los archivos adjuntos a objetos estándar y las asignaciones de campos sugeridas adicionales, seleccione entre los campos de fuentes de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

   > **Note**
   >
   > Es necesaria una asignación de índice a `_document_body`. No puede cambiar la asignación entre el campo `Salesforce ID` y el campo `_document_id` de Amazon Kendra .

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Salesforce Amazon Kendra

Debe especificar lo siguiente en la SalesforceConfigurationAPI:

- URL del servidor: la URL de la instancia para el sitio de Salesforce que desea indexar.
- Nombre secreto del recurso de Amazon (ARN): proporcione el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta de Salesforce. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "authenticationUrl": "OAUTH endpoint that Amazon Kendra connects to get an
 OAUTH token",
    "consumerKey": "Application public key generated when you created your
 Salesforce application",
    "consumerSecret": "Application private key generated when you created your
 Salesforce application.",
    "password": "Password associated with the user logging in to the Salesforce
 instance",
    "securityToken": "Token associated with the user account logging in to the
 Salesforce instance",
    "username": "User name of the user logging in to the Salesforce instance"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Salesforce y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Salesforce.
- Debe proporcionar información de configuración para indexar al menos uno de los objetos estándar, artículos de conocimiento o fuentes de chat.

- Objetos estándar: si decide rastrear los Objetos estándar, debe especificar el nombre del objeto estándar y el nombre del campo de la tabla de objetos estándar que contiene el contenido del documento.
- Artículos de conocimiento: si decide rastrear los Artículos de conocimiento, debe especificar los tipos de artículos de conocimiento que desea indexar, los estados de los artículos de conocimiento que desea indexar y si quiere indexar los campos estándar de todos los artículos de conocimiento o solo los campos de un tipo de artículo personalizado.
- Fuentes de Chatter: si decide rastrear las fuentes de Chatter, debe especificar el nombre de la columna de la FeedItem tabla de Salesforce que contiene el contenido que se va a indexar.

También puede añadir las siguientes características opcionales:

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos adjuntos.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Asignaciones de campos: elija asignar los campos del origen de datos de Salesforce a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio para Amazon Kendra poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice `_document_body`. Todos los demás campos son opcionales.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de sus documentos, si tiene una ACL para ellos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

Salesforce Connector V2.0

Salesforce es una herramienta de gestión de relaciones con los clientes (CRM) para administrar los equipos de soporte, ventas y marketing. Puede utilizarlos Amazon Kendra para indexar sus objetos estándar de Salesforce e incluso objetos personalizados.

El conector de fuentes de datos de Amazon Kendra Salesforce es compatible con las siguientes ediciones de Salesforce: Developer Edition y Enterprise Edition.

> **Note**
>
> Está previsto que el soporte para el conector SalesforceConfiguration V1.0/API de Salesforce finalice en 2023. Recomendamos migrar o utilizar el conector V2.0/API de Salesforce. TemplateConfiguration

Para solucionar problemas de su conector de fuente de datos de Amazon Kendra Salesforce, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra El conector de fuente de datos de Salesforce admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión

- Sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar su fuente de datos de Salesforce, realice estos cambios en Salesforce y en sus cuentas. AWS

En Salesforce, asegúrese de que:

- Ha creado una cuenta administrativa de Salesforce y ha anotado el nombre de usuario y la contraseña que utiliza para conectarse a Salesforce.
- Ha copiado el token de seguridad de Salesforce asociado a la cuenta utilizada para conectarse a Salesforce.
- Ha creado una cuenta de la aplicación Salesforce Connected con OAuth activada y ha copiado la clave de consumidor (ID de cliente) y el secreto de consumidor (secreto de cliente) asignados a la aplicación Salesforce Connected. El ID de cliente y el secreto del cliente se utilizan como credenciales de autenticación almacenadas en un AWS Secrets Manager secreto. Consulte la Documentación de Salesforce sobre aplicaciones conectadas para obtener más información.

  > **Note**
  >
  > Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Ha copiado la URL de la instancia de Salesforce que desea indexar. Normalmente, es *`https://<company>.salesforce.com/`*. El servidor debe ejecutar una aplicación conectada de Salesforce.
- Se agregaron credenciales a su servidor de Salesforce para un usuario con acceso de solo lectura a Salesforce. Para ello, clonó el ReadOnly perfil y, a continuación, agregó los permisos Ver todos los datos y Administrar artículos. Estas credenciales identifican al usuario que realiza la conexión y a la aplicación conectada de Salesforce a la que se conecta. Amazon Kendra
- Ha comprobado que cada documento es único en Salesforce y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no

debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En la suya Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Salesforce en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de datos de Salesforce. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Salesforce, debe proporcionar los detalles necesarios de su fuente de datos de Salesforce para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Salesforce para consultarlo. Amazon Kendra Requisitos previos

Console

Para conectarse Amazon Kendra a Salesforce:

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector de Salesforce y, a continuación, selecciona Añadir conector. Si usa la versión 2 (si corresponde), elija el conector de Salesforce con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. URL de Salesforce: introduzca la URL de la instancia para el sitio de Salesforce que desea indexar.
   b. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.
   c. Introduzca un secreto existente o, si crea uno nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

- Autenticación: introduzca la siguiente información en la ventana Crear un AWS Secrets Manager secreto:

  A. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-Salesforce-' se añade automáticamente a su nombre secreto.

  B. En Nombre de usuario, Contraseña, Token de seguridad, Clave de consumidor, Secreto del consumidor y URL de autenticación, introduzca los valores de las credenciales de autenticación que generó y descargó de la cuenta de Salesforce.

  > **Note**
  >
  > Si utiliza Salesforce Developer Edition, utilice **`https://login.salesforce.com/services/oauth2/token`** la URL de inicio de sesión de My Domain (por ejemplo, ***`https://MyCompany.my.salesforce.com`***) como URL de autenticación. Si utiliza Salesforce Sandbox Edition, utilice **`https://test.salesforce.com/services/oauth2/token`** o la URL de inicio de sesión de My Domain (por ejemplo, ***`MyDomainName--`***.sandbox.my.salesforce.com) como URL de autenticación. SandboxName

  C. Seleccione Guardar autenticación.

d. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

e. Rastreador de identidad: especifique si se debe activar el rastreador de identidad. Amazon Kendra El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

f. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

g. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. En Rastrear archivos adjuntos: seleccione esta opción para rastrear todos los objetos de Salesforce adjuntos.

   b. En Objetos estándar, Objetos estándar con archivos adjuntos, Objetos estándar sin archivos adjuntos y Artículos de conocimiento: seleccione las entidades o los tipos de contenido de Salesforce que desee rastrear.

   c. Debe proporcionar información de configuración para indexar al menos uno de los objetos estándar, artículos de conocimiento o fuentes de chat. Si decide rastrear los Artículos de conocimiento, debe especificar los tipos de artículos de conocimiento que desea indexar. Puede elegir entre artículos publicados, archivados, borradores y archivos adjuntos.

      Filtro de regex: especifique un patrón de regex para incluir elementos específicos del catálogo.

8. En Configuración adicional:

   - Información de ACL: todas las listas de control de acceso se incluyen de forma predeterminada. Al anular la selección de una lista de control de acceso, todos los archivos de esa categoría serán públicos.
   - Patrones regex: añada patrones de expresiones regulares para incluir o excluir determinados archivos. Puede agregar hasta 100 patrones.

   Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización

completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

9. Elija Siguiente.
10. En la página Establecer asignaciones de campos, especifique la siguiente información:

    a. Para ver el artículo de información estándar, los archivos adjuntos a objetos estándar y otras sugerencias de mapeo de campos: seleccione entre los campos de fuentes de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

    > **Note**
    >
    > Es necesaria una asignación de índice a `_document_body`. No puede cambiar la asignación entre el campo `Salesforce ID` y el campo `_document_id` de Amazon Kendra . Puedes asignar cualquier campo de Salesforce al título o al cuerpo del documento (campos de índice reservados o predeterminados de Amazon Kendra).
    > Si asigna cualquier campo de Salesforce a los campos de título y cuerpo del documento de Amazon Kendra, Amazon Kendra utilizará los datos de los campos de título y cuerpo del documento en las respuestas de búsqueda.

    b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

    c. Elija Siguiente.

11. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Salesforce Amazon Kendra

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `SALESFORCEV2` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- URL del host: especifique la URL del host de la instancia de Salesforce.
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Nombre secreto del recurso de Amazon (ARN): proporcione el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta de Salesforce. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "authenticationUrl": "OAUTH endpoint that Amazon Kendra connects to get an
 OAUTH token",
    "consumerKey": "Application public key generated when you created your
 Salesforce application",
    "consumerSecret": "Application private key generated when you created your
 Salesforce application",
    "password": "Password associated with the user logging in to the Salesforce
 instance",
    "securityToken": "Token associated with the user account logging in to the
 Salesforce instance",
    "username": "User name of the user logging in to the Salesforce instance"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Salesforce y. Amazon Kendra Para obtener más información, consulte Roles de IAM para orígenes de datos de Salesforce.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir o excluir determinados documentos, cuentas, campañas, casos, contactos, clientes potenciales, oportunidades, soluciones, tareas, grupos, chats y archivos de entidades personalizados.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Rastreador de identidades: especifique si se debe activar el rastreador de identidades. Amazon Kendra El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.
- Asignaciones de campos: elija asignar los campos del origen de datos de Salesforce a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

> **Note**
>
> Es necesaria una asignación de índice a `_document_body`. No puede cambiar la asignación entre el campo `Salesforce ID` y el campo `_document_id` de Amazon Kendra . Puedes asignar cualquier campo de Salesforce al título o al cuerpo del documento (campos de índice reservados o predeterminados de Amazon Kendra). Si asigna cualquier campo de Salesforce a los campos de título y cuerpo del documento de Amazon Kendra, Amazon Kendra utilizará los datos de los campos de título y cuerpo del documento en las respuestas de búsqueda.

Para ver una lista de otras claves JSON importantes que debe configurar, consulte el esquema de plantillas de Salesforce.

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Salesforce, consulte:

- Anunciamos el conector de Salesforce actualizado (V2) para Amazon Kendra

ServiceNow

ServiceNow proporciona un sistema de administración de servicios basado en la nube para crear y administrar flujos de trabajo a nivel de organización, como los servicios de TI, los sistemas de venta de entradas y el soporte. Puede utilizarlo Amazon Kendra para indexar sus ServiceNow catálogos, artículos de conocimiento, incidentes y sus archivos adjuntos.

Puede conectarse Amazon Kendra a su fuente de ServiceNow datos mediante la Amazon Kendra consola, la TemplateConfigurationAPI o la ServiceNowConfigurationAPI.

Amazon Kendra tiene dos versiones del ServiceNow conector. Las características compatibles de cada versión incluyen:

ServiceNow conector V1.0/API ServiceNowConfiguration

- Asignaciones de campo
- ServiceNow versiones de instancia: Londres, Otras
- Filtros de inclusión/exclusión

ServiceNow conector V2.0/API TemplateConfiguration

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronización de contenido completa e incremental
- ServiceNow versiones de instancia: Roma, Sandiego, Tokio, otras
- Nube privada virtual (VPC)

> **Note**
>
> Está previsto que el soporte para el ServiceNow conector ServiceNowConfiguration V1.0/API finalice en 2023. Recomendamos migrar o utilizar el ServiceNow conector V2.0/API. TemplateConfiguration

Para solucionar problemas del conector de la fuente de Amazon Kendra ServiceNow datos, consulte. Solución de problemas con los orígenes de datos

Temas

- ServiceNow conector V1.0
- ServiceNow conector V2.0
- Especificar los documentos que se van a indexar con una consulta

ServiceNow conector V1.0

ServiceNow proporciona un sistema de administración de servicios basado en la nube para crear y administrar flujos de trabajo a nivel de organización, como los servicios de TI, los sistemas de emisión de tickets y el soporte. Puede utilizarlo Amazon Kendra para indexar sus ServiceNow catálogos, artículos de conocimiento y sus anexos.

> **Note**
>
> Está previsto que el soporte para el ServiceNow conector ServiceNowConfiguration V1.0/API finalice en 2023. Recomendamos migrar o utilizar el ServiceNow conector V2.0/API. TemplateConfiguration

Para solucionar problemas del conector de la fuente de Amazon Kendra ServiceNow datos, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión

- Más información

Características admitidas

Amazon Kendra ServiceNow el conector de fuente de datos admite las siguientes funciones:

- ServiceNow versiones de instancia: Londres, Otras
- Patrones de inclusión/exclusión: catálogos de servicios, artículos de conocimiento y sus archivos adjuntos

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de ServiceNow datos, realiza estos cambios en tus AWS cuentas ServiceNow y.

En ServiceNow, asegúrate de tener:

- Creó una cuenta de ServiceNow administrador y creó una ServiceNow instancia.
- Ha copiado el host de la URL de la ServiceNow instancia. Por ejemplo, si la URL de la instancia es *`https://your-domain.service-now.com`*, el formato de la URL del host que introduzca es *`your-domain.service-now.com`*.
- Apuntó sus credenciales de autenticación básicas, que incluyen un nombre de usuario y una contraseña Amazon Kendra para poder conectarse a la ServiceNow instancia.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Opcional: configuraste un token de credenciales de OAuth 2.0 que puede identificar Amazon Kendra y generar un nombre de usuario, una contraseña, un ID de cliente y un secreto de cliente. El nombre de usuario y la contraseña deben proporcionar acceso a la base de ServiceNow conocimientos y al catálogo de servicios. Consulte ServiceNow la documentación sobre la autenticación de OAuth 2.0 para obtener más información.
- Ha agregado los siguientes permisos:

- kb_category
- kb_knowledge
- kb_knowledge_base
- kb_uc_cannot_read_mtom
- kb_uc_can_read_mtom
- sc_catalog
- sc_category
- sc_cat_item
- sys_attachment
- sys_attachment_doc
- sys_user_role

- Comprobaste que cada documento es único en ServiceNow las demás fuentes de datos que planeas usar para el mismo índice y entre ellas. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Guardó sus credenciales de ServiceNow autenticación en un AWS Secrets Manager secreto y, si usa la API, anotó el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda

volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de ServiceNow datos. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de ServiceNow datos, debe proporcionar los detalles necesarios de la fuente de ServiceNow datos para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configurado ServiceNow , Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a ServiceNow

1. Inicie sesión en la consola AWS de administración y abra la Amazon Kendra consola.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el ServiceNowconector V1.0 y, a continuación, elija Agregar fuente de datos.
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. ServiceNow host: introduzca la URL del ServiceNow host.

b. ServiceNow versión: seleccione su ServiceNow versión.

c. Elija entre Autenticación básica y Autenticación OAuth 2.0 según el caso de uso.

d. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de ServiceNow autenticación. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

i. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-ServiceNow -' se añade automáticamente a tu nombre secreto.

ii. Si utilizas la autenticación básica, introduce el nombre secreto, el nombre de usuario y la contraseña de tu cuenta. ServiceNow

Si utilizas la autenticación OAuth2, introduce el nombre secreto, el nombre de usuario, la contraseña, el ID de cliente y el secreto de cliente que creaste en tu cuenta. ServiceNow

iii. Haga clic en Guardar y agregar secreto.

e. IAM rol: elige un IAM rol existente o crea uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

f. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. Incluir artículos de conocimiento: elija si desea indexar los artículos de conocimiento.

b. Tipo de artículos de conocimiento: elija entre incluir solo artículos públicos e Incluir artículos según una consulta de ServiceNow filtro según su caso de uso. Si selecciona Incluir artículos según una consulta de ServiceNow filtro, debe introducir una consulta de filtro copiada de su ServiceNow cuenta.

c. Incluir archivos adjuntos de artículos de conocimiento: elija si desea indexar los archivos adjuntos de los artículos de conocimiento. También puede seleccionar tipos de archivos específicos para indexarlos.

d. Incluir elementos del catálogo: elija si desea indexar los elementos del catálogo.

e. Incluir archivos adjuntos de elementos del catálogo: elija si desea indexar los archivos adjuntos de los elementos del catálogo. También puede seleccionar tipos de archivos específicos para indexarlos.

f. Frecuencia: la frecuencia con la Amazon Kendra que se sincronizará con la fuente de datos.

g. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Artículos de conocimiento y catálogo de servicios: seleccione entre los campos de fuentes de datos predeterminados Amazon Kendra generados y otras asignaciones de campos sugeridas que desee asignar a su índice.

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra ServiceNow

Debe especificar lo siguiente mediante la ServiceNowConfiguration API:

- URL de la fuente de datos: especifique la ServiceNow URL. El punto de conexión del host debe tener el siguiente aspecto *your-domain.service-now.com*.
- Instancia de host de la fuente de datos: especifique la versión de la instancia de ServiceNow host como oLONDON. OTHERS
- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. ServiceNow

  Si está utilizando la autenticación básica, el secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  Si está utilizando la autenticación OAuth2, el secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "username": "user name",
      "password": "password",
      "clientId": "client id",
      "clientSecret": "client secret"
  }
  ```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector y. ServiceNow Amazon Kendra Para obtener más información, consulte las IAM funciones de las fuentes ServiceNow de datos.

También puede añadir las siguientes características opcionales:

- Asignaciones de campos: elija asignar los campos de la fuente de ServiceNow datos a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice `_document_body`. Todos los demás campos son opcionales.

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos adjuntos de catálogos y artículos de conocimiento.

> **Note**
>
> La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Parámetros de indexación: también puede optar por especificar si desea:
  - Indexar artículos de conocimiento y catálogos de servicios, o ambos. Si decide indexar artículos de conocimiento y elementos del catálogo de servicios, debe proporcionar el nombre del ServiceNow campo que está asignado al campo de contenido del documento de índice en el Amazon Kendra índice.
  - Indexar los archivos adjuntos de los artículos de conocimiento y los elementos del catálogo.
  - Utilice una ServiceNow consulta que seleccione documentos de una o más bases de conocimiento. Las bases de conocimiento pueden ser públicas o privadas. Para obtener más información, consulte Especificar documentos a indexar con una consulta.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de ServiceNow datos, consulte:

- Cómo empezar a usar Amazon Kendra ServiceNow Online Connector

ServiceNow conector V2.0

ServiceNow proporciona un sistema de administración de servicios basado en la nube para crear y administrar flujos de trabajo a nivel de organización, como los servicios de TI, los sistemas de emisión de tickets y el soporte. Puede utilizarlo Amazon Kendra para indexar sus ServiceNow catálogos, artículos de conocimiento, incidentes y sus archivos adjuntos.

Para solucionar problemas del conector de la fuente de Amazon Kendra ServiceNow datos, consulteSolución de problemas con los orígenes de datos.

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra ServiceNow el conector de fuente de datos admite las siguientes funciones:

- Asignaciones de campo
- control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronización de contenido completa e incremental
- ServiceNow versiones de instancia: Roma, Sandiego, Tokio, otras
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de ServiceNow datos, realiza estos cambios en tus AWS cuentas ServiceNow y.

En ServiceNow, asegúrate de tener:

- Ha creado una instancia de desarrollador personal o empresarial y dispone de una ServiceNow instancia con una función administrativa.

- Ha copiado el host de la URL de la ServiceNow instancia. El formato de la URL del host que introduzca es *your-domain.service-now.com*. Necesitas la URL de tu ServiceNow instancia para conectarte Amazon Kendra.
- Apuntó sus credenciales de autenticación básicas, es decir, un nombre de usuario y una contraseña Amazon Kendra para poder conectarse a la ServiceNow instancia.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Opcional: credenciales de cliente de OAuth 2.0 configuradas que pueden identificarse Amazon Kendra mediante un nombre de usuario, una contraseña y un identificador de cliente generado y un secreto de cliente. Consulte ServiceNow la documentación sobre la autenticación de OAuth 2.0 para obtener más información.
- Comprobaste que cada documento es único en ServiceNow las demás fuentes de datos que planeas usar para el mismo índice y entre ellas. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Guardó sus credenciales de ServiceNow autenticación en un AWS Secrets Manager secreto y, si usa la API, anotó el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de ServiceNow datos. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de ServiceNow datos, debe proporcionar los detalles necesarios de la fuente de ServiceNow datos para que Amazon Kendra pueda acceder a sus datos. Si aún no lo ha configurado ServiceNow , Amazon Kendra consulteRequisitos previos.

Console

Para conectarse Amazon Kendra a ServiceNow

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

   > **Note**
   >
   > Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el ServiceNow conector y, a continuación, elija Agregar conector. Si utiliza la versión 2 (si corresponde), elija el ServiceNow conector con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

a. ServiceNow host: introduzca la URL del ServiceNow host. El formato de la URL del host que introduzca es *your-domain.service-now.com*.

b. ServiceNow versión: seleccione la versión de la ServiceNow instancia. Puede seleccionar entre Roma, Sandiego, Tokio u otros.

c. Autorización: active o desactive la información de la lista de control de acceso (ACL) para sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

d. Autenticación: elija entre la autenticación básica y la autenticación Oauth 2.0.

e. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo para almacenar sus Secrets Manager credenciales de autenticación. ServiceNow Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager . En la ventana, introduzca la siguiente información:

i. Nombre del secreto: un nombre para su secreto. El prefijo 'AmazonKendra-ServiceNow -' se añade automáticamente a tu nombre secreto.

ii. Si utilizas la autenticación básica, introduce el nombre secreto, el nombre de usuario y la contraseña de tu cuenta. ServiceNow

Si utilizas la autenticación OAuth2.0, introduce el nombre secreto, el nombre de usuario, la contraseña, el ID de cliente y el secreto de cliente que creaste en tu cuenta. ServiceNow

iii. Guarda y añade tu secreto.

f. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

g. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

h. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

i. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. En Artículos de Knowledge, elija entre las siguientes opciones:

- Artículos de conocimiento: elija si desea indexar los artículos de conocimiento.
- Archivos adjuntos de artículos de conocimiento: elija si desea indexar los archivos adjuntos de los artículos de conocimiento.
- Tipo de artículos de conocimiento: elija entre Solo artículos públicos y artículos de conocimiento basados en una consulta de ServiceNow filtro según su caso de uso. Si selecciona Incluir artículos según una consulta de ServiceNow filtro, debe introducir una consulta de filtro copiada de su ServiceNow cuenta. Entre los ejemplos de consultas de filtro se incluye: `workflow_state=draft^EQ`,

kb_knowledge_base=dfc19531bf2021003f07e2c1ac0739ab^textISNOTEMPTY^EQ, *article_type=text^active=true^EQ*.

> **Important**
>
> Si elige rastrear solo los artículos públicos, Amazon Kendra rastreará solo los artículos de conocimiento a los que se haya asignado una función de acceso público en. ServiceNow

- Incluir artículos según el filtro de descripción breve: especifique patrones de expresión regular para incluir o excluir artículos específicos.

b. En Elementos del catálogo de servicios:

- Elementos del catálogo de servicios: elija si desea indexar los elementos del catálogo de servicios.
- Archivos adjuntos de elementos del catálogo de servicios: elija si desea indexar los archivos adjuntos de los elementos del catálogo de servicios.
- Elementos activos del catálogo de servicios: elija si desea indexar los elementos activos del catálogo de servicios.
- Elementos inactivos del catálogo de servicios: elija si desea indexar los elementos inactivos del catálogo de servicios.
- Consulta de filtrado: elija incluir los elementos del catálogo de servicios en función de un filtro definido en su instancia. ServiceNow Entre los ejemplos de consultas de filtro se incluye: *short_descriptionLIKEAccess^category=2809952237b1300054b6a3549dbe5dd4nameSTARTSWITHService^active=true^EQ*.
- Incluir los elementos del catálogo de servicios según un filtro de descripción breve: especifique un patrón de regex para incluir elementos del catálogo específicos.

c. En Incidentes:

- Incidentes: elija si desea indexar los incidentes de servicio.
- Archivos adjuntos de incidentes: elija si desea indexar los archivos adjuntos de incidentes.
- Incidentes activos: elija si desea indexar los incidentes activos.
- Incidentes inactivos: elija si desea indexar los incidentes inactivos.

- Tipo de incidente activo: elija entre Todos los incidentes, Incidentes abiertos, Incidentes abiertos: no asignados e Incidentes resueltos, según el caso de uso.
- Consulta de filtro: elija incluir los incidentes en función de un filtro definido en la instancia ServiceNow . Entre los ejemplos de consultas de filtro se incluye: *short_descriptionLIKETest^urgency=3^state=1^EQ*, *priority=2^category=software^EQ*.
- Incluir incidentes según el filtro de descripción breve: especifique un patrón de regex para incluir incidentes específicos.

d. En Configuración adicional:

- Información de ACL: las listas de control de acceso de las entidades que ha seleccionado se incluyen de forma predeterminada. Al anular la selección de una lista de control de acceso, todos los archivos de esa categoría serán públicos. Las opciones de ACL se desactivan automáticamente para las entidades no seleccionadas. En el caso de los artículos públicos, no se aplica la ACL.
- Para Tamaño máximo de archivo: especifique el límite de tamaño de archivo en MB que Amazon Kendra rastreará. Amazon Kendra rastreará solo los archivos dentro del límite de tamaño que usted defina. El tamaño predeterminado del archivo es de 50 MB. El tamaño máximo del archivo debe ser superior a 0 MB e inferior o igual a 50 MB.
- Patrones de regex de archivos adjuntos: añada patrones de expresiones regulares para incluir o excluir determinados archivos adjuntos de catálogos, artículos de conocimiento e incidentes. Puede agregar hasta 100 patrones.

e. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no eliges la sincronización completa como opción de modo de sincronización.

- Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

f. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

g. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Asignaciones de campos predeterminadas: seleccione entre las fuentes de datos predeterminadas Amazon Kendra generadas los campos que desee asignar a su índice.

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse a Amazon Kendra ServiceNow

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `SERVICENOWV2` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- URL del host: especifique la versión de la instancia del ServiceNow host. Por ejemplo, *`your-domain.service-now.com`*.
- Tipo de autenticación: especifique el tipo de autenticación que utiliza, ya sea `OAuth2` para su ServiceNow instancia `basicAuth` o para ella.
- ServiceNow versión de instancia: especifique la ServiceNow instancia que utiliza, `Tokyo` `SandiegoRome`, o. `Others`
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizarse el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no

eliges la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:

- `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
- `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

- Nombre secreto de recurso de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación que creó en su cuenta. ServiceNow

  Si utiliza la autenticación básica, el secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

- Si utiliza las credenciales de cliente de OAuth2, el secreto se almacena en una estructura JSON con las siguientes claves:

  ```
  {
      "username": "user name",
      "password": "password",
      "clientId": "client id",
      "clientSecret": "client secret"
  }
  ```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar a un IAM rol permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el ServiceNow conector y. Amazon Kendra Para obtener más información, consulte las IAM funciones de las fuentes ServiceNow de datos.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Filtros de inclusión y exclusión: puede especificar si desea incluir o excluir determinados archivos adjuntos mediante los nombres y tipos de archivos de los artículos de conocimiento, los catálogos de servicios y los incidentes.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Documentos específicos para indexar: puede utilizar una ServiceNow consulta para especificar los documentos que desee de una o más bases de conocimiento, incluidas las bases de conocimiento privadas. El acceso a las bases de conocimiento lo determina el usuario que utilice para conectarse a la ServiceNow instancia. Para obtener más información, consulte Especificar documentos a indexar con una consulta.
- Parámetros de indexación: también puede optar por especificar si desea:
  - Indexar artículos de conocimiento, catálogos de servicios e incidentes, o todos ellos. Si decide indexar artículos de conocimiento, artículos del catálogo de servicios e incidentes, debe proporcionar el nombre del ServiceNow campo que se asigna al campo de contenido del documento de Amazon Kendra índice del índice.
  - Indexar los archivos adjuntos de los artículos de conocimiento, los elementos del catálogo de servicios y los incidentes.
  - Incluir artículos de conocimiento, elementos del catálogo de servicios e incidentes según el patrón de filtrado `short description`.
  - Elegir filtrar los elementos e incidentes del catálogo de servicios activos e inactivos.
  - Elegir filtrar los incidentes en función de su tipo.
  - Elegir de qué entidades se debe rastrear la ACL.
  - Puede utilizar una ServiceNow consulta para especificar los documentos que desee de una o más bases de conocimiento, incluidas las bases de conocimiento privadas. El acceso a las

bases de conocimiento lo determina el usuario que utilice para conectarse a la ServiceNow instancia. Para obtener más información, consulte Especificar documentos a indexar con una consulta.

- Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.
- Asignaciones de campos: elija asignar los campos de la fuente de ServiceNow datos a los campos de índice. Amazon Kendra Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para ver una lista de otras claves JSON importantes que debes configurar, consulta el esquema ServiceNow de la plantilla.

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de ServiceNow datos, consulte:

- Para empezar, Amazon Kendra anunciamos el ServiceNow conector actualizado (V2) para Amazon Kendra

Especificar los documentos que se van a indexar con una consulta

Puede usar una ServiceNow consulta para especificar los documentos que desea incluir en un Amazon Kendra índice. Cuando utiliza una consulta, puede especificar varias bases de conocimiento, incluidas las bases de conocimiento privadas. El acceso a las bases de conocimiento lo determina el usuario que utilice para conectarse a la ServiceNow instancia.

Para crear una consulta, utilice el generador de ServiceNow consultas. Puede usar el creador para generar la consulta y comprobar que devuelve la lista correcta de documentos.

Para crear una consulta mediante la ServiceNow consola

1. Inicie sesión en la ServiceNow consola.
2. En el menú de la izquierda, seleccione Knowledge, luego Articles y, a continuación, All.
3. En la parte superior de la página, elija el icono del filtro.
4. Utilice el creador de consultas para crear la consulta.
5. Cuando la consulta esté completa, haga clic con el botón derecho en ella y seleccione Copy query para copiarla del creador de consultas. Guarde esta consulta para utilizarla en ella Amazon Kendra.

Recuerde no cambiar ningún parámetro de la consulta al copiarla. Si no se reconoce alguno de los parámetros de la consulta, ServiceNow trata el parámetro como vacío y no lo usa para filtrar los resultados.

Slack

Slack es una aplicación de comunicación empresarial que permite a los usuarios enviar mensajes y archivos adjuntos a través de varios canales públicos y privados. Puedes usarlo Amazon Kendra para indexar tus canales públicos y privados de Slack, guardar y almacenar mensajes, archivos y archivos adjuntos y mensajes directos y grupales. También puede elegir contenido específico para filtrar.

> **Note**
>
> Amazon Kendra ahora es compatible con un conector de Slack actualizado.
> La consola se ha actualizado automáticamente para ti. Todos los conectores nuevos que cree en la consola utilizarán la arquitectura actualizada. Si usa la API, ahora debe usar el TemplateConfigurationobjeto en lugar del `SlackConfiguration` objeto para configurar el conector.
> Los conectores configurados con la antigua arquitectura de consola y API seguirán funcionando tal y como estaban configurados. Sin embargo, no podrá editarlos ni actualizarlos. Si desea editar o actualizar la configuración del conector, debe crear un conector nuevo.
> Se recomienda migrar el flujo de trabajo del conector a la versión actualizada. Está previsto que el soporte para los conectores configurados con la arquitectura anterior finalice en junio de 2024.

Puedes conectarte Amazon Kendra a tu fuente de datos de Slack mediante la Amazon Kendra consola o la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuente de datos de Amazon Kendra Slack, consulta. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra El conector de fuente de datos de Slack admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- Sincronizaciones de contenido completas e incrementales

- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar tu fuente de datos de Slack, realiza estos cambios en tu cuenta y en tu cuenta de Slack. AWS

En Slack, asegúrese de que:

- Has configurado un token OAuth de usuario de Slack Bot o un token de OAuth de usuario de Slack. Puedes elegir cualquier token para Amazon Kendra conectarte a tu fuente de datos de Slack. Se necesita un token para usarlo como credenciales de autenticación. Consulte la documentación de Slack sobre los tokens de acceso para obtener más información.

> **Note**
>
> Si utiliza el token de bot como parte de sus credenciales de Slack, no podrá indexar los mensajes directos ni los mensajes de grupo y deberá añadir el token de bot al canal que desee indexar.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Anote el ID de equipo de su espacio de trabajo Slack en la URL de la página principal de su espacio de trabajo Slack. *Por ejemplo, https://app.slack.com/client/T0123456789/ …* donde *T0123456789* es el ID del equipo.
- Se agregaron los siguientes alcances/permisos de Oauth:

Ámbito del token de usuario	Ámbito del token del bot
• channels:history • channels:read	• channels:history • channels:manage

Ámbito del token de usuario	Ámbito del token del bot
• emoji:read • files:read • groups:history • groups:read • im:history • im:read • mpim:history • mpim:read • team:read • users.profile:read • users:read • users:read.email	• channels:read • conversations.connect:manage • conversations.connect: leer • files:read • groups:history • groups:read • im:history • im:read • mpim:history • mpim:read • reacciones: leer • team:read • usergroups:read • users.profile:read • users:read • users:read.email

- Ha comprobado que cada documento es único en Slack y en otros orígenes de datos que vaya a utilizar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En tu Cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Slack en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tienes un IAM rol o un secreto existentes, puedes usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar tu fuente de datos de Slack. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarte Amazon Kendra a tu fuente de datos de Slack, debes proporcionar los detalles necesarios de tu fuente de datos de Slack para que Amazon Kendra puedas acceder a tus datos. Si aún no has configurado Slack para Amazon Kendra, consulta. Requisitos previos

Console

Para conectarte Amazon Kendra a Slack

1. Inicia sesión en la Amazon Kendra consola AWS Management Console y ábrela.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Añadir fuente de datos, selecciona el conector de Slack y, a continuación, selecciona Añadir conector. Si utilizas la versión 2 (si corresponde), elige el conector de Slack con la etiqueta «V2.0".

5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:

   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.

   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.

   c. En el idioma predeterminado: elige un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.

   e. Elija Siguiente.

6. En la página Definir acceso y seguridad, introduzca la siguiente información:

   a. Para el ID de equipo del espacio de trabajo de Slack: el ID del equipo de tu espacio de trabajo de Slack. Puedes encontrar el ID de tu equipo en la URL de la página principal de tu espacio de trabajo de Slack. *Por ejemplo, https://app.slack.com/client/T0123456789/…* donde *T0123456789* es el ID del equipo.

   b. Autorización: activa o desactiva la información de la lista de control de acceso (ACL) en tus documentos si tienes una ACL y quieres usarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

   c. AWS Secrets Manager secreto: elige un secreto existente o crea uno nuevo Secrets Manager para almacenar tus credenciales de autenticación de Slack. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

      i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

         A. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Slack» se añade automáticamente a tu nombre secreto.

         B. Para el token de Slack: introduce los valores de las credenciales de autenticación que configuraste en Slack.

      ii. Guarda y añade tu secreto.

d. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

e. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

f. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

g. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

a. Selecciona el tipo de contenido: selecciona las entidades o los tipos de contenido de Slack que quieres rastrear. Puedes elegir entre todos los canales: canales públicos, canales privados, mensajes grupales y mensajes privados.

b. Selecciona la fecha de inicio del rastreo: introduce la fecha en la que quieres empezar a rastrear tu contenido.

c. Para una configuración adicional: elige incluir mensajes archivados y de bots y utiliza patrones de expresiones regulares para incluir o excluir cierto contenido.

> **Note**
>
> Si decides incluirlos tanto para los ID como para los nombres de los canales, el conector de Amazon Kendra Slack priorizará los ID de los canales por encima de los nombres de los canales.
> Si has decidido incluir algunos mensajes privados y grupales, el conector de Amazon Kendra Slack ignorará todos los mensajes privados y grupales y solo rastreará los mensajes privados y grupales que especifiques.

d. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Cuando sincronizas tu fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no eliges la sincronización completa como opción de modo de sincronización.

   - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
   - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

e. En el programa de ejecución sincronizado, para Frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

f. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

   b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

   c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la

información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarte Amazon Kendra a Slack

Debe especificar un JSON del esquema del origen de datos mediante la API TemplateConfiguration. Debe proporcionar la siguiente información:

- Fuente de datos: especifica el tipo de fuente de datos como `SLACK` cuando usas el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- ID del equipo del espacio de trabajo Slack: El ID del equipo de Slack que copió de la URL de su página principal de Slack.
- Fecha inicial: la fecha en la que empezarás a rastrear los datos de tu equipo de espacio de trabajo de Slack. La fecha debe seguir este formato:. yyyy-mm-dd
- Modo de sincronización: especifique cómo Amazon Kendra debe actualizar su índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización. Puede elegir entre las siguientes opciones:
  - `FORCED_FULL_CRAWL`para indexar todo el contenido de forma actualizada, sustituyendo el contenido existente cada vez que la fuente de datos se sincronice con el índice.
  - `FULL_CRAWL`para indexar solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
  - `CHANGE_LOG`para indexar solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
- Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del

usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

- Nombre secreto del recurso de Amazon (ARN): proporciona el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de tu cuenta de Slack. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
    "slackToken": "token"
}
```

- IAM rol: especifica `RoleArn` cuándo llamas `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a tu Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Slack y. Amazon Kendra Para obtener más información, consulte Roles de Slack IAM para orígenes de datos de Slack.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Canales específicos: filtra por canales públicos o privados y especifica determinados canales por su ID.
- Tipos de canales y mensajes: si Amazon Kendra debes indexar tus canales públicos y privados, tus mensajes grupales y directos, y tus mensajes bots y archivados. Si utiliza un token de bot como parte de sus credenciales de autenticación de Slack, deberá añadir el token de bot al canal que desee indexar. No puede indexar mensajes directos y mensajes de grupo utilizando un token de bot.
- Mira hacia atrás: puedes configurar un `lookBack` parámetro para que el conector de Slack rastree el contenido actualizado o eliminado hasta un número específico de horas antes de la última sincronización del conector.

- Filtros de inclusión y exclusión: especifica si deseas incluir o excluir determinado contenido de Slack. Si utiliza un token de bot como parte de sus credenciales de autenticación de Slack, deberá añadir el token de bot al canal que desee indexar. No puede indexar mensajes directos y mensajes de grupo utilizando un token de bot.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Asignaciones de campos: elija asignar los campos del origen de datos de Slack a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

  > **Note**
  >
  > El campo del cuerpo del documento o el cuerpo del documento equivalente de tus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para obtener una lista de otras claves JSON importantes que debe configurar, consulte el Esquema de plantilla de Slack.

Más información

Para obtener más información sobre la integración Amazon Kendra con tu fuente de datos de Slack, consulta:

- Desentrañe los conocimientos de los espacios de trabajo de Slack con la búsqueda inteligente mediante el conector Slack Amazon Kendra

Zendesk

Zendesk es un sistema de gestión de relaciones con los clientes que ayuda a las empresas a automatizar y mejorar las interacciones de atención al cliente. Puede usarlo Amazon Kendra para indexar los tickets de soporte de Zendesk, los comentarios de los tickets, los archivos adjuntos de los tickets, los artículos del centro de ayuda, los comentarios de los artículos, los archivos adjuntos a los comentarios de los artículos, los temas de la comunidad, las publicaciones de la comunidad y los comentarios de las publicaciones de la comunidad.

Puede filtrar por nombre de organización si quiere indexar los tickets que solo están dentro de una organización específica. También puede elegir establecer una fecha de rastreo para comenzar a rastrear los datos de Zendesk.

Puede conectarse Amazon Kendra a su fuente de datos de Zendesk mediante la Amazon Kendra consola y la TemplateConfigurationAPI.

Para solucionar problemas del conector de fuentes de datos de Amazon Kendra Zendesk, consulte. Solución de problemas con los orígenes de datos

Temas

- Características admitidas
- Requisitos previos
- Instrucciones de conexión
- Más información

Características admitidas

Amazon Kendra El conector de fuente de datos de Zendesk admite las siguientes funciones:

- Asignaciones de campo
- Control de acceso de usuarios
- Filtros de inclusión/exclusión
- El registro de cambios y las sincronizaciones de contenido completas e incrementales
- Nube privada virtual (VPC)

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de datos de Zendesk, realice estos cambios en su cuenta de Zendesk y en sus cuentas. AWS

En Zendesk, asegúrese de que:

- Creó una cuenta administrativa de Zendesk Suite (Professional/Enterprise).
- Apuntó la URL de su servidor de Zendesk. *Por ejemplo, https://{sub-domain}.zendesk.com/.*

  > **Note**
  >
  > (local o en el servidor) Amazon Kendra comprueba si la información de punto final incluida AWS Secrets Manager es la misma que la información de punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a evitar el problema del suplente confuso, que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, pero utiliza Amazon Kendra como proxy para acceder al secreto configurado y realizar la acción. Si más adelante cambia la información de punto de conexión, debe crear un nuevo secreto para sincronizar esta información.

- Se configuró un token de OAuth 2.0 que contiene un ID de cliente, un secreto de cliente, un nombre de usuario y una contraseña. Es necesario utilizar el token de OAuth 2.0 como credenciales de autenticación. Consulte la documentación de Zendesk sobre la configuración de los tokens de OAuth 2.0 para obtener más información.

  > **Note**
  >
  > Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

- Se agregó el siguiente ámbito de OAuth 2.0:
  - leer
- Opcional: se instaló un certificado SSL para permitir a Amazon Kendra la conexión.
- Marcó que cada documento es único en Zendesk y en otras Origen de datos que planea usar para el mismo índice. Cada origen de datos que desee utilizar para un índice no debe contener el

mismo documento en varios orígenes de datos. Los ID de documento son globales para un índice y deben ser únicos por índice.

En el suyo Cuenta de AWS, asegúrese de tener:

- Creó un Amazon Kendra índice y, si usa la API, anotó el ID del índice.
- Creó un IAM rol para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

> **Note**
>
> Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al ID AWS Secrets Manager secreto correcto.

- Ha guardado sus credenciales de autenticación de Zendesk en un secreto de AWS Secrets Manager y, si utiliza la API, ha anotado el ARN del secreto.

> **Note**
>
> Le recomendamos que actualice o modifique con regularidad las credenciales y el secreto. Por su propia seguridad, proporcione solo el nivel de acceso necesario. No se recomienda volver a utilizar las credenciales y los datos secretos en varios orígenes de datos ni en las versiones 1.0 y 2.0 del conector (si procede).

Si no tiene un IAM rol o secreto existente, puede usar la consola para crear un nuevo IAM rol y un Secrets Manager secreto al conectar su fuente de datos de Zendesk. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Zendesk, debe proporcionar los detalles necesarios de su fuente de datos de Zendesk para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Zendesk para Amazon Kendra, consulte. Requisitos previos

Console

Para conectarse Amazon Kendra a Zendesk

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala.
2. En el panel de navegación izquierdo, elija Índices y, a continuación, elija el índice que desee usar de la lista de índices.

> **Note**
>
> Puede elegir configurar o editar los ajustes de Control de acceso de usuarios en la Configuración del índice.

3. En la página Introducción, seleccione Agregar origen de datos.
4. En la página Agregar fuente de datos, elija el conector de Zendesk y, a continuación, elija Agregar conector. Si usa la versión 2 (si corresponde), elija el conector de Zendesk con la etiqueta «V2.0".
5. En la página Especificar detalles del origen de datos, introduzca la siguiente información:
   a. En Nombre y descripción, en Nombre del origen de datos: introduzca un nombre para el origen de datos. Puede incluir guiones, pero no espacios.
   b. (Opcional) Descripción: introduzca una descripción opcional para el origen de datos.
   c. En el idioma predeterminado: elija un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
   d. En Etiquetas, para añadir una nueva etiqueta: incluya etiquetas opcionales para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costes.
   e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
   a. URL de Zendesk: introduzca la URL de su cuenta de Zendesk. Por ejemplo, *https://{sub-domain}.zendesk.com/*.
   b. Autorización: active o desactive la información de la lista de control de acceso (ACL) de sus documentos, si tiene una ACL y desea utilizarla para el control de acceso. La ACL especifica a qué documentos pueden acceder los usuarios y los grupos. La información de la ACL se utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

c. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de autenticación de Zendesk. Si decide crear un secreto nuevo, se abrirá una ventana de secreto de AWS Secrets Manager .

   i. Introduzca la siguiente información en la ventana Crear un secreto de AWS Secrets Manager :

      A. Nombre del secreto: un nombre para su secreto. El prefijo «AmazonKendra-Zendesk» se agrega automáticamente a su nombre secreto.

      B. Para el ID de cliente, el secreto del cliente, el nombre de usuario y la contraseña: introduzca los valores de las credenciales de autenticación configurados en Zendesk.

   ii. Guarde y añada su secreto.

d. Nube privada virtual (VPC): puede optar por utilizar una VPC. Si es así, debe agregar Subredes y Grupos de seguridad de VPC.

e. Rastreador de identidad: especifique si se debe activar el rastreador Amazon Kendra de identidad. El rastreador de identidades utiliza la información de la lista de control de acceso (ACL) de los documentos para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Si tiene una ACL para sus documentos y decide utilizarla, también puede optar por activar el rastreador de identidades para configurar el filtrado Amazon Kendra de los resultados de búsqueda según el contexto del usuario. De lo contrario, si el rastreador de identidades está desactivado, se pueden realizar búsquedas públicas en todos los documentos. Si quieres usar el control de acceso para tus documentos y el rastreador de identidad está desactivado, también puedes usar la PutPrincipalMappingAPI para cargar información de acceso de usuarios y grupos para filtrar el contexto de los usuarios.

f. IAM rol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio y al contenido del índice.

> **Note**
>
> IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se utiliza para un índice o para las preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

g. Elija Siguiente.

7. En la página Configurar ajustes de sincronización, introduzca la siguiente información:

   a. Seleccione el contenido: seleccione los tipos de contenido que desea rastrear desde los tickets, los artículos del centro de ayuda, los temas de la comunidad y mucho más.

   b. Nombre de la organización: introduzca los nombres de las organizaciones de Zendesk para filtrar el contenido.

   c. Fecha de inicio de la sincronización: introduzca la fecha a partir de la cual quiere empezar a rastrear el contenido.

   d. Patrones regex: añada patrones de expresiones regulares para incluir o excluir determinados archivos. Puede agregar hasta 100 patrones.

   e. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido del origen de datos. Al sincronizar la fuente de datos con Amazon Kendra por primera vez, todo el contenido se rastrea e indexa de forma predeterminada. Debes realizar una sincronización completa de los datos si la sincronización inicial ha fallado, incluso si no seleccionas la sincronización completa como opción de modo de sincronización.

      - Sincronización completa: indexa todo el contenido de forma inmediata y reemplaza el contenido existente cada vez que la fuente de datos se sincronice con el índice.
      - Sincronización nueva y modificada: indexe solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.
      - Sincronización nueva, modificada o eliminada: indexe solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. Amazon Kendra puede usar el mecanismo de la fuente de datos para realizar un seguimiento de los cambios en el contenido e indexar el contenido que ha cambiado desde la última sincronización.

   f. Calendario de ejecución sincronizado para la frecuencia: elija la frecuencia con la que desea sincronizar el contenido de la fuente de datos y actualizar el índice.

   g. Elija Siguiente.

8. En la página Establecer asignaciones de campos, especifique la siguiente información:

   a. Campos de fuente de datos predeterminados: seleccione entre los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.

b. Agregar campo: para agregar campos de origen de datos personalizados para crear un nombre de campo de índice al que asignarlos y el tipo de datos del campo.

c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Añadir origen de datos. También puede elegir editar la información desde esta página. El origen de datos aparecerá en la página Orígenes de datos una vez que el origen de datos se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Zendesk

Debe especificar un JSON del esquema de la fuente de datos mediante la TemplateConfigurationAPI. Debe proporcionar la siguiente información:

- Fuente de datos: especifique el tipo de fuente de datos como `ZENDESK` cuando utiliza el esquema TemplateConfigurationJSON. Especifique también la fuente de datos `TEMPLATE` al llamar a la CreateDataSourceAPI.
- URL del host: proporcione la URL de su host de Zendesk como parte de la configuración de la conexión o de los detalles del punto de conexión del repositorio. Por ejemplo, *`https://yoursubdomain.zendesk.com`*.
- Registro de cambios: si se Amazon Kendra debe usar el mecanismo de registro de cambios de la fuente de datos de Zendesk para determinar si un documento debe actualizarse en el índice.

> **Note**
>
> Utilice el registro de cambios si no quiere que Amazon Kendra digitalice todos los documentos. Si el registro de cambios es grande, es posible que se tarde Amazon Kendra menos en escanear los documentos de la fuente de datos de Zendesk que en procesar el registro de cambios. Si está sincronizando el origen de datos de Zendesk con su índice por primera vez, se escanean todos los documentos.

- Nombre secreto del recurso de Amazon (ARN): proporcione el nombre del recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta de Zendesk. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{
```

```
    "hostUrl": "https://yoursubdomain.zendesk.com",
    "clientId": "client ID",
    "clientSecret": "Zendesk client secret",
    "userName": "Zendesk user name",
    "password": "Zendesk password"
}
```

- IAM rol: especifique `RoleArn` cuándo llama `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Zendesk y. Amazon Kendra Para obtener más información, consulte las IAM funciones de los orígenes de datos de Zendesk.

También puede añadir las siguientes características opcionales:

- Nube privada virtual (VPC): especifique a `VpcConfiguration` cuándo llamar a `CreateDataSource`. Para obtener más información, consulte Configuración Amazon Kendra para usar un Amazon VPC.
- Tipos de documentos o contenido: especifique si desea rastrear:
  - Tickets de soporte, comentarios de tickets y/o archivos adjuntos de comentarios de tickets
  - Artículos del centro de ayuda, anexos y comentarios de artículos
  - Guía los temas, las publicaciones o los comentarios de la comunidad
- Filtros de inclusión y exclusión: especifican si se debe incluir o excluir determinado contenido de Slack. Si utiliza un token de bot como parte de sus credenciales de autenticación de Slack, deberá añadir el token de bot al canal que desee indexar. No puede indexar mensajes directos y mensajes de grupo utilizando un token de bot.

  > **Note**
  >
  > La mayoría de los orígenes de datos utilizan patrones de expresiones regulares, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexará el contenido que coincida con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Filtrado por contexto de usuario y control de acceso:Amazon Kendra rastrea la lista de control de acceso (ACL) de tus documentos, si tienes una ACL para ellos. La información de la ACL se

utiliza para filtrar los resultados de búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte Filtrado de contexto de usuario.

- Asignaciones de campos: elija asignar los campos del origen de datos de Zendesk a los campos de índice de Amazon Kendra . Para obtener más información, consulte Asignación de campos de origen de datos.

> **Note**
>
> El campo del cuerpo del documento o el cuerpo del documento equivalente de sus documentos es obligatorio Amazon Kendra para poder buscarlos. Debe asignar el nombre del campo del cuerpo del documento en la fuente de datos al nombre del campo de índice`_document_body`. Todos los demás campos son opcionales.

Para ver una lista de otras claves JSON importantes que debe configurar, consulte el esquema de plantillas de Zendesk.

Más información

Para obtener más información sobre la integración Amazon Kendra con la fuente de datos de Zendesk, consulte:

- Descubra información de Zendesk con Amazon Kendra la búsqueda inteligente

Asignación de campos de origen de datos

Amazon Kendra Los conectores de fuentes de datos pueden asignar campos de documentos o contenido de la fuente de datos a los campos Amazon Kendra del índice. De forma predeterminada, cada conector está diseñado para rastrear campos de origen de datos específicos. Los campos de origen de datos predeterminados y sus propiedades no se pueden cambiar ni personalizar. En la Amazon Kendra consola, los campos predeterminados y las propiedades de los campos predeterminados que no se pueden editar aparecen atenuados.

Amazon Kendra Los conectores también le permiten asignar campos de contenido o documentos personalizados de la fuente de datos a los campos personalizados del índice. Por ejemplo, si tiene un campo en su origen de datos llamado “dept” que contiene información de departamento de un

documento, puede asignarlo a un campo de índice denominado “Department”. De esta forma, puede utilizar el campo al consultar documentos.

También puede mapear campos Amazon Kendra reservados o comunes, como`_created_at`. Si la fuente de datos tiene un campo denominado «fecha de creación», puede asignarlo al campo Amazon Kendra reservado equivalente denominado. `_created_at` Para obtener más información sobre los campos Amazon Kendra reservados, consulte Atributos o campos del documento.

Puede asignar campos para la mayoría de orígenes de datos. Puede crear asignaciones de campos para los siguientes orígenes de datos:

- Adobe Experience Manager
- Alfresco
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (Windows)
- Amazon FSx (NetApp ONTAP)
- Amazon RDS/Aurora
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon Kendra Rastreador web
- Amazon WorkDocs
- Box (Cuadro)
- Confluence
- Dropbox
- Drupal
- GitHub
- Unidades de Workspace de Google
- Gmail
- IBM DB2

- Jira
- Microsoft Exchange
- Microsoft OneDrive
- Microsoft SharePoint
- Microsoft Teams
- Microsoft SQL Server
- Microsoft Yammer
- MySQL
- Oracle Database
- PostgreSQL
- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

Si almacena los documentos en un bucket de S3 o en un origen de datos de S3, especifica los campos mediante un archivo de metadatos JSON. Para obtener más información, consulte Conector de origen de datos de S3.

La asignación de los campos de origen de datos a un campo de índice es un proceso de tres pasos:

1. Cree un índice. Para obtener más información, consulte Creación de un índice.
2. Actualice el índice para añadir campos.
3. Cree una fuente de datos e incluya asignaciones de campos para asignar los campos reservados y cualquier campo personalizado a los campos de Amazon Kendra indexación.

Para actualizar el índice y añadir campos personalizados, utilice la consola para editar las asignaciones de campos de la fuente de datos y añadir un campo personalizado o utilice la API. UpdateIndex Puede añadir un total de 500 campos personalizados a su índice.

Para los orígenes de datos de la base de datos, si el nombre de la columna de la base de datos coincide con el nombre de un campo reservado, el campo y la columna se asignan automáticamente.

Con la UpdateIndexAPI, puede añadir campos reservados y personalizados mediante. `DocumentMetadataConfigurationUpdates`

En el siguiente ejemplo de JSON se utiliza `DocumentMetadataConfigurationUpdates` para agregar al índice un campo denominado “Department”.

```
"DocumentmetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE"
   }
]
```

Al crear el campo, tiene la opción de configurar cómo se utiliza el campo en las búsquedas. Puede elegir entre las siguientes opciones:

- Visualizable: determina si el campo se devuelve en la respuesta de la consulta. El valor predeterminado es `true`.
- Facetable: indica que el campo se puede utilizar para crear facetas. El valor predeterminado es `false`.
- Buscable: determina si el campo se utiliza en la búsqueda. El valor predeterminado es `true` para los campos de cadena y `false` para los campos de número y fecha.
- Ordenable: indica que el campo se puede utilizar para ordenar los resultados de búsqueda. Solo se puede configurar para campos de fecha, número y cadena. No se puede configurar para los campos de lista de cadenas.

En el siguiente ejemplo de JSON se utiliza `DocumentMetadataConfigurationUpdates` para agregar al índice un campo denominado “Department” y marcarlo como facetable.

```
"DocumentMetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE",
       "Search": {
           "Facetable": true
       }
   }
]
```

Uso de campos de documentos comunes o Amazon Kendra reservados

Con la UpdateIndex API, puede crear campos reservados o comunes utilizando `DocumentMetadataConfigurationUpdates` y especificando el nombre del campo de índice Amazon Kendra reservado para asignarlos al atributo o nombre de campo del documento equivalente. También puede crear campos personalizados. Si utiliza un conector de fuente de datos, la mayoría incluye asignaciones de campos que asignan los campos del documento de la fuente de datos a campos de indexación. Amazon Kendra Si utiliza la consola, los campos se actualizan seleccionando el origen de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de asignación de campos para configurar el origen de datos.

Puede configurar el objeto `Search` para establecer un campo como visualizable, facetable, buscable y ordenable. Puede configurar el objeto `Relevance` para establecer el orden de clasificación, duración de potenciación o período de tiempo de un campo para aplicarlos a los valores de potenciación, actualización, valor de importancia y valores de importancia asignados a valores de campo específicos. Si utiliza la consola, puede configurar los ajustes de búsqueda de un campo seleccionando la opción de faceta en el menú de navegación. Para configurar el ajuste de relevancia, seleccione la opción de buscar en su índice en el menú de navegación, introduzca una consulta y utilice las opciones del panel lateral para ajustar la relevancia de la búsqueda. No puede cambiar el tipo de campo una vez que este se ha creado.

Amazon Kendra tiene los siguientes campos de documento reservados o comunes que puede usar:

- `_authors`: una lista de uno o más autores responsables del contenido del documento.
- `_category`: una categoría que coloca un documento en un grupo específico.
- `_created_at`: la fecha y hora en formato ISO 8601 de creación del documento. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en el horario de Europa Central.
- `_data_source_id`: el identificador del origen de datos que contiene el documento.
- `_document_body`: el contenido del documento.
- `_document_id`: un identificador único del documento.
- `_document_title`: el título del documento.
- `_excerpt_page_number`: el número de página de un archivo PDF en el que aparece el extracto del documento. Si el índice se creó antes del 8 de septiembre de 2020, debe volver a indexar los documentos antes de poder utilizar este atributo.

- `_faq_id`: si se trata de un documento tipo pregunta-respuesta (preguntas frecuentes), un identificador único para las preguntas frecuentes.
- `_file_type`: el tipo de archivo del documento, como pdf o doc.
- `_last_updated_at`: la fecha y hora en formato ISO 8601 de última actualización del documento. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en el horario de Europa Central.
- `_source_uri`: el URI en el que está disponible el documento. Por ejemplo, el URI del documento en el sitio web de una empresa.
- `_version`: un identificador de la versión específica de un documento.
- `_view_count`: el número de veces que se ha visto el documento.
- `_language_code` (cadena): el código de un idioma que se aplica al documento. Este valor se define por defecto en inglés si no especifica un idioma. Para obtener más información acerca de los idiomas admitidos, incluidos sus códigos, consulte Adición de documentos en idiomas distintos del inglés.

En el caso de campos personalizados, estos campos se crean mediante `DocumentMetadataConfigurationUpdates` con la API `UpdateIndex`, del mismo modo que cuando se crea un campo reservado o común. Debe establecer el tipo de datos adecuado para el campo personalizado. Si utiliza la consola, los campos se actualizan seleccionando el origen de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de asignación de campos para configurar el origen de datos. Algunos orígenes de datos no admiten la adición de campos nuevos o campos personalizados. No puede cambiar el tipo de campo una vez que este se ha creado.

Los siguientes son los tipos que puede configurar para los campos personalizados:

- Date
- Número
- Cadena
- Lista de cadenas

Si ha añadido documentos al índice mediante la BatchPutDocumentAPI, `Attributes` muestra los campos/atributos de los documentos y crea campos con el `DocumentAttribute` objeto.

En el caso de los documentos indexados a partir de una fuente de Amazon S3 datos, los campos se crean mediante un archivo de metadatos JSON que incluye la información de los campos.

Si utiliza una base de datos compatible como origen de datos, puede configurar los campos mediante la opción de asignación de campos.

Adición de documentos en idiomas distintos del inglés

Puede indexar documentos en varios idiomas. Si no especifica un idioma, Amazon Kendra indexa los documentos en inglés de forma predeterminada. El código de idioma de un documento se incluye en los metadatos del documento como un campo. Consulte Asignaciones de campos y Atributos personalizados para obtener más información sobre el campo `_language_code` de un documento.

Cuando llame CreateDataSource, puede especificar el código de idioma de todos los documentos de la fuente de datos. Si un documento no contiene un código de idioma especificado en un campo de metadatos, el documento se indexa utilizando el código de idioma especificado para todos los documentos en el nivel de origen de datos. En la consola, solo puede indexar documentos en un idioma compatible en el nivel del origen de datos. Vaya a Orígenes de datos, luego a la página Especificar detalles del origen de datos y elija un idioma en el menú desplegable Idioma.

También puede buscar o consultar documentos en un idioma compatible. Para obtener más información, consulte Buscar en idiomas.

Se admiten los siguientes idiomas y sus códigos (el inglés o `en` se admite de forma predeterminada si no especifica un idioma). En esta tabla se incluyen los idiomas Amazon Kendra compatibles con la búsqueda semántica completa, así como los idiomas que solo admiten la coincidencia simple de palabras clave. Los idiomas que admiten la búsqueda semántica completa se marcan con un asterisco y aparecen en negrita en la tabla siguiente. La búsqueda semántica completa también admite el inglés (idioma predeterminado).

Nombre del idioma	Código de idioma
Árabe	`ar`
Armenio	`hy`
Euskera	`eu`
Bengalí	`bn`

Nombre del idioma	Código de idioma
Búlgaro	bg
Catalán	ca
Chino: simplificado y tradicional*	zh
Checo	cs
Danés	da
Neerlandés	nl
Finés	fi
Francés: incluye francés (Canadá)*	fr
Gallego	gl
Alemán*	de
Griego	el
Hindi	hi
Húngaro	hu
Indonesio	id
Irlandés	ga
Italiano	it
Japonés*	ja
Coreano*	ko
Letón	lv
Lituano	lt

Nombre del idioma	Código de idioma
Noruego	`no`
Persa	`fa`
Portugués	`pt`
Portugués (Brasil)*	`pt-BR`
Rumano	`ro`
Ruso	`ru`
Sorani	`ckb`
Español: incluye español (México)*	`es`
Sueco	`sv`
Turco	`tr`

*Se admite la búsqueda semántica en este idioma.

Para los idiomas que admiten la búsqueda semántica, se admiten las siguientes funciones.

- La relevancia del documento va más allá de la simple coincidencia de palabras clave.
- Preguntas frecuentes más allá de la simple coincidencia de palabras clave.
- Extraer las respuestas de los documentos en Amazon Kendra función de su comprensión lectora.
- Buckets de confianza (muy alta, alta, media y baja) de los resultados de búsqueda.

En el caso de los idiomas que no admiten la búsqueda semántica, se admite la búsqueda simple de palabras clave para determinar la relevancia del documento y las preguntas frecuentes.

Los sinónimos (incluidos los sinónimos personalizados), el aprendizaje gradual y los comentarios y las sugerencias de consultas solo se admiten en inglés (idioma predeterminado).

Configuración Amazon Kendra para usar un Amazon VPC

Amazon Kendra puede conectarse a una nube privada virtual (VPC) que haya creado Amazon Virtual Private Cloud para indexar el contenido almacenado en las fuentes de datos que se ejecutan en su nube privada. Al crear un conector de origen de datos, puede proporcionar identificadores de subred y grupo de seguridad para la subred que contiene el origen de datos. Con esta información, Amazon Kendra crea una interfaz de red elástica que utiliza para comunicarse de forma segura con la fuente de datos de la VPC.

Para configurar un conector de fuente de Amazon Kendra datos Amazon VPC, puede utilizar la operación AWS Management Console o la CreateDataSourceAPI. Si usa la consola, conecta una VPC durante el proceso de configuración del conector.

> **Note**
>
> La Amazon VPC función es opcional al configurar un conector de fuente de Amazon Kendra datos. Si se puede acceder a la fuente de datos desde la Internet pública, no es necesario que habilite la Amazon VPC función. No todos los conectores Amazon Kendra de fuentes de datos son compatibles Amazon VPC.

Si la fuente de datos no se está ejecutando Amazon VPC y no se puede acceder a ella desde la Internet pública, primero debe conectar la fuente de datos a la VPC mediante una red privada virtual (VPN). A continuación, puede conectar su fuente de datos Amazon Kendra mediante una combinación de Amazon VPC y AWS Virtual Private Network. Para obtener información sobre la configuración de una VPN, consulte la Documentación de AWS VPN.

Temas

- Configuración del Amazon VPC soporte para Amazon Kendra conectores
- Configure una fuente Amazon Kendra de datos a la que conectarse Amazon VPC
- Conexión a una base de datos en una VPC
- Solución de problemas de conexión de VPC

Configuración del Amazon VPC soporte para Amazon Kendra conectores

Para Amazon VPC configurarlo para su uso con Amazon Kendra los conectores, siga los siguientes pasos.

Pasos

- Paso 1. Cree Amazon VPC subredes para Amazon Kendra
- Paso 2. Cree grupos de Amazon VPC seguridad para Amazon Kendra
- Paso 3. Configure su fuente de datos externa y Amazon VPC

Paso 1. Cree Amazon VPC subredes para Amazon Kendra

Cree o elija una Amazon VPC subred existente que Amazon Kendra pueda usar para acceder a su fuente de datos. Las subredes preparadas deben estar en una de las siguientes zonas de disponibilidad Regiones de AWS y en una de las siguientes zonas:

- Oeste de EE. UU. (Oregón)/us-west-2—usw2-az1, usw2-az2, usw2-az3
- Este de EE. UU. (Norte de Virginia)/us-east-1—use1-az1, use1-az2, use1-az4
- Este de EE. UU. (Ohio)/us-east-2—use2-az1, use2-az2, use2-az3
- Asia-Pacífico (Tokio)/ap-northeast-1—apne1-az1, apne1-az2, apne1-az4
- Asia-Pacífico (Bombay)/ap-south-1—aps1-az1, aps1-az2, aps1-az3
- Asia-Pacífico (Singapur)/ap-southeast-1—apse1-az1, apse1-az2, apse1-az3
- Asia-Pacífico (Sídney)/ap-southeast-2—apse2-az1, apse2-az2, apse2-az3
- Canadá (centro)/ca-central-1—cac1-az1, cac1-az2, cac1-az4
- Europa (Irlanda)/eu-west-1—euw1-az1, uew1-az2, euw1-az3
- Europa (Londres)/eu-west-2—usw2-az1, usw2-az2, usw2-az3

Debe poder acceder a su fuente de datos desde las subredes que proporcionó al Amazon Kendra conector.

Para obtener más información sobre cómo configurar las Amazon VPC subredes, consulte Subnets for your Amazon VPC en la Guía del usuario de Amazon VPC.

Si Amazon Kendra debe enrutar la conexión entre dos o más subredes, puede preparar varias subredes. Por ejemplo, la subred que contiene el origen de datos no tiene direcciones IP. En ese caso, puede proporcionar Amazon Kendra una subred adicional que tenga suficientes direcciones IP y esté conectada a la primera subred. Si enumera varias subredes, las subredes deben poder comunicarse entre sí.

Paso 2. Cree grupos de Amazon VPC seguridad para Amazon Kendra

Para conectar el conector de la fuente de Amazon Kendra datos Amazon VPC, debe preparar uno o más grupos de seguridad de la VPC para asignarlos. Amazon Kendra Los grupos de seguridad se asociarán a la interfaz de red elástica creada por Amazon Kendra. Esta interfaz de red controla el tráfico entrante y saliente Amazon Kendra al acceder a las Amazon VPC subredes.

Asegúrese de que las reglas de salida de su grupo de seguridad permitan que el tráfico de los conectores de las fuentes de Amazon Kendra datos acceda a las subredes y a la fuente de datos con las que se va a sincronizar. Por ejemplo, puede usar un conector de MySQL para sincronizar desde una base de datos de MySQL. Si utiliza el puerto predeterminado, los grupos de seguridad deben permitir el acceso Amazon Kendra al puerto 3306 del host que ejecuta la base de datos.

Se recomienda configurar un grupo de seguridad predeterminado con los siguientes valores Amazon Kendra para su uso:

- Reglas de entrada: si decide dejar este campo vacío, se bloqueará todo el tráfico entrante.
- Reglas de salida: agregue una regla para permitir que todo el tráfico saliente Amazon Kendra pueda iniciar las solicitudes de sincronización desde su fuente de datos.
  - Versión de IP: IPv4
  - Tipo: todo el tráfico
  - Protocolo: todo el tráfico
  - Rango de puertos: todos
  - Destino: 0.0.0.0/0

Para obtener más información sobre cómo configurar los grupos Amazon VPC de seguridad, consulte Reglas de grupos de seguridad en la Guía del usuario de Amazon VPC.

Paso 3. Configure su fuente de datos externa y Amazon VPC

Asegúrese de que la fuente de datos externa tenga la configuración de permisos y los ajustes de red correctos para acceder Amazon Kendra a ella. Encontrará instrucciones detalladas sobre cómo configurar los orígenes de datos en la sección de requisitos previos de la página de cada conector.

Además, compruebe la Amazon VPC configuración y asegúrese de que se pueda acceder a la fuente de datos externa desde la subred a la que vaya a realizar la asignación. Amazon Kendra Para ello, le recomendamos que cree una Amazon EC2 instancia en la misma subred con los mismos grupos de

seguridad y pruebe el acceso a la fuente de datos desde esta instancia. Amazon EC2 Para obtener más información, consulta Solución de problemas de Amazon VPC conexión.

Configure una fuente Amazon Kendra de datos a la que conectarse Amazon VPC

Al añadir una nueva fuente de datos Amazon Kendra, puede utilizar la Amazon VPC función si el conector de fuente de datos seleccionado es compatible con esta función.

Puede configurar una nueva fuente de Amazon Kendra datos si Amazon VPC está habilitada mediante la API AWS Management Console o la Amazon Kendra API. En concreto, utilice la operación de la API CreateDataSource y, a continuación, utilice el parámetro `VpcConfiguration` para proporcionar la siguiente información:

- `SubnetIds`— Una lista de identificadores de subredes Amazon VPC
- `SecurityGroupIds`— Una lista de identificadores de grupos de seguridad Amazon VPC

Si utiliza la consola, proporciona la Amazon VPC información necesaria durante la configuración del conector. Para utilizar la consola para habilitar la característica Amazon VPC para un conector, primero debe elegir una Amazon VPC. A continuación, debe proporcionar los identificadores de cualquier subred de Amazon VPC y los identificadores de cualquier grupo de seguridad de Amazon VPC. Puede elegir las subredes de Amazon VPC y los grupos de seguridad de Amazon VPC que creó en Configuración de Amazon VPC o utilizar cualquiera de los existentes.

Temas

- Visualización de los identificadores de Amazon VPC
- Verificación de su rol de IAM del origen de datos

Visualización de los identificadores de Amazon VPC

Los identificadores de las subredes y los grupos de seguridad se configuran en la Amazon VPC consola. Para ver los identificadores, utilice los siguientes procedimientos.

Para ver los identificadores de subred

1. Inicie sesión en la consola de Amazon VPC AWS Management Console y ábrala en https://console.aws.amazon.com/vpc/.

2. En el panel de navegación, elija Subnets (Subredes).
3. En la lista Subredes, elija la subred que contiene el servidor de base de datos.
4. En la pestaña Detalles, tome nota del identificador del campo ID de subred.

Para ver los identificadores de grupos de seguridad

1. Inicie sesión en la consola de Amazon VPC AWS Management Console y ábrala en https://console.aws.amazon.com/vpc/.
2. En el panel de navegación, elija Grupos de seguridad.
3. En la lista de grupos de seguridad, elija el grupo para el que desea el identificador.
4. En la pestaña Detalles, tome nota del identificador del campo ID de grupo de seguridad.

Verificación de su rol de IAM del origen de datos

Asegúrese de que su función de fuente de datos AWS Identity and Access Management IAM (conector) contenga permisos para acceder a su Amazon VPC.

Si usa la consola para crear un nuevo rol para su IAM rol, agrega Amazon Kendra automáticamente los permisos correctos a su IAM rol en su nombre. Si utilizas la API o utilizas un IAM rol existente, comprueba que tu rol contenga permisos de acceso Amazon VPC. Para comprobar que tiene los permisos correctos, consulte IAM Funciones para VPC.

Puede modificar un origen de datos existente para usar una subred de Amazon VPC diferente. Sin embargo, compruebe la IAM función de la fuente de datos y, si es necesario, modifíquela para que refleje el cambio y que el conector de la fuente de Amazon Kendra datos funcione correctamente.

Conexión a una base de datos en una VPC

En el ejemplo siguiente se muestra cómo conectar una base de datos de MySQL que se ejecuta en una nube privada virtual (VPC). En el ejemplo se supone que está empezando por la VPC predeterminada y que necesita crear una base de datos de MySQL. Si ya tiene una VPC, asegúrese de que esté configurada como se muestra. Si tiene una base de datos de MySQL, puede utilizarla en lugar de crear una nueva.

Pasos

- Paso 1: Configurar una VPC

- Paso 2: crear y configurar grupos de seguridad
- Paso 3: Crear una base de datos
- Paso 4: crear un conector de origen de datos

Paso 1: Configurar una VPC

Configure su VPC de modo que tenga una subred privada y un grupo de seguridad para acceder Amazon Kendra a una MySQL base de datos que se ejecuta en la subred. Las subredes proporcionadas en la configuración de la VPC deben estar en la región Oeste de EE. UU. (Oregón), en la región Este de EE. UU. (Norte de Virginia) o en la región Europa (Irlanda).

Para configurar una VPC mediante Amazon VPC

1. Inicie sesión en la consola de Amazon VPC AWS Management Console y ábrala en https://console.aws.amazon.com/vpc/.
2. En el panel de navegación, elija Route Tables (Tablas de enrutamiento) y, a continuación, elija Create route table (Crear tabla de enrutamiento).
3. En el campo Nombre, introduzca **`Private subnet route table`**. En el menú desplegable VPC, seleccione su VPC y, a continuación, elija Crear tabla de enrutamiento. Elija Close (Cerrar) para volver a la lista de tablas de enrutamiento.
4. En el panel de navegación, elija Puertas de enlace NAT y luego elija Crear puerta de enlace NAT.
5. En el menú desplegable Subred, seleccione la subred que es la subred pública. Anote los ID de subred.
6. Si no tiene una dirección IP elástica, elija Create New EIP (Crear nueva EIP), elija Create a NAT Gateway (Crear gateway NAT) y, a continuación, elija Close (Cerrar).
7. En el panel de navegación, elija Tablas de enrutamiento.
8. En la lista de tablas de enrutamiento, elija la tabla de enrutamiento de la subred privada creada en el paso 3. En Acciones, elija Editar rutas.
9. Seleccione Add route (Añadir ruta). Para el destino, introduzca **`0.0.0.0/0`** para permitir todo el tráfico saliente a Internet. En Target (Destino), elija NAT Gateway (Gateway NAT) y luego, el gateway creado en el paso 4. Elija Guardar cambios y después Cerrar.
10. En el menú Actions (Acciones), elija Edit subnet associations (Editar asociaciones de subred).
11. Elija las subredes que quiere que sean privadas. No elija la subred con la gateway NAT que ha indicado anteriormente. Elija Guardar asociaciones cuando haya terminado.

Paso 2: crear y configurar grupos de seguridad

A continuación, configure los grupos de seguridad para su base de datos.

Para crear y configurar grupos de seguridad

1. Inicie sesión en la consola de Amazon VPC AWS Management Console y ábrala en https://console.aws.amazon.com/vpc/.
2. En la descripción de la VPC, anote el CIDR de IPv4.
3. En el panel de navegación, elija Grupos de seguridad y, a continuación, elija Crear un grupo de seguridad.
4. En Nombre del grupo de seguridad, introduzca **`DataSourceInboundSecurityGroup`**. Proporcione una descripción y, a continuación, elija su VPC en la lista. Elija Crear grupo de seguridad y luego seleccione Cerrar.
5. Elija la pestaña Inbound rules (Reglas de entrada).
6. Elija Editar reglas de entrada y, a continuación, Añadir regla.
7. En una base de datos, escriba el número de puerto para Rango de puertos. Por ejemplo, para MySQL es **`3306`**, y, para HTTPS, es **`443`**. Para Source (Origen), escriba el enrutamiento entre dominios sin clases (CIDR) de la VPC. Elija Save (Guardar) y, a continuación, elija Close (Cerrar).

El grupo de seguridad permite que cualquier persona de la VPC se conecte a la base de datos y permite conexiones salientes a Internet.

Paso 3: Crear una base de datos

Cree una base de datos para guardar los documentos, o puede utilizar la base de datos existente.

Para obtener instrucciones sobre cómo crear una base de datos de MySQL, consulte MySQL.

Paso 4: crear un conector de origen de datos

Después de configurar la VPC y crear la base de datos, puede crear un conector de origen de datos para la base de datos. Para obtener información sobre los conectores de bases de datos Amazon Kendra compatibles, consulte Conectores compatibles.

Para la base de datos, asegúrese de configurar la VPC, las subredes privadas que creó en la VPC y el grupo de seguridad que creó en la VPC.

Solución de problemas de conexión de VPC

Si tiene algún problema con la conexión de la nube privada virtual (VPC), compruebe que IAM los permisos, la configuración del grupo de seguridad y las tablas de enrutamiento de la subred estén configurados correctamente.

Una posible causa de un error en la sincronización del conector de la fuente de datos es que es posible que no se pueda acceder a la fuente de datos desde la subred a la que la asignó. Amazon Kendra Para solucionar este problema, te recomendamos que crees una Amazon EC2 instancia con la misma configuración. Amazon VPC A continuación, intenta acceder a la fuente de datos desde esta Amazon EC2 instancia mediante llamadas a la API REST u otros métodos (según el tipo específico de fuente de datos).

Si accedes correctamente a la fuente de datos desde la Amazon EC2 instancia que has creado, significa que se puede acceder a la fuente de datos desde esta subred. Por lo tanto, el problema de sincronización no está relacionado con que la fuente de datos no pueda acceder a ella. Amazon VPC

Si no puedes acceder a tu Amazon EC2 instancia desde la configuración de tu VPC y validarla con la Amazon EC2 instancia que has creado, tendrás que seguir solucionando los problemas. Por ejemplo, si tienes un Amazon S3 conector cuya sincronización ha fallado debido a errores relacionados con problemas de conexión, puedes configurar una Amazon EC2 instancia con la misma Amazon VPC configuración que asignaste a tu Amazon S3 conector. A continuación, utilice esta instancia de Amazon EC2 para comprobar si la Amazon VPC se ha configurado correctamente.

A continuación, se muestra un ejemplo de cómo configurar una Amazon EC2 instancia para solucionar problemas de Amazon VPC conexión con una fuente de Amazon S3 datos.

Temas

- Paso 1: lanza una instancia Amazon EC2
- Paso 2: Conectarse a la Amazon EC2 instancia
- Paso 3: probar el acceso de Amazon S3

Paso 1: lanza una instancia Amazon EC2

1. Inicie sesión en la consola Amazon EC2 AWS Management Console y ábrala en https://console.aws.amazon.com/ec2/.
2. Seleccione Lanzar una instancia.

3. Elija Configuración de red, luego elija Editar y realice lo siguiente:

   a. Elija la misma VPC y la misma subred a las que las asignó. Amazon Kendra

   b. En Firewall (grupos de seguridad), elija Seleccionar un grupo de seguridad existente. A continuación, seleccione el grupo de seguridad al que lo asignó. Amazon Kendra

      > **Note**
      >
      > El grupo de seguridad debe permitir que el tráfico saliente a Amazon S3.

   c. Configure la opción Asignar automáticamente una IP pública en Desactivar.

   d. En Detalles avanzados, haga lo siguiente:

      - En el perfil de instancia de IAM, seleccione Crear un nuevo perfil de IAM para crear y adjuntar un perfil de IAM instancia a su instancia. Asegúrese de que el perfil tenga permisos de acceso a Amazon S3. Para obtener más información, consulta ¿Cómo puedo conceder a mi Amazon EC2 instancia acceso a un Amazon S3 bucket? en AWS re:Post.
      - Deje el resto de la configuración predeterminada.

   e. Revisa y lanza la Amazon EC2 instancia.

Paso 2: Conectarse a la Amazon EC2 instancia

Una vez que la Amazon EC2 instancia esté en ejecución, vaya a la página de detalles de la instancia y conéctese a ella. Para ello, siga estos pasos en Conéctese a las instancias sin necesidad de una dirección IPv4 pública mediante el punto de conexión de EC2 Instance Connect en la Guía del usuario de instancias de Linux de Amazon EC2 .

Paso 3: probar el acceso de Amazon S3

Una vez que te hayas conectado al terminal de tu Amazon EC2 instancia, ejecuta un AWS CLI comando para probar la conexión desde esta subred privada a tu Amazon S3 bucket.

Para probar el Amazon S3 acceso, escribe el siguiente AWS CLI comando en: AWS CLI `aws s3 ls`

Cuando se ejecute el AWS CLI comando, revise lo siguiente:

- Si has configurado correctamente los IAM permisos necesarios y tu Amazon S3 configuración es correcta, deberías ver una lista de tus Amazon S3 depósitos.

- Si ves errores de permisos, por ejemplo`Access Denied`, es probable que la configuración de tu VPC sea correcta, pero hay algún problema con tus IAM permisos o Amazon S3 con tu política de bucket.

Si se agota el tiempo de espera del comando, es probable que se esté agotando el tiempo de espera de la conexión porque la configuración de la VPC es incorrecta y la instancia de Amazon EC2 no puede acceder a Amazon S3 desde la subred. Vuelva a configurar la VPC e inténtelo de nuevo.

Eliminar un índice, un origen de datos o documentos cargados por lotes

En esta sección, se muestra cómo eliminar un índice, un repositorio de orígenes de datos de los documentos del índice o los documentos del índice que ha cargado por lotes.

Temas

- Eliminación de un índice
- Eliminación de un origen de datos
- Eliminar documentos cargados por lotes

Eliminación de un índice

Puede eliminar un índice de Amazon Kendra cuando ya no lo utilice. Por ejemplo, elimine un índice cuando:

- Ya no usa el índice y desea reducir los cargos en su cuenta de AWS. Un índice de Amazon Kendra acumula cargos mientras se está ejecutando, independientemente de si realiza consultas en el índice o no.
- Desea volver a configurar el índice para una edición diferente de Amazon Kendra. Elimine el índice existente y, a continuación, cree uno nuevo con la edición diferente.
- Ha alcanzado el número máximo de índices en su cuenta y no desea superar su cuota. Elimine un índice existente y añada uno nuevo. Para obtener más información acerca del número máximo de índices que puede crear consulte Cuotas.

Para eliminar un índice, utilice la consola, la AWS Command Line Interface, el script de AWS CloudFormation o la API `DeleteIndex`. Al eliminar un índice, se eliminan el índice y todos los orígenes de datos y datos de documentos asociados. Al eliminar un índice no se eliminan los documentos originales del almacenamiento.

La eliminación de un índice es una operación asincrónica. Al empezar a eliminar un índice, el estado del índice cambia a `DELETING`. Permanece en el estado `DELETING` hasta que se elimine toda la información relacionada con el índice. Una vez que se elimina el índice, deja de aparecer en los resultados de una llamada a la API ListIndices. Si llama a la API DescribeIndex con el identificador del índice eliminado, recibe una excepción `ResourceNotFound`.

Para eliminar un índice (consola)

1. Inicie sesión en AWS Management Console y abra la consola de Amazon Kendra en https://console.aws.amazon.com/kendra/.
2. En el panel de navegación, seleccione Índices y, a continuación, elija el índice que se va a eliminar.
3. Elija Eliminar para eliminar el índice seleccionado.

Para eliminar un índice (CLI)

- En la AWS CLI, utilice el siguiente comando. El comando tiene formato para Linux y macOS. Si está usando Windows, reemplace el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

```
aws kendra delete-index \
   --id index-id
```

Eliminación de un origen de datos

Se elimina un origen de datos cuando se quiere quitar la información contenida en el origen de datos del índice de Amazon Kendra. Por ejemplo, elimine un origen de datos cuando:

- Un origen de datos está configurado incorrectamente. Elimine el origen de datos, espere a que termine de eliminarse y, a continuación, vuelva a crearlo.
- Ha migrado documentos de un origen de datos a otro. Elimine el origen de datos original y vuelva a crearlo en la nueva ubicación.
- Ha alcanzado el límite de orígenes de datos para un índice. Elimine uno de los orígenes de datos existentes y añada uno nuevo. Para obtener más información acerca del número de orígenes de datos que puede crear, consulte Cuotas.

Para eliminar un origen de datos, utilice la consola, la AWS Command Line Interface (AWS CLI), la API `DeleteDataSource` o un script de AWS CloudFormation. Al eliminar un origen de datos, se elimina del índice toda la información sobre el origen de datos. Si solo desea detener la sincronización del origen de datos, cambie la programación de sincronización del origen de datos a “ejecución bajo demanda”.

La eliminación de un origen de datos es una operación asincrónica. Al empezar a eliminar un origen de datos, el estado del origen de datos cambia a `DELETING`. Permanece en el estado `DELETING` hasta que se elimine la información relacionada con el origen de datos. Una vez eliminado el origen de datos, ya no aparece en los resultados de una llamada a la API ListDataSources. Si llama a la API DescribeDataSource con el identificador del origen de datos eliminado, recibirá una excepción `ResourceNotFound`.

> **Note**
>
> Eliminar un origen de datos completo o volver a sincronizar el índice después de eliminar documentos específicos de un origen de datos puede tardar hasta una hora o más, según la cantidad de documentos que desee eliminar.

Para eliminar un origen de datos (consola)

1. Inicie sesión en AWS Management Console y abra la consola de Amazon Kendra en https://console.aws.amazon.com/kendra/.
2. En el panel de navegación, elija Índices y, a continuación, elija el índice que contiene el origen de datos que desea eliminar.
3. En el panel de navegación, elija Orígenes de datos.
4. Elija el origen de datos que desee eliminar.
5. Elija Eliminar para eliminar el origen de datos.

Para eliminar un origen de datos (CLI)

- En la AWS Command Line Interface, utilice el siguiente comando. El comando tiene formato para Linux y macOS. Si está usando Windows, reemplace el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

```
aws kendra delete-data-source \
   --id data-source-id \
   --index-id index-id
```

Al eliminar un origen de datos, Amazon Kendra elimina toda la información almacenada sobre el origen de datos. Amazon Kendraelimina todos los datos de documentos almacenados en el índice y

todos los historiales de ejecución y métricas asociados al origen de datos. Al eliminar un origen de datos, no se eliminan los documentos originales del almacenamiento.

Los documentos del origen de datos pueden incluirse en el recuento de documentos devuelto por la API `DescribeIndex`, mientras que Amazon Kendra elimina un origen de datos. Los documentos del origen de datos pueden aparecer en los resultados de búsqueda mientras Amazon Kendra elimina el origen de datos.

Amazon Kendra libera los recursos de un origen de datos tan pronto como se llama a la API `DeleteDataSource` o se decide eliminar el origen de datos de la consola. Si va a eliminar el origen de datos para reducir el número de orígenes de datos por debajo de su límite, puede crear un nuevo origen de datos de inmediato.

Si va a eliminar un origen de datos y, a continuación, crear otro origen de datos para los datos del documento, espere a que se elimine el primer origen de datos antes de sincronizar el nuevo origen de datos.

Puede eliminar un origen de datos que se esté sincronizando con Amazon Kendra. La sincronización se detiene y el origen de datos se elimina. Si intenta iniciar una sincronización cuando se está eliminando el origen de datos, recibirá una excepción `ConflictException`.

No puede eliminar un origen de datos si el índice asociado está en el estado `DELETING`. Al eliminar un índice, se eliminan todos los orígenes de datos del índice. Puede empezar a eliminar un índice mientras el origen de datos de ese índice esté en el estado `DELETING`.

Si tiene dos orígenes de datos que apuntan a los mismos documentos, por ejemplo, dos orígenes de datos que apuntan al mismo bucket de Amazon S3, es posible que los documentos del índice no sean coherentes si se elimina uno de los orígenes de datos. Cuando dos orígenes de datos hacen referencia a los mismos documentos, solo se almacena una copia de los datos del documento en el índice. Al eliminar un origen de datos, se eliminan los datos del índice de los documentos. El otro origen de datos no sabe que los documentos se han eliminado, por lo que no Amazon Kendra volverá a indexarlos correctamente la próxima vez que se sincronice. Si tiene dos orígenes de datos que apuntan a la misma ubicación de documento, debe eliminar ambos orígenes de datos y, a continuación, volver a crear uno.

Eliminar documentos cargados por lotes

Puede eliminar documentos directamente de un índice mediante la API BatchDeleteDocument. No puede eliminar documentos directamente con la consola. Si utiliza la consola, puede eliminar

documentos específicos del repositorio del origen de datos y volver a sincronizarlos con el índice o eliminar el conector de todo el origen de datos.

La eliminación de documentos de un índice con `BatchDeleteDocument` es una operación asincrónica. Después de llamar a la API `BatchDeleteDocument`, utilice la API BatchGetDocumentStatus para supervisar el progreso de la eliminación de los documentos. Cuando se elimina un documento del índice, Amazon Kendra devuelve `NOT_FOUND` como estado.

> **Note**
>
> Eliminar documentos de un índice utilizando `BatchDeleteDocument` puede tardar hasta una hora o más, en función del número de documentos que desee eliminar.

Para eliminar documentos cargados por lotes de un índice (CLI)

- En la AWS Command Line Interface, utilice el siguiente comando. El comando tiene formato para Linux y macOS. Si está usando Windows, reemplace el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

```
aws kendra batch-delete-document \
   --index-id index-id \
   --document-id-list 'doc-id-1' 'doc-id-2'
```

Enriquecimiento de sus documentos durante la ingesta

Puede modificar el contenido y los campos de metadatos o atributos del documento durante el proceso de ingesta de documentos. Con la característica Enriquecimiento de documentos personalizado de Amazon Kendra, puede crear, modificar o eliminar atributos y contenido del documento durante la ingesta de documentos en Amazon Kendra. Esto significa que puedes manipular e ingerir sus datos según lo necesite.

Esta característica otorga el control sobre cómo se tratan e ingieren sus documentos en Amazon Kendra. Por ejemplo, puede eliminar la información de identificación personal en los metadatos del documento mientras se ingieren los documentos en Amazon Kendra.

Otra forma de utilizar esta característica es invocar una función de Lambda en AWS Lambda para ejecutar el reconocimiento óptico de caracteres (OCR) en imágenes, traducción en texto y otras tareas para preparar los datos para la búsqueda o el análisis. Por ejemplo, puede invocar una función para ejecutar OCR en imágenes. La función podría interpretar el texto de las imágenes y tratar cada imagen como un documento textual. Una empresa que recibe encuestas de clientes enviadas por correo y las almacena como imágenes, podría ingerirlas como documentos textuales en Amazon Kendra. A continuación, la empresa puede buscar información valiosa de la encuesta de clientes en Amazon Kendra.

Puede utilizar operaciones básicas para aplicarlas como primer análisis de los datos y, a continuación, utilizar una función de Lambda para aplicar operaciones más complejas a los datos. Por ejemplo, puede utilizar una operación básica para eliminar simplemente todos los valores del campo de metadatos del documento “Customer_ID” y, a continuación, aplicar una función de Lambda para extraer texto de las imágenes del texto en los documentos.

Cómo funciona Custom Document Enrichment

El proceso general de Custom Document Enrichment es el siguiente:

1. Custom Document Enrichment se configura al crear o actualizar el origen de datos o indexar los documentos directamente en Amazon Kendra.
2. Amazon Kendra aplica configuraciones en línea o lógica básica para modificar los datos. Para obtener más información, consulte the section called “Operaciones básicas para cambiar los metadatos”.

3. Si elige configurar la manipulación avanzada de datos, Amazon Kendra puede aplicarlo en sus documentos originales, sin procesar o en los documentos estructurados y analizados. Para obtener más información, consulte the section called “Funciones de Lambda: extraer y cambiar metadatos o contenido”.
4. Los documentos modificados se ingieren en Amazon Kendra.

En cualquier momento de este proceso, si la configuración no es válida, Amazon Kendra arroja un error.

Al llamar a la API CreateDataSource, UpdateDataSource o BatchPutDocument, se proporciona la configuración de Custom Document Enrichment. Si se llama a `BatchPutDocument`, se debe configurar Custom Document Enrichment con cada solicitud. Si utiliza la consola, seleccione el índice y, a continuación, seleccione Document enrichments (Enriquecimientos de documentos) para configurar Custom Document Enrichment.

Si usa Enriquecimientos de documento en la consola, puede elegir configurar solo las operaciones básicas o solo las funciones de Lambda o ambas, del mismo modo que puede usar la API. Puede seleccionar Siguiente en los pasos de la consola para elegir no configurar las operaciones básicas y solo las funciones de Lambda, incluida la opción de aplicarlas a los datos originales (antes de la extracción) o estructurados (después de la extracción). Solo puede guardar las configuraciones si completa todos los pasos de la consola. Las configuraciones de sus documentos no se guardan si no completa todos los pasos.

Operaciones básicas para cambiar los metadatos

Puede manipular los campos y el contenido del documento mediante la lógica básica. Esto incluye la eliminación de valores de un campo, la modificación de los valores de un campo mediante una condición o la creación de un campo. Para manipulaciones avanzadas que van más allá de lo que puedes manipular con la lógica básica, invoque una función Lambda. Para obtener más información, consulte the section called “Funciones de Lambda: extraer y cambiar metadatos o contenido”.

Para aplicar la lógica básica, especifique el campo de destino que desea manipular mediante el objeto DocumentAttributeTarget. Proporcione la clave de atributo. Por ejemplo, la clave “Department” es un campo o atributo que contiene todos los nombres de departamento asociados a los documentos. También puede especificar un valor que se utilizará en el campo de destino si se cumple una condición determinada. Establezca la condición mediante el objeto DocumentAttributeCondition. Por ejemplo, si el campo “source_URI” contiene “financiero” en su

valor URI, rellene previamente el campo de destino “Department” con el valor objetivo “Finance” del documento. También puede eliminar los valores del atributo del documento de destino.

Para aplicar la lógica básica mediante la consola, seleccione el índice y, a continuación, seleccione Document enrichments(Enriquecimiento de documentos) en el menú de navegación. Vaya a Configurar operaciones básicas para aplicar manipulaciones básicas a los campos y el contenido del documento.

A continuación se muestra un ejemplo del uso de la lógica básica para eliminar todos los números de identificación de clientes del campo del documento denominado “Customer_ID”.

Ejemplo 1: Eliminación de números de identificación de clientes asociados a los documentos

Datos antes de aplicar la manipulación básica.

Document_ID	Body_Text	Customer_ID
1	Lorem Ipsum.	CID1234
2	Lorem Ipsum.	CID1235
3	Lorem Ipsum.	CID1236

Datos después de aplicar la manipulación básica.

Document_ID	Body_Text	Customer_ID
1	Lorem Ipsum.	
2	Lorem Ipsum.	
3	Lorem Ipsum.	

A continuación se muestra un ejemplo de uso de la lógica básica para crear un campo denominado “Department” y rellenar previamente este campo con los nombres de departamento en función de la información del campo “Source_URI”. Por ejemplo, si el campo “source_URI” contiene “financial” en su valor URI, rellene previamente el campo de destino “Department” con el valor objetivo “Finance” para el documento.

Ejemplo 2: Crear el campo “Department” y rellenarlo previamente con nombres de departamento asociados a los documentos mediante una condición.

Datos antes de aplicar la manipulación básica.

Document_ID	Body_Text	URI de origen
1	Lorem Ipsum.	financial/1
2	Lorem Ipsum.	financial/2
3	Lorem Ipsum.	financial/3

Datos después de aplicar la manipulación básica.

Document_ID	Body_Text	URI de origen	Department
1	Lorem Ipsum.	financial/1	Finance
2	Lorem Ipsum.	financial/2	Finance
3	Lorem Ipsum.	financial/3	Finance

> **Note**
>
> Amazon Kendra no puede crear un campo de documento de destino si aún no se ha creado como campo de índice. Después de crear el campo de índice, puede crear un campo de documento mediante `DocumentAttributeTarget`. Amazon Kendra luego asigna el campo de metadatos de documento recién creado al campo de índice.

El código siguiente es un ejemplo de configuración de la manipulación básica de datos para eliminar los números de identificación de clientes asociados a los documentos.

Console

Para configurar la manipulación básica de datos para eliminar números de identificación de clientes

1. En el panel de navegación izquierdo, en Indexes (Índices), seleccione Document enrichments (Enriquecimiento de documentos) y luego seleccione Add document enrichment (Añadir enriquecimiento de documentos).
2. En la página Configurar operaciones básicas, elija en el menú desplegable el origen de datos que desea modificar los campos de los documentos y el contenido. A continuación, elija en el menú desplegable el nombre del campo de documento “Customer_ID”, seleccione en el menú desplegable el nombre del campo de índice “Customer_ID” y seleccione en el menú desplegable la acción de destino Eliminar. A continuación, seleccione Add basic operation (Añadir operación básica).

CLI

Para configurar la manipulación básica de datos para eliminar números de identificación de clientes

```
aws kendra create-data-source \
 --name data-source-name \
 --index-id index-id \
 --role-arn arn:aws:iam::account-id:role/role-name \
 --type S3 \
 --configuration '{"S3Configuration":{"BucketName":"S3-bucket-name"}}' \
 --custom-document-enrichment-configuration '{"InlineConfigurations":[{"Target":
{"TargetDocumentAttributeKey":"Customer_ID", "TargetDocumentAttributeValueDeletion":
 true}}]}'
```

Python

Para configurar la manipulación básica de datos para eliminar números de identificación de clientes

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time
```

```
kendra = boto3.client("kendra")

print("Create a data source with customizations")

# Provide the name of the data source
name = "data-source-name"
# Provide the index ID for the data source
index_id = "index-id"
# Provide the IAM role ARN required for data sources
role_arn = "arn:aws:iam::${account-id}:role/${role-name}"
# Provide the data source connection information
data_source_type = "S3"
S3_bucket_name = "S3-bucket-name"
# Configure the data source with Custom Document Enrichment
configuration = {"S3Configuration":
        {
            "BucketName": S3_bucket_name
        }
    }
custom_document_enrichment_configuration = {"InlineConfigurations":[
        {
            "Target":{"TargetDocumentAttributeKey":"Customer_ID",
                      "TargetDocumentAttributeValueDeletion": True}
        }]
    }

try:
    data_source_response = kendra.create_data_source(
        Name = name,
        IndexId = index_id,
        RoleArn = role_arn,
        Type = data_source_type
        Configuration = configuration
        CustomDocumentEnrichmentConfiguration =
 custom_document_enrichment_configuration
    )

    pprint.pprint(data_source_response)

    data_source_id = data_source_response["Id"]

    print("Wait for Amazon Kendra to create the data source with your
 customizations.")
```

```
    while True:
        # Get the details of the data source, such as the status
        data_source_description = kendra.describe_data_source(
            Id = data_source_id,
            IndexId = index_id
        )
        status = data_source_description["Status"]
        print(" Creating data source. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

    print("Synchronize the data source.")

    sync_response = kendra.start_data_source_sync_job(
        Id = data_source_id,
        IndexId = index_id
    )

    pprint.pprint(sync_response)

    print("Wait for the data source to sync with the index.")

    while True:

        jobs = kendra.list_data_source_sync_jobs(
            Id= data_source_id,
            IndexId= index_id
        )

        # For this example, there should be one job
        status = jobs["History"][0]["Status"]

        print(" Syncing data source. Status: "+status)
        time.sleep(60)
        if status != "SYNCING":
            break

except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

Para configurar la manipulación básica de datos para eliminar números de identificación de clientes

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.DataSourceStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJob;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJobStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.S3DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;

public class CreateDataSourceWithCustomizationsExample {

    public static void main(String[] args) throws InterruptedException {
        System.out.println("Create a data source with customizations");

        String dataSourceName = "data-source-name";
        String indexId = "index-id";
        String dataSourceRoleArn = "arn:aws:iam::account-id:role/role-name";
        String s3BucketName = "S3-bucket-name"

        KendraClient kendra = KendraClient.builder().build();

        CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
            .builder()
            .name(dataSourceName)
```

```
            .description(experienceDescription)
            .roleArn(experienceRoleArn)
            .type(DataSourceType.S3)
            .configuration(
                DataSourceConfiguration
                    .builder()
                    .s3Configuration(
                        S3DataSourceConfiguration
                            .builder()
                            .bucketName(s3BucketName)
                            .build()
                    ).build()
            )
            .customDocumentEnrichmentConfiguration(
                CustomDocumentEnrichmentConfiguration
                    .builder()
                    .inlineConfigurations(Arrays.asList(
                        InlineCustomDocumentEnrichmentConfiguration
                            .builder()
                            .target(
                                DocumentAttributeTarget
                                    .builder()
                                    .targetDocumentAttributeKey("Customer_ID")
                                    .targetDocumentAttributeValueDeletion(true)
                                    .build())
                            .build()
                    )).build();

        CreateDataSourceResponse createDataSourceResponse =
kendra.createDataSource(createDataSourceRequest);
        System.out.println(String.format("Response of creating data source: %s",
createDataSourceResponse));

        String dataSourceId = createDataSourceResponse.id();
        System.out.println(String.format("Waiting for Kendra to create the data
source %s", dataSourceId));
        DescribeDataSourceRequest describeDataSourceRequest =
DescribeDataSourceRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
```

```
            DescribeDataSourceResponse describeDataSourceResponse =
kendra.describeDataSource(describeDataSourceRequest);

            DataSourceStatus status = describeDataSourceResponse.status();
            System.out.println(String.format("Creating data source. Status: %s",
status));
            TimeUnit.SECONDS.sleep(60);
            if (status != DataSourceStatus.CREATING) {
                break;
            }
        }

        System.out.println(String.format("Synchronize the data source %s",
dataSourceId));
        StartDataSourceSyncJobRequest startDataSourceSyncJobRequest =
StartDataSourceSyncJobRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();
        StartDataSourceSyncJobResponse startDataSourceSyncJobResponse =
kendra.startDataSourceSyncJob(startDataSourceSyncJobRequest);
        System.out.println(String.format("Waiting for the data
source to sync with the index %s for execution ID %s", indexId,
startDataSourceSyncJobResponse.executionId()));

        // For this example, there should be one job
        ListDataSourceSyncJobsRequest listDataSourceSyncJobsRequest =
ListDataSourceSyncJobsRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            ListDataSourceSyncJobsResponse listDataSourceSyncJobsResponse =
kendra.listDataSourceSyncJobs(listDataSourceSyncJobsRequest);
            DataSourceSyncJob job = listDataSourceSyncJobsResponse.history().get(0);
            System.out.println(String.format("Syncing data source. Status: %s",
job.status()));

            TimeUnit.SECONDS.sleep(60);
            if (job.status() != DataSourceSyncJobStatus.SYNCING) {
                break;
```

```
            }

        }

        System.out.println("Data source creation with customizations is complete");
    }
}
```

Funciones de Lambda: extraer y cambiar metadatos o contenido

Puede manipular los campos y el contenido de los documentos mediante las funciones de Lambda. Esto resulta útil si desea ir más allá de la lógica básica y aplicar manipulaciones avanzadas de datos. Por ejemplo, mediante el reconocimiento óptico de caracteres (OCR), que interpreta el texto de las imágenes y trata cada imagen como un documento textual. O bien, recuperar la fecha y hora actual en una zona horaria determinada e insertar la fecha y hora donde haya un valor vacío para un campo de fecha.

Puede aplicar primero la lógica básica y, a continuación, utilizar una función de Lambda para manipular aún más los datos o viceversa. También puede optar por aplicar solo una función de Lambda.

Amazon Kendra puede invocar una función Lambda para aplicar manipulaciones avanzadas de datos durante el proceso de ingesta como parte de suCustomDocumentEnrichmentConfiguration. Especifique un rol que incluya permiso para ejecutar la función de Lambda y acceder a su bucket de Amazon S3 para almacenar la salida de las manipulaciones de datos; consulte roles de acceso de IAM.

Amazon Kendra puede aplicar una función de Lambda en los documentos originales, sin procesar o en los documentos estructurados y analizados. Puede configurar una función Lambda que tome sus datos originales o sin procesar y aplique sus manipulaciones de datos mediante PreExtractionHookConfiguration. También puede configurar una función Lambda que tome sus documentos estructurados y aplique sus manipulaciones de datos mediante PostExtractionHookConfiguration. Amazon Kendra extrae los metadatos y el texto del documento para estructurar los documentos. Sus funciones Lambda deben seguir las estructuras obligatorias de solicitud y respuesta. Para obtener más información, consulte the section called “Contratos de datos para funciones Lambda”.

Para configurar una función Lambda en la consola, seleccione el índice y, a continuación, seleccione Document enrichments(Enriquecimiento de documentos) en el menú de navegación. Vaya a Configurar funciones Lambda para configurar una función Lambda.

Solo puede configurar una función Lambda para `PreExtractionHookConfiguration` y solo una función Lambda para `PostExtractionHookConfiguration`. Sin embargo, la función Lambda puede invocar otras funciones que requiere. Puede configurar ambos `PreExtractionHookConfiguration` y `PostExtractionHookConfiguration`, o cualquiera de los dos. La función Lambda para `PreExtractionHookConfiguration` no debe exceder un tiempo de ejecución de 5 minutos y su función Lambda para `PostExtractionHookConfiguration` no debe exceder un tiempo de ejecución de 1 minuto. La configuración de Custom Document Enrichment tarda más tiempo en ingerir sus documentos de forma natural en Amazon Kendra que si no lo configurara.

Puede configurar Amazon Kendra para invocar una función Lambda solo si se cumple una condición. Por ejemplo, puede especificar una condición que si hay valores de fecha y hora vacíos, Amazon Kendra debería invocar una función que insertara la fecha y hora actuales.

A continuación se muestra un ejemplo de uso de una función Lambda para ejecutar OCR para interpretar texto de imágenes y almacenar este texto en un campo denominado “Document_Image_Text”.

Ejemplo 1: Extracción de texto de imágenes para crear documentos textuales

Datos antes de aplicar la manipulación avanzada.

Document_ID	Document_Image
1	image_1.png
2	image_2.png
3	image_3.png

Datos después de aplicar la manipulación avanzada.

Document_ID	Document_Image	Document_Image_Text
1	image_1.png	Mailed survey response

Document_ID	Document_Image	Document_Image_Text
2	image_2.png	Mailed survey response
3	image_3.png	Mailed survey response

A continuación se muestra un ejemplo de uso de una función Lambda para insertar la fecha y hora actual para valores de fecha vacíos. Utiliza la condición de que si el valor de un campo de fecha es “null”, se sustituye por la fecha y hora actuales.

Ejemplo 2: Sustitución de valores vacíos en el campo Last_Updated por la fecha y hora actuales.

Datos antes de aplicar la manipulación avanzada.

Document_ID	Body_Text	Last_Updated
1	Lorem Ipsum.	1 de enero de 2020
2	Lorem Ipsum.	
3	Lorem Ipsum.	July 1, 2020

Datos después de aplicar la manipulación avanzada.

Document_ID	Body_Text	Last_Updated
1	Lorem Ipsum.	1 de enero de 2020
2	Lorem Ipsum.	December 1, 2021
3	Lorem Ipsum.	July 1, 2020

El siguiente código es un ejemplo de configuración de una función Lambda para la manipulación avanzada de datos en los datos originales y sin procesar.

Console

Para configurar una función Lambda para la manipulación avanzada de datos en los datos originales sin procesar

1. En el panel de navegación izquierdo, en Indexes (Índices), seleccione Document enrichments (Enriquecimiento de documentos) y luego seleccione Add document enrichment (Añadir enriquecimiento de documentos).
2. En la página Configurar funciones Lambda, en la sección Lambda para preextracción, seleccione en los menús desplegables su ARN de la función Lambda y su bucket de Amazon S3. Añada su rol de acceso de IAM seleccionando la opción de crear un nuevo rol en el menú desplegable. Esto crea los permisos de Amazon Kendra necesarios para crear el enriquecimiento de documentos.

CLI

Para configurar una función Lambda para la manipulación avanzada de datos en los datos originales sin procesar

```
aws kendra create-data-source \
 --name data-source-name \
 --index-id index-id \
 --role-arn arn:aws:iam::account-id:role/role-name \
 --type S3 \
 --configuration '{"S3Configuration":{"BucketName":"S3-bucket-name"}}' \
 --custom-document-enrichment-configuration '{"PreExtractionHookConfiguration":
{"LambdaArn":"arn:aws:iam::account-id:function/function-name", "S3Bucket":"S3-
bucket-name"}, "RoleArn": "arn:aws:iam:account-id:role/cde-role-name"}'
```

Python

Para configurar una función Lambda para la manipulación avanzada de datos en los datos originales sin procesar

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")
```

```
print("Create a data source with customizations.")

# Provide the name of the data source
name = "data-source-name"
# Provide the index ID for the data source
index_id = "index-id"
# Provide the IAM role ARN required for data sources
role_arn = "arn:aws:iam::${account-id}:role/${role-name}"
# Provide the data source connection information
data_source_type = "S3"
S3_bucket_name = "S3-bucket-name"
# Configure the data source with Custom Document Enrichment
configuration = {"S3Configuration":
        {
            "BucketName": S3_bucket_name
        }
    }
custom_document_enrichment_configuration = {"PreExtractionHookConfiguration":
        {
            "LambdaArn":"arn:aws:iam::account-id:function/function-name",
            "S3Bucket":"S3-bucket-name"
        }
    "RoleArn":"arn:aws:iam::account-id:role/cde-role-name"
    }

try:
    data_source_response = kendra.create_data_source(
        Name = name,
        IndexId = index_id,
        RoleArn = role_arn,
        Type = data_source_type
        Configuration = configuration
        CustomDocumentEnrichmentConfiguration =
 custom_document_enrichment_configuration
    )

    pprint.pprint(data_source_response)

    data_source_id = data_source_response["Id"]

    print("Wait for Amazon Kendra to create the data source with your
 customizations.")
```

```
    while True:
        # Get the details of the data source, such as the status
        data_source_description = kendra.describe_data_source(
            Id = data_source_id,
            IndexId = index_id
        )
        status = data_source_description["Status"]
        print(" Creating data source. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

    print("Synchronize the data source.")

    sync_response = kendra.start_data_source_sync_job(
        Id = data_source_id,
        IndexId = index_id
    )

    pprint.pprint(sync_response)

    print("Wait for the data source to sync with the index.")

    while True:

        jobs = kendra.list_data_source_sync_jobs(
            Id = data_source_id,
            IndexId = index_id
        )

        # For this example, there should be one job
        status = jobs["History"][0]["Status"]

        print(" Syncing data source. Status: "+status)
        time.sleep(60)
        if status != "SYNCING":
            break

except  ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

Para configurar una función Lambda para la manipulación avanzada de datos en los datos originales sin procesar

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.DataSourceStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJob;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJobStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.S3DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;


public class CreateDataSourceWithCustomizationsExample {

    public static void main(String[] args) throws InterruptedException {
        System.out.println("Create a data source with customizations");

        String dataSourceName = "data-source-name";
        String indexId = "index-id";
        String dataSourceRoleArn = "arn:aws:iam::account-id:role/role-name";
        String s3BucketName = "S3-bucket-name"

        KendraClient kendra = KendraClient.builder().build();

        CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
            .builder()
```

```
            .name(dataSourceName)
            .description(experienceDescription)
            .roleArn(experienceRoleArn)
            .type(DataSourceType.S3)
            .configuration(
                DataSourceConfiguration
                    .builder()
                    .s3Configuration(
                        S3DataSourceConfiguration
                            .builder()
                            .bucketName(s3BucketName)
                            .build()
                    ).build()
            )
            .customDocumentEnrichmentConfiguration(
                CustomDocumentEnrichmentConfiguration
                    .builder()
                    .preExtractionHookConfiguration(
                        HookConfiguration
                            .builder()
                            .lambdaArn("arn:aws:iam::account-id:function/function-
name")
                            .s3Bucket("S3-bucket-name")
                            .build())
                    .roleArn("arn:aws:iam::account-id:role/cde-role-name")
                    .build();

        CreateDataSourceResponse createDataSourceResponse =
 kendra.createDataSource(createDataSourceRequest);
        System.out.println(String.format("Response of creating data source: %s",
 createDataSourceResponse));

        String dataSourceId = createDataSourceResponse.id();
        System.out.println(String.format("Waiting for Kendra to create the data
 source %s", dataSourceId));
        DescribeDataSourceRequest describeDataSourceRequest =
 DescribeDataSourceRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
```

```
            DescribeDataSourceResponse describeDataSourceResponse =
kendra.describeDataSource(describeDataSourceRequest);

            DataSourceStatus status = describeDataSourceResponse.status();
            System.out.println(String.format("Creating data source. Status: %s",
status));
            TimeUnit.SECONDS.sleep(60);
            if (status != DataSourceStatus.CREATING) {
                break;
            }
        }

        System.out.println(String.format("Synchronize the data source %s",
dataSourceId));
        StartDataSourceSyncJobRequest startDataSourceSyncJobRequest =
StartDataSourceSyncJobRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();
        StartDataSourceSyncJobResponse startDataSourceSyncJobResponse =
kendra.startDataSourceSyncJob(startDataSourceSyncJobRequest);
        System.out.println(String.format("Waiting for the data
source to sync with the index %s for execution ID %s", indexId,
startDataSourceSyncJobResponse.executionId()));

        // For this example, there should be one job
        ListDataSourceSyncJobsRequest listDataSourceSyncJobsRequest =
ListDataSourceSyncJobsRequest
            .builder()
            .indexId(indexId)
            .id(dataSourceId)
            .build();

        while (true) {
            ListDataSourceSyncJobsResponse listDataSourceSyncJobsResponse =
kendra.listDataSourceSyncJobs(listDataSourceSyncJobsRequest);
            DataSourceSyncJob job = listDataSourceSyncJobsResponse.history().get(0);
            System.out.println(String.format("Syncing data source. Status: %s",
job.status()));

            TimeUnit.SECONDS.sleep(60);
            if (job.status() != DataSourceSyncJobStatus.SYNCING) {
                break;
```

```
            }

        }

        System.out.println("Data source creation with customizations is complete");
    }
}
```

Contratos de datos para funciones Lambda

Sus funciones Lambda para la manipulación avanzada de datos interactúan con contratos de datos de Amazon Kendra. Los contratos son las estructuras de solicitud y respuesta obligatorias de sus funciones Lambda. Si sus funciones Lambda no siguen estas estructuras, Amazon Kendra arroja un error.

La función Lambda para `PreExtractionHookConfiguration` debería esperar la siguiente estructura de solicitud:

```
{
    "version": <str>,
    "dataBlobStringEncodedInBase64": <str>, //In the case of a data blob
    "s3Bucket": <str>, //In the case of an S3 bucket
    "s3ObjectKey": <str>, //In the case of an S3 bucket
    "metadata": <Metadata>
}
```

La estructura de `metadata`, que incluye la estructura de `CustomDocumentAttribute`, es la siguiente:

```
{
    "attributes": [<CustomDocumentAttribute<]
}

CustomDocumentAttribute
{
    "name": <str>,
    "value": <CustomDocumentAttributeValue>
}

CustomDocumentAttributeValue
```

```
{
    "stringValue": <str>,
    "integerValue": <int>,
    "longValue": <long>,
    "stringListValue": list<str>,
    "dateValue": <str>
}
```

La función Lambda para `PreExtractionHookConfiguration` debe cumplir la siguiente estructura de respuesta:

```
{
    "version": <str>,
    "dataBlobStringEncodedInBase64": <str>, //In the case of a data blob
    "s3ObjectKey": <str>, //In the case of an S3 bucket
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

La función Lambda para `PostExtractionHookConfiguration` debería esperar la siguiente estructura de solicitud:

```
{
    "version": <str>,
    "s3Bucket": <str>,
    "s3ObjectKey": <str>,
    "metadata": <Metadata>
}
```

La función Lambda para `PostExtractionHookConfiguration` debe cumplir la siguiente estructura de respuesta:

```
PostExtractionHookConfiguration Lambda Response
{
    "version": <str>,
    "s3ObjectKey": <str>,
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

El documento modificado se carga en su bucket de Amazon S3. El documento modificado debe seguir el formato que se muestra en the section called “Formato del documento estructurado”.

Formato del documento estructurado

Amazon Kendra carga su documento estructurado en el bucket de Amazon S3 determinado. El documento estructurado sigue este formato:

```
Kendra document

{
   "textContent": <TextContent>
}

TextContent
{
  "documentBodyText": <str>
}
```

Ejemplo de una función Lambda que se adhiere a los contratos de datos

El siguiente código de Python es un ejemplo de una función Lambda que aplica manipulación avanzada de los campos de metadatos `_authors`, `_document_title` y el contenido del cuerpo de los documentos originales o sin procesar.

En el caso del contenido del cuerpo que reside en un bucket de Amazon S3

```
import json
import boto3

s3 = boto3.client("s3")

# Lambda function for advanced data manipulation
def lambda_handler(event, context):
    # Get the value of "S3Bucket" key name or item from the given event input
    s3_bucket = event.get("s3Bucket")
    # Get the value of "S3ObjectKey" key name or item from the given event input
    s3_object_key = event.get("s3ObjectKey")

    content_object_before_CDE = s3.get_object(Bucket = s3_bucket, Key = s3_object_key)
    content_before_CDE = content_object_before_CDE["Body"].read().decode("utf-8");
    content_after_CDE = "CDEInvolved " + content_before_CDE

    # Get the value of "metadata" key name or item from the given event input
```

```
    metadata = event.get("metadata")
    # Get the document "attributes" from the metadata
    document_attributes = metadata.get("attributes")

    s3.put_object(Bucket = s3_bucket, Key = "dummy_updated_kendra_document",
 Body=json.dumps(content_after_CDE))
    return {
        "version": "v0",
        "s3ObjectKey": "dummy_updated_kendra_document",
        "metadataUpdates": [
            {"name":"_document_title", "value":
{"stringValue":"title_from_pre_extraction_lambda"}},
            {"name":"_authors", "value":{"stringListValue":["author1", "author2"]}}
        ]
    }
```

En el caso del contenido del cuerpo que reside en un blob de datos

```
import json
import boto3
import base64

# Lambda function for advanced data manipulation
def lambda_handler(event, context):

    # Get the value of "dataBlobStringEncodedInBase64" key name or item from the given
 event input
    data_blob_string_encoded_in_base64 = event.get("dataBlobStringEncodedInBase64")
    # Decode the data blob string in UTF-8
    data_blob_string =
 base64.b64decode(data_blob_string_encoded_in_base64).decode("utf-8")
    # Get the value of "metadata" key name or item from the given event input
    metadata = event.get("metadata")
    # Get the document "attributes" from the metadata
    document_attributes = metadata.get("attributes")

    new_data_blob = "This should be the modified data in the document by pre processing
 lambda ".encode("utf-8")
    return {
        "version": "v0",
        "dataBlobStringEncodedInBase64":
 base64.b64encode(new_data_blob).decode("utf-8"),
        "metadataUpdates": [
```

```
            {"name":"_document_title", "value":
{"stringValue":"title_from_pre_extraction_lambda"}},
            {"name":"_authors", "value":{"stringListValue":["author1", "author2"]}}
        ]
    }
```

El siguiente código de Python es un ejemplo de una función Lambda que aplica manipulación avanzada de los campos de metadatos `_authors`, `_document_title` y el contenido del cuerpo de los documentos estructurados o analizados.

```
import json
import boto3
import time

s3 = boto3.client("s3")

# Lambda function for advanced data manipulation
def lambda_handler(event, context):

    # Get the value of "S3Bucket" key name or item from the given event input
    s3_bucket = event.get("s3Bucket")
    # Get the value of "S3ObjectKey" key name or item from the given event input
    s3_key = event.get("s3ObjectKey")
    # Get the value of "metadata" key name or item from the given event input
    metadata = event.get("metadata")
    # Get the document "attributes" from the metadata
    document_attributes = metadata.get("attributes")

    kendra_document_object = s3.get_object(Bucket = s3_bucket, Key = s3_key)
    kendra_document_string = kendra_document_object['Body'].read().decode('utf-8')
    kendra_document = json.loads(kendra_document_string)
    kendra_document["textContent"]["documentBodyText"] = "Changing document body to a
 short sentence."

    s3.put_object(Bucket = s3_bucket, Key = "dummy_updated_kendra_document",
 Body=json.dumps(kendra_document))

    return {
        "version" : "v0",
        "s3ObjectKey": "dummy_updated_kendra_document",
        "metadataUpdates": [
            {"name": "_document_title", "value":{"stringValue":
 "title_from_post_extraction_lambda"}},
```

```
            {"name": "_authors", "value":{"stringListValue":["author1", "author2"]}}
        ]
    }
```

Búsqueda en un índice

Para buscar un Amazon Kendra índice, se utiliza la API de consultas. La API `Query` devuelve información sobre los documentos indexados que utiliza en su aplicación. En esta sección se muestra cómo realizar una consulta, aplicar filtros e interpretar la respuesta que obtiene de la API `Query`.

Para buscar documentos con Amazon Kendra los que ha indexado Amazon Lex, utilice AMAZON.KendraSearchIntent. Para ver un ejemplo de configuración Amazon Kendra con Amazon Lex, consulte Creación de un bot de preguntas frecuentes para un Amazon Kendra índice.

Temas

- Consulta de un índice
- Navegar por un índice
- Destacar resultados de búsqueda
- Búsqueda tabular de HTML
- Sugerencias de consulta
- Corrector ortográfico de las consultas
- Filtrado y búsqueda por facetas
- Filtrar por contexto de usuario
- Respuestas a las consultas y tipos de respuestas
- Ajustar y ordenar las respuestas
- Contraer o expandir los resultados de la consulta

Consulta de un índice

Cuando busca en su índice, Amazon Kendra utiliza toda la información que ha proporcionado sobre sus documentos para determinar los documentos más relevantes para los términos de búsqueda introducidos. Algunos de los elementos que se tienen Amazon Kendra en cuenta son:

- El texto o el cuerpo del documento.
- El título del documento.
- Campos de texto personalizados que ha marcado como que se pueden buscar.

- El campo de fecha que ha indicado que debe usarse para determinar la “antigüedad” de un documento.
- Cualquier otro campo que pueda proporcionar información relevante.

Amazon Kendra también puede filtrar la respuesta en función de cualquier filtro de campo o atributo que haya establecido para la búsqueda. Por ejemplo, si tiene un campo personalizado denominado “department”, puede filtrar la respuesta para que muestre únicamente los documentos de un departamento denominado “legal”. Para obtener más información, consulte Campos o atributos personalizados.

Los resultados de la búsqueda devueltos se ordenan según la relevancia que se Amazon Kendra determine para cada documento. Los resultados están paginados para que pueda mostrar una página a la vez al usuario.

Para buscar documentos con Amazon Kendra los que ha indexado Amazon Lex, utilice AMAZON.KendraSearchIntent. Para ver un ejemplo de configuración Amazon Kendra con Amazon Lex, consulte Creación de un bot de preguntas frecuentes para un Amazon Kendra índice.

El siguiente ejemplo muestra cómo buscar en un índice. Amazon Kendra determina el tipo de resultado de la búsqueda (respuesta, documento, pregunta-respuesta) que mejor se adapta a la consulta. No se puede configurar Amazon Kendra para que devuelva un tipo específico de respuesta de búsqueda (respuesta, documento, pregunta-respuesta) a una consulta.

Para obtener más información acerca de las respuestas a las consultas, vea Respuestas a las consultas y tipos de respuestas.

Requisitos previos

Antes de usar la API de consulta para consultar un índice:

- Configure los permisos necesarios para un índice y conéctese a su origen de datos o cargue sus documentos por lotes. Para obtener más información, consulte Roles de IAM. Debe utilizar el nombre de recursos de Amazon del rol cuando llama a la API para crear un conector de índice y origen de datos o para cargar documentos por lotes.
- Configura un SDK o ve a la Amazon Kendra consola. AWS Command Line Interface Para más información, consulte Configuración Amazon Kendra.

- Cree un índice y conéctese a un origen de datos de documentos o cargue documentos por lotes. Para obtener más información, consulte Creación de un índice y Creación de un conector de origen de datos.

Buscar en un índice (consola)

Puedes usar la Amazon Kendra consola para buscar y probar tu índice. Puede realizar consultas y ver los resultados.

Para buscar en un índice con la consola

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala en http://console.aws.amazon.com/kendra/.
2. En el panel de navegación, elija Índices.
3. Elija su índice.
4. En el menú de navegación, elija la opción para buscar en el índice.
5. Escriba una consulta en el cuadro de texto y, a continuación, pulse Intro.
6. Amazon Kendra devuelve los resultados de la búsqueda.

También puede obtener el ID de consulta para la búsqueda seleccionando el icono de la bombilla en el panel lateral.

Buscar en un índice (SDK)

Para buscar en un índice con Python o Java

- En el siguiente ejemplo se busca en un índice. Cambie el valor de `query` a su consulta de búsqueda y `index_id` o `indexId` al identificador de índice del índice en el que desee buscar.

  También puede obtener el ID de consulta de la búsqueda como parte de los elementos de respuesta cuando llama a la API de consulta.

  Python

```
import boto3
import pprint

kendra = boto3.client("kendra")
```

```
# Provide the index ID
index_id = "index-id"
# Provide the query text
query = "query text"

response = kendra.query(
        QueryText = query,
        IndexId = index_id)

print("\nSearch results for query: " + query + "\n")

for query_result in response["ResultItems"]:

    print("-------------------")
    print("Type: " + str(query_result["Type"]))

    if query_result["Type"]=="ANSWER" or
 query_result["Type"]=="QUESTION_ANSWER":
        answer_text = query_result["DocumentExcerpt"]["Text"]
        print(answer_text)

    if query_result["Type"]=="DOCUMENT":
        if "DocumentTitle" in query_result:
            document_title = query_result["DocumentTitle"]["Text"]
            print("Title: " + document_title)
        document_text = query_result["DocumentExcerpt"]["Text"]
        print(document_text)

    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResponse;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
```

```
        KendraClient kendra = KendraClient.builder().build();

        String query = "query text";
        String indexId = "index-id";

        QueryRequest queryRequest = QueryRequest
            .builder()
            .queryText(query)
            .indexId(indexId)
            .build();

        QueryResponse queryResponse = kendra.query(queryRequest);

        System.out.println(String.format("\nSearch results for query: %s",
 query));
        for(QueryResultItem item: queryResponse.resultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Type: %s", item.type()));

            switch(item.type()) {
                case QUESTION_ANSWER:
                case ANSWER:
                    String answerText = item.documentExcerpt().text();
                    System.out.println(answerText);
                    break;
                case DOCUMENT:
                    String documentTitle = item.documentTitle().text();
                    System.out.println(String.format("Title: %s",
 documentTitle));
                    String documentExcerpt = item.documentExcerpt().text();
                    System.out.println(String.format("Excerpt: %s",
 documentExcerpt));
                    break;
                default:
                    System.out.println(String.format("Unknown query result type:
 %s", item.type()));

            }

            System.out.println("----------------------\n");
        }
    }
}
```

Buscar en un índice (Postman)

Puede usar Postman para consultar y probar su Amazon Kendra índice.

Para buscar en un índice mediante Postman

1. Cree una nueva colección en Postman y establezca el tipo de solicitud en POST.
2. Introduzca la URL del punto de conexión. Por ejemplo, https://kendra.<region>.amazonaws.com.
3. Seleccione la pestaña Autorización e introduzca la siguiente información.

   - Tipo: seleccione la firma de AWS .
   - AccessKey—Introduzca la clave de acceso generada al crear un IAM usuario.
   - SecretKey—Introduzca la clave secreta generada al crear un IAM usuario.
   - AWS Región: introduzca la región de su índice. Por ejemplo, us-west-2.
   - Nombre del servicio: introduzca kendra. Se distingue entre mayúsculas y minúsculas, por lo que debe estar en minúsculas.

   > **Warning**
   >
   > Si escribe un nombre de servicio incorrecto o no utiliza minúsculas, aparecerá un mensaje de error al seleccionar Enviar para enviar la solicitud: “Credential should be scoped to the correct service 'kendra'”.
   > También debe comprobar que ha introducido la clave de acceso y la clave secreta.

4. Seleccione la pestaña Encabezados e introduzca la siguiente información de clave y valor.

   - Clave: X-Amz-Target

     Valor: com.amazonaws.kendra. AWSKendraFrontendService.Consulta
   - Clave: Content-Encoding

     Valor: amz-1.0

5. Seleccione la pestaña Cuerpo y haga lo siguiente.

   - Elija el tipo JSON sin procesar para el cuerpo de la solicitud.
   - Introduzca un JSON que incluya su ID de índice y el texto de la consulta.

   ```
   {
   ```

```
    "IndexId": "index-id",
    "QueryText": "enter a query here"
}
```

> **Warning**
>
> Si tu JSON no usa la indentación correcta, aparece un error: "». SerializationException Compruebe la indentación en su JSON.

6. Seleccione Enviar (cerca de la esquina superior derecha).

Búsqueda con una sintaxis de consulta avanzada

Puede crear consultas que sean más específicas que las consultas simples de palabras clave o lenguaje natural mediante operadores o sintaxis de consulta avanzados. Por ejemplo, puede utilizar rangos, operadores booleanos, caracteres comodín y mucho más. Al usar operadores, puede dar más contexto a la consulta y refinar aún más los resultados de búsqueda.

Amazon Kendra admite los siguientes operadores.

- Operadores booleanos: lógica para limitar o ampliar la búsqueda. Por ejemplo, `amazon AND sports` limita la búsqueda para que solo busque documentos que contengan ambos términos.
- Paréntesis: lee los términos de consulta anidados en orden de prioridad. Por ejemplo, `(amazon AND sports) NOT rainforest` lee `(amazon AND sports)` antes que `NOT rainforest`.
- Rangos: valores de rango numérico o de fecha. Los rangos pueden ser inclusivos, exclusivos o ilimitados. Por ejemplo, puede buscar documentos que se actualizaron por última vez entre el 1 de enero de 2020 y el 31 de diciembre de 2020, con ambas fechas incluidas.
- Campos: utiliza un campo específico para limitar la búsqueda. Por ejemplo, puede buscar documentos que tengan “Estados Unidos” en el campo “ubicación”.
- Caracteres comodín: coinciden parcialmente con una cadena de texto. Por ejemplo, `Cloud*` podría coincidir CloudFormation. Amazon Kendra actualmente solo admite caracteres comodín al final.
- Citas exactas: coinciden exactamente con una cadena de texto. Por ejemplo, los documentos que contienen `"Amazon Kendra" "pricing"`.

Puede utilizar combinaciones de cualquiera de los operadores anteriores.

Tenga en cuenta que el uso excesivo de operadores o de consultas muy complejas podría afectar a la latencia de las consultas. Los caracteres comodín son algunos de los operadores que más afectan en términos de latencia. Por regla general, cuantos más términos y operadores utilice, mayor será el impacto potencial en la latencia. Otros factores que afectan a la latencia son el tamaño medio de los documentos indexados, el tamaño del índice, cualquier filtrado de los resultados de búsqueda y la carga total del índice. Amazon Kendra

Booleano

Puede combinar o excluir palabras mediante los operadores booleanos `AND`, `OR` y `NOT`.

Los siguientes ejemplos muestran el uso de los operadores booleanos.

`amazon AND sports`

Devuelve los resultados de búsqueda que contienen los términos “amazon” y “sports” en el texto, como vídeos de deporte de Amazon Prime u otro contenido similar.

`sports OR recreation`

Devuelve los resultados de búsqueda que contienen los términos “sports” o “recreation”, o ambos, en el texto.

`amazon NOT rainforest`

Devuelve los resultados de búsqueda que contienen el término “amazon” pero no el término “rainforest” en el texto. Se utiliza para buscar documentos sobre la empresa Amazon, no sobre la selva amazónica.

Paréntesis

Puede consultar palabras anidadas en orden de prioridad utilizando paréntesis. Los paréntesis indican Amazon Kendra cómo debe leerse una consulta.

Los siguientes ejemplos muestran el uso de los paréntesis como operadores.

`(amazon AND sports) NOT rainforest`

Devuelve los documentos que contienen los términos “amazon” y “sports” en el texto, pero no el término “rainforest”. Se utiliza para buscar contenidos como vídeos de deporte de Amazon Prime u otro contenido similar, y no para buscar deportes de aventura en la selva amazónica. Los paréntesis

indican que `amazon AND sports` debe leerse antes que `NOT rainforest`. La consulta no debe leerse como `amazon AND (sports NOT rainforest)`.

`(amazon AND (sports OR recreation)) NOT rainforest`

Devuelve documentos que contienen los términos “sports” o “recreation”, o ambos, y el término “amazon”. Sin embargo, no incluye el término “rainforest”. Se utiliza para buscar vídeos de deporte de Amazon Prime u otros contenidos de ocio, y no para buscar deportes de aventura en la selva amazónica. Los paréntesis indican que `sports OR recreation` debe leerse antes de combinarlo con “amazon”, que se lee antes que `NOT rainforest`. La consulta no debe leerse como `amazon AND (sports OR (recreation NOT rainforest))`.

Rangos

Puede utilizar un rango de valores para filtrar los resultados de búsqueda. Debe especificar un atributo y el rango de valores. Puede ser de tipo numérico o de fecha.

Los rangos de fechas deben tener los siguientes formatos:

- Epoch
- AAAA
- AAAA-mm
- AAAA-mm-dd
- AAAA-mm-dd'T'HH

También puede especificar si desea incluir o excluir los valores inferior y superior del rango.

Los siguientes ejemplos muestran el uso de los rangos como operadores.

`_processed_date:>2019-12-31 AND _processed_date:<2021-01-01`

Devuelve los documentos que se procesaron en 2020, es decir, más tarde que el 31 de diciembre de 2019 y antes que el 1 de enero de 2021.

`_processed_date:>=2020-01-01 AND _processed_date:<=2020-12-31`

Devuelve los documentos que se procesaron en 2020, es decir, entre el 1 de enero de 2020 o más tarde y hasta el 31 de diciembre de 2020 o antes.

`_document_likes:<1`

Devuelve los documentos sin likes o sin valoraciones de los usuarios (menos de 1 like).

Puede especificar si un rango debe considerarse inclusivo o exclusivo de los valores del rango dados.

Inclusivo

`_last_updated_at:[2020-01-01 TO 2020-12-31]`

Devuelve los documentos actualizados por última vez en 2020; incluye los días 1 de enero de 2020 y 31 de diciembre de 2020.

Exclusivo

`_last_updated_at:{2019-12-31 TO 2021-01-01}`

Devuelve los documentos actualizados por última vez en 2020; excluye los días 31 de diciembre de 2019 y 1 de enero de 2021.

Para usar rangos ilimitados que no sean inclusivos ni exclusivos, simplemente use los operadores **<** y **>**. Por ejemplo, `_last_updated_at:>2019-12-31 AND _last_updated_at:<2021-01-01`

Campos

Puede limitar la búsqueda para que solo se devuelvan los documentos que cumplan un valor en un campo específico. El campo puede ser de cualquier tipo.

A continuación se muestran ejemplos del uso de operadores de contexto en el campo.

`status:"Incomplete" AND financial_year:2021`

Devuelve los documentos del ejercicio fiscal de 2021 con el estado incompleto.

`(sports OR recreation) AND country:"United States" AND level:"professional"`

Devuelve documentos relacionados con el ocio o el deporte profesional en los Estados Unidos.

Caracteres comodín

Puede ampliar la búsqueda para incluir variantes de palabras y frases utilizando el operador de carácter comodín. Esto resulta útil cuando se buscan variantes de nombres. Amazon Kendra actualmente solo admite caracteres comodín al final. El número de caracteres del prefijo de un carácter comodín final debe ser superior a dos.

Los siguientes ejemplos muestran el uso de los caracteres comodín como operadores.

`Cloud*`

Devuelve documentos que contienen variantes como CloudFormation y. CloudWatch

`kendra*aws`

Devuelve documentos que contienen variantes como kendra.amazonaws.

`kendra*aws*`

Devuelve documentos que contienen variantes como kendra.amazonaws.com.

Citas exactas

Puede utilizar las comillas para buscar una coincidencia exacta de un fragmento de texto.

A continuación se muestran ejemplos del uso de las comillas.

`"Amazon Kendra" "pricing"`

Devuelve los documentos que contienen “Amazon Kendra” y el término “pricing”. Los documentos deben incluir tanto “Amazon Kendra” como “pricing” para poder mostrarlos en los resultados.

`"Amazon Kendra" "pricing" cost`

Devuelve los documentos que contienen “Amazon Kendra” y el término “pricing”, y opcionalmente el término “cost”. Los documentos deben incluir tanto “Amazon Kendra” como “pricing” para poder mostrarlos en los resultados, pero no deben incluir necesariamente “cost”.

Sintaxis de consulta no válida

Amazon Kendra emite una advertencia si hay problemas con la sintaxis de la consulta o si la consulta no es compatible actualmente con Amazon Kendra. Para obtener más información, consulte la documentación de la API sobre las advertencias de las consultas.

Las siguientes consultas muestran ejemplos de sintaxis de consulta no válida.

`_last_updated_at:<2021-12-32`

Fecha no válida. El día 32 no existe en el calendario gregoriano, que es el que utiliza Amazon Kendra.

`_view_count:ten`

Valor numérico no válido. Se deben usar dígitos para representar valores numéricos.

`nonExistentField:123`

Búsqueda de campo no válida. El campo debe existir para poder utilizar la búsqueda de campos.

`Product:[A TO D]`

Rango no válido. Se deben usar valores numéricos o fechas para los rangos.

`OR Hello`

Operador booleano no válido. Los operadores deben usarse con términos y colocarse entre términos.

Buscar en otros idiomas

Puede buscar documentos en un idioma compatible. Debe introducir el código de idioma AttributeFilterpara que se devuelvan los documentos filtrados en el idioma que elija. Puede escribir la consulta en un idioma compatible.

Si no especifica un idioma, Amazon Kendra consulta los documentos en inglés de forma predeterminada. Para obtener más información acerca de los idiomas admitidos, incluidos sus códigos, consulte Adición de documentos en idiomas distintos del inglés.

Para buscar documentos en un idioma compatible en la consola, seleccione el índice y, a continuación, seleccione la opción de buscar en el índice en el menú de navegación. Elija el idioma en el que desea que se muestren los documentos. Para ello, seleccione la configuración de búsqueda y, a continuación, seleccione un idioma en el menú desplegable Idioma.

En los siguientes ejemplos se muestra cómo buscar documentos en español.

Para buscar un índice en español en la consola

1. Inicie sesión AWS Management Console y abra la Amazon Kendra consola en http://console.aws.amazon.com/kendra/.
2. En el menú de navegación, elija Índices y, a continuación, elija su índice.
3. En el menú de navegación, elija la opción para buscar en el índice.

4. En la configuración de búsqueda, seleccione el menú desplegable Idiomas y elija español.
5. Escriba una consulta en el cuadro de texto y, a continuación, pulse Intro.
6. Amazon Kendra devuelve los resultados de la búsqueda en español.

Para buscar un índice en español mediante la CLI, Python o Java

- En el siguiente ejemplo se busca en un índice en español. Cambie el valor de `searchString` a su consulta de búsqueda y el valor de `indexID` al identificador del índice en el que desee buscar. El código de idioma de español es `es`. Puede reemplazarlo por el código de su idioma.

  CLI

```
{
  "EqualsTo":{
    "Key": "_language_code",
    "Value": {
    "StringValue": "es"
    }
  }
}
```

  Python

```
import boto3
import pprint

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the query text
query = "search-string"

# Includes the index ID, query text, and language attribute filter
response = kendra.query(
        QueryText = query,
        IndexId = index_id,
        AttributeFilter = {
            "EqualsTo": {
                "Key": "_language_code",
                "Value": {
```

```
                    "StringValue": "es"
                    }
                }
            })

print ("\nSearch results|Resultados de la búsqueda: " + query + "\n")

for query_result in response["ResultItems"]:

    print("-------------------")
    print("Type: " + str(query_result["Type"]))

    if query_result["Type"]=="ANSWER" or
 query_result["Type"]=="QUESTION_ANSWER":
        answer_text = query_result["DocumentExcerpt"]["Text"]
        print(answer_text)

    if query_result["Type"]=="DOCUMENT":
        if "DocumentTitle" in query_result:
            document_title = query_result["DocumentTitle"]["Text"]
            print("Title: " + document_title)
        document_text = query_result["DocumentExcerpt"]["Text"]
        print(document_text)

    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResponse;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String query = "searchString";
        String indexId = "indexID";
```

```
        QueryRequest queryRequest = QueryRequest.builder()
            .queryText(query)
            .indexId(indexId)
            .attributeFilter(
                 AttributeFilter.builder()
                     .withEqualsTo(
                         DocumentAttribute.builder()
                             .withKey("_language_code")
                             .withValue("es")
                             .build())
                     .build())
            .build();

        QueryResponse queryResponse = kendra.query(queryRequest);

        System.out.println(String.format("\nSearch results|
                                          Resultados de la búsqueda: %s",
query));
        for(QueryResultItem item: queryResponse.resultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Type: %s", item.type()));

            switch(item.type()) {
                case QUESTION_ANSWER:
                case ANSWER:
                    String answerText = item.documentExcerpt().text();
                    System.out.println(answerText);
                    break;
                case DOCUMENT:
                    String documentTitle = item.documentTitle().text();
                    System.out.println(String.format("Title: %s",
documentTitle));
                    String documentExcerpt = item.documentExcerpt().text();
                    System.out.println(String.format("Excerpt: %s",
documentExcerpt));
                    break;
                default:
                    System.out.println(String.format("Unknown query result type:
%s", item.type()));

            }

            System.out.println("----------------------\n");
```

```
        }
    }
}
```

Recuperación de pasajes

Puede utilizar la API Retrieve como recuperador para los sistemas de generación aumentada de recuperación (RAG).

Los sistemas RAG utilizan inteligencia artificial generativa para crear aplicaciones de preguntas y respuestas. Los sistemas RAG constan de un recuperador y modelos de lenguaje grandes (LLM). Al realizar una consulta, el recuperador identifica los fragmentos de texto más relevantes de un corpus de documentos y los envía al LLM para proporcionar la respuesta más útil. Luego, el LLM analiza los fragmentos de texto relevantes y genera una respuesta integral para la consulta.

La API `Retrieve` analiza los fragmentos de texto o los extractos, que se denominan pasajes, y devuelve los pasajes principales que son más relevantes para la consulta.

Al igual que la API Query, la API `Retrieve` también busca información relevante mediante la búsqueda semántica. La búsqueda semántica tiene en cuenta el contexto de la consulta de búsqueda, además de toda la información disponible en los documentos indexados. Sin embargo, de forma predeterminada, la API `Query` solo devuelve fragmentos o pasajes de hasta 100 palabras simbólicas. Con la API `Retrieve`, puede recuperar pasajes más largos de hasta 200 palabras simbólicas y hasta 100 pasajes semánticamente relevantes. Esto no incluye las respuestas de tipo pregunta-respuesta o preguntas frecuentes de su índice. Los pasajes son extractos de texto que se pueden extraer semánticamente de varios documentos y de varias partes del mismo documento. Si, en casos extremos, sus documentos no devuelven ningún pasaje mediante la API `Retrieve`, también puede utilizar la API `Query` y sus tipos de respuestas.

También puede utilizar la API `Retrieve` para hacer lo siguiente:

- Anular la priorización en el índice
- Filtrar en función de los campos o atributos del documento
- Filtrar en función del acceso del usuario o su grupo a los documentos
- Consultar el bucket de puntuación de confianza para el resultado de un pasaje recuperado. El bucket de confianza proporciona una clasificación relativa que indica el grado de confianza de Amazon Kendra en que la respuesta es relevante para la consulta.

> **Note**
>
> Por el momento, los buckets de puntuación de confianza solo están disponibles en inglés.

También puede incluir algunos campos en la respuesta que podrían proporcionar información adicional útil.

Actualmente, la API `Retrieve` no admite todas las características que admite `Query`. No se admiten las siguientes características: las consultas con una sintaxis de consulta avanzada, las sugerencias de correcciones ortográficas para las consultas, la creación de facetas, las sugerencias de consultas para completar automáticamente las consultas de búsqueda y el aprendizaje incremental. Tenga en cuenta que no todas las funciones se aplican a la API. `Retrieve` Todas las versiones futuras de la `Retrieve` API se documentarán en esta guía.

La API `Retrieve` comparte el número de unidades de capacidad de consulta que establezca para su índice. Para obtener información sobre lo que incluye una unidad de capacidad única y la capacidad base por defecto de un índice, consulte Ajuste de la capacidad.

> **Note**
>
> No se puede añadir capacidad si se utiliza la Amazon Kendra Developer Edition; solo se puede añadir capacidad cuando se utiliza la Amazon Kendra Enterprise Edition. Para obtener más información sobre lo que se incluye en las ediciones Developer y Enterprise, consulte Ediciones de Amazon Kendra.

A continuación se muestra un ejemplo del uso de la API `Retrieve` para recuperar los 100 pasajes más relevantes de los documentos de un índice para la consulta "how does amazon kendra work?".

Python

```
import boto3
import pprint

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
```

```
# Provide the query text
query = "how does amazon kendra work?"
# You can retrieve up to 100 relevant passages
# You can paginate 100 passages across 10 pages, for example
page_size = 10
page_number = 10

result = kendra.retrieve(
        IndexId = index_id,
        QueryText = query,
        PageSize = page_size,
        PageNumber = page_number)

print("\nRetrieved passage results for query: " + query + "\n")

for retrieve_result in result["ResultItems"]:

    print("-------------------")
    print("Title: " + str(retrieve_result["DocumentTitle"]))
    print("URI: " + str(retrieve_result["DocumentURI"]))
    print("Passage content: " + str(retrieve_result["Content"]))
    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.RetrieveRequest;
import software.amazon.awssdk.services.kendra.model.RetrieveResult;
import software.amazon.awssdk.services.kendra.model.RetrieveResultItem;

public class RetrievePassageExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indxId = "index-id";
        String query = "how does amazon kendra work?";
        Integer pgSize = 10;
        Integer pgNumber = 10;

        RetrieveRequest retrieveRequest = retrieveRequest
            .builder()
```

```
            .indexId(indxId)
            .queryText(query)
            .pageSize(pgSize)
            .pageNumber(pgNumber)
            .build();

        RetrieveResult retrieveResult = kendra.retrieve(retrieveRequest);

        System.out.println(String.format("\nRetrieved passage results for query:
 %s", query));
        for(RetrieveResultItem item: retrieveResult.resultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Title: %s", documentTitle));
            System.out.println(String.format("URI: %s", documentURI));
            System.out.println(String.format("Passage content: %s", content));
            System.out.println("----------------------\n");
        }
    }
}
```

Navegar por un índice

Puede examinar los documentos por sus atributos o facetas sin tener que escribir una consulta de búsqueda. Amazon Kendra Navegar por el índice puede ayudar a sus usuarios a descubrir documentos al navegar libremente por un índice sin tener en mente una consulta específica. Esto también ayuda a los usuarios a navegar ampliamente por un índice como punto de partida en su búsqueda.

La navegación por un índice solo se puede utilizar para buscar por atributo o faceta del documento con un tipo de ordenación. No puede buscar en un índice completo mediante la navegación por el índice. Si falta el texto de la consulta, Amazon Kendra solicita un filtro de atributos del documento o una faceta y un tipo de ordenación.

Para permitir la navegación por índices mediante la API de consultas, debe incluir AttributeFiltero Facet y. SortingConfiguration Para permitir la navegación por el índice en la consola, seleccione su índice en Índices en el menú de navegación y, a continuación, seleccione la opción de buscar en su índice. En el cuadro de búsqueda, presione la tecla Intro dos veces. Seleccione el menú desplegable Filtrar los resultados de búsqueda para elegir un filtro y seleccione el menú desplegable Ordenar para elegir un tipo de ordenación.

El siguiente ejemplo muestra cómo navegar por un índice para buscar documentos en español en orden descendente según la fecha de creación del documento.

CLI

```
aws kendra query \
--index-id "index-id" \
--attribute-filter '{
    "EqualsTo":{
      "Key": "_language_code",
      "Value": {
        "StringValue": "es"
      }
    }
  }' \
--sorting-configuration '{
    "DocumentAttributeKey": "_created_at",
    "SortOrder": "DESC"
  }'
```

Python

```
import boto3

kendra = boto3.client("kendra")

# Must include the index ID, the attribute filter, and sorting configuration
response = kendra.query(
        IndexId = "index-id",
        AttributeFilter = {
            "EqualsTo": {
                "Key": "_language_code",
                "Value": {
                    "StringValue": "es"
                    }
                }
            },
        SortingConfiguration = {
            "DocumentAttributeKey": "_created_at",
            "SortOrder": "DESC"})

print("\nSearch results|Resultados de la búsqueda: \n")
```

```
for query_result in response["ResultItems"]:

    print("-------------------")
    print("Type: " + str(query_result["Type"]))

    if query_result["Type"]=="ANSWER" or query_result["Type"]=="QUESTION_ANSWER":
        answer_text = query_result["DocumentExcerpt"]["Text"]
        print(answer_text)

    if query_result["Type"]=="DOCUMENT":
        if "DocumentTitle" in query_result:
            document_title = query_result["DocumentTitle"]["Text"]
            print("Title: " + document_title)
        document_text = query_result["DocumentExcerpt"]["Text"]
        print(document_text)

    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResult;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();
        QueryRequest queryRequest = QueryRequest.builder()
            .withIndexId("index-id")
            .withAttributeFilter(AttributeFilter.builder()
                .withEqualsTo(DocumentAttribute.builder()
                    .withKey("_language_code")
                    .withValue(DocumentAttributeValue.builder()
                        .withStringValue("es")
                        .build())
                    .build())
                .build())
            .withSortingConfiguration(SortingConfiguration.builder()
                .withDocumentAttributeKey("_created_at")
                .withSortOrder("DESC")
```

```
                .build())
            .build());

        QueryResult queryResult = kendra.query(queryRequest);
        for (QueryResultItem item : queryResult.getResultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Type: %s", item.getType()));

            switch (item.getType()) {
                case QueryResultType.QUESTION_ANSWER:
                case QueryResultType.ANSWER:
                    String answerText = item.getDocumentExcerpt().getText();
                    System.out.println(answerText);
                    break;
                case QueryResultType.DOCUMENT:
                    String documentTitle = item.getDocumentTitle().getText();
                    System.out.println(String.format("Title: %s", documentTitle));
                    String documentExcerpt = item.getDocumentExcerpt().getText();
                    System.out.println(String.format("Excerpt: %s",
 documentExcerpt));
                    break;
                default:
                    System.out.println(String.format("Unknown query result type:
 %s", item.getType()));
            }
            System.out.println("----------------------\n");
        }
    }
}
```

Destacar resultados de búsqueda

Puede destacar determinados documentos en los resultados de búsqueda cuando sus usuarios realicen determinadas consultas. Esto ayuda a que los resultados sean más visibles y destacados para los usuarios. Los resultados destacados se separan de la lista habitual de resultados y se muestran en la parte superior de la página de búsqueda. Puede probar destacando diferentes documentos para distintas consultas o asegurarse de que determinados documentos tengan la visibilidad que se merecen.

Debe asignar consultas específicas a documentos específicos para destacarlos en los resultados. Si una consulta contiene una coincidencia exacta, en los resultados de la búsqueda se destacan uno o más documentos específicos.

Por ejemplo, si los usuarios escriben la consulta “nuevos productos de 2023”, puede seleccionar los documentos titulados “Novedades” y “Próximamente” para que se destaquen en la parte superior de la página de resultados de búsqueda. Esto ayuda a garantizar que estos documentos sobre los nuevos productos tengan la visibilidad que se merecen.

Amazon Kendra no duplica los resultados de la búsqueda si un resultado ya está seleccionado para que aparezca en la parte superior de la página de resultados de la búsqueda. Un resultado destacado no vuelve a clasificarse como primer resultado si ya aparece destacado por encima de todos los demás resultados.

Para destacar determinados resultados, debe especificar una coincidencia exacta de una consulta de texto completo, no una coincidencia parcial de una consulta que utilice una palabra clave o frase incluida en una consulta. Por ejemplo, si solo especifica la consulta “Kendra” en un conjunto de resultados destacados, consultas como “¿Cómo clasifica Kendra semánticamente los resultados?” no mostrará los resultados destacados. Los resultados destacados están diseñados para consultas específicas, en lugar de consultas con un alcance demasiado amplio. Amazon Kendra gestiona de forma natural las consultas de tipos de palabras clave para clasificar los documentos más útiles en los resultados de búsqueda, evitando que los resultados aparezcan excesivamente basados en palabras clave simples.

Si hay determinadas consultas que sus usuarios utilizan con frecuencia, puede especificarlas para los resultados destacados. Por ejemplo, si analiza sus consultas principales con Amazon Kendra Analytics y encuentra consultas específicas, como “¿Cómo clasifica Kendra los resultados semánticamente?” y la «búsqueda semántica de kendra» se utilizan con frecuencia, por lo que puede ser útil especificar estas consultas para incluir el documento titulado «búsqueda 101».Amazon Kendra

Amazon Kendra trata las consultas de resultados destacados sin distinguir entre mayúsculas y minúsculas. Amazon Kendra convierte una consulta a minúsculas y reemplaza los espacios en blanco finales por un solo espacio. Amazon Kendra coincide con todos los demás caracteres tal y como aparecen al especificar las consultas para los resultados destacados.

Se crea un conjunto de resultados destacados que se asignan a determinadas consultas mediante la CreateFeaturedResultsSetAPI. Si utiliza la consola, seleccione su índice y, a continuación, seleccione Resultados destacados en el menú de navegación para crear un conjunto de resultados

destacados. Puede crear hasta 50 conjuntos de resultados destacados por índice, destacar hasta cuatro documentos por conjunto y hasta 49 textos de consulta por conjunto de resultados destacados. Puede solicitar un aumento de estos límites poniéndose en contacto con Soporte.

Puede seleccionar el mismo documento en varios conjuntos de resultados destacados. Sin embargo, no debe utilizar el mismo texto de consulta de coincidencia exacta en varios conjuntos. Las consultas que especifique para los resultados destacados deben ser únicas para cada conjunto de resultados destacados de cada índice.

Puede organizar el orden de los documentos al seleccionar hasta cuatro documentos destacados. Si utiliza la API, el orden en el que publica los documentos destacados es el mismo que se muestra en los resultados destacados. Si utiliza la consola, solo tiene que arrastrar y soltar el orden de los documentos al seleccionar los documentos para destacarlos en los resultados.

El control de acceso, según el cual algunos usuarios y grupos tienen acceso a determinados documentos y otros no, se sigue respetando a la hora de configurar los resultados destacados. Esto también es válido para el filtrado por contexto de usuario. Por ejemplo, el usuario A pertenece al grupo de la empresa “Becarios”, que no debería acceder a los documentos que contienen secretos de empresa. Si el usuario A introduce una consulta que hace destacar un documento secreto de la empresa, el usuario A no verá este documento destacado en sus resultados. Esto también se aplica a cualquier otro resultado de la página de resultados de búsqueda. También puede usar etiquetas para controlar el acceso a un conjunto de resultados destacados, que es un recurso de Amazon Kendra para el cual tiene el control del acceso.

A continuación se muestra un ejemplo de cómo crear un conjunto de resultados destacados con las consultas “new products 2023” y “new products available” asignadas a los documentos titulados “What's new” (doc-id-1) and “Coming soon” (doc-id-2).

CLI

```
aws kendra create-featured-results-set \
 --featured-results-set-name 'New product docs to feature' \
 --description "Featuring What's new and Coming soon docs" \
 --index-id index-id \
 --query-texts 'new products 2023' 'new products available' \
 --featured-documents '{"Id":"doc-id-1", "Id":"doc-id-2"}'
```

Python

```
import boto3
```

```
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create a featured results set.")

# Provide a name for the featured results set
featured_results_name = "New product docs to feature"
# Provide an optional decription for the featured results set
description = "Featuring What's new and Coming soon docs"
# Provide the index ID for the featured results set
index = "index-id"
# Provide a list of query texts for the featured results set
queries = ['new products 2023', 'new products available']
# Provide a list of document IDs for the featured results set
featured_doc_ids = [{"Id":"doc-id-1"}, {"Id":"doc-id-2"}]

try:
    featured_results_set_response = kendra.create_featured_results_set(
        FeaturedResultsSetName = featured_results_name,
        Decription = description,
        Index = index,
        QueryTexts = queries,
        FeaturedDocuments = featured_doc_ids
    )

    pprint.pprint(featured_results_set_response)

    featured_results_set_id = featured_results_set_response["FeaturedResultsSetId"]

    while True:
        # Get the details of the featured results set, such as the status
        featured_results_set_description = kendra.describe_featured_results_set(
            Id = featured_results_set_id
        )
        status = featured_results_set_description["Status"]
        print(" Featured results set status: "+status)

except  ClientError as e:
        print("%s" % e)
```

```
print("Program ends.")
```

Búsqueda tabular de HTML

La característica de búsqueda tabular de Amazon Kendra permite buscar y extraer respuestas de tablas incrustadas en documentos HTML. Cuando busques en tu índice, Amazon Kendra incluye un extracto de una tabla si es relevante para la consulta y proporciona información útil.

Amazon Kendra examina toda la información del cuerpo del texto de un documento, incluida la información útil de las tablas. Por ejemplo, un índice contiene informes empresariales con tablas sobre los costes de operación, los ingresos y otra información financiera. Para la consulta, «¿cuál es el costo operativo anual de 2020 a 2022?» , Amazon Kendra puede devolver un extracto de una tabla que contenga las columnas pertinentes de la tabla «Operaciones (millones de USD)» y «Ejercicio financiero», y filas de la tabla que contengan los valores de ingresos de 2020, 2021 y 2022. El extracto de la tabla se incluye en el resultado, junto con el título del documento, un enlace al documento completo y cualquier otro campo del documento que desee incluir.

Los extractos de tablas se pueden mostrar en los resultados de búsqueda tanto si la información se encuentra en una celda de la tabla como en varias celdas. Por ejemplo, Amazon Kendra puede mostrar un extracto de una tabla adaptado a cada uno de estos tipos de consultas:

- “tarjeta de crédito con la tasa de interés más alta de 2020”
- “tarjeta de crédito con la tasa de interés más alta de 2020 a 2022”
- “las 3 tarjetas de crédito con la tasa de interés más alta de 2020 a 2022”
- “tarjetas de crédito con tasas de interés inferiores al 10 %”
- “todas las tarjetas de crédito con intereses bajos disponibles”

Amazon Kendra resalta la celda o celdas de la tabla que son más relevantes para la consulta. En el resultado de búsqueda se muestran las celdas más relevantes con sus filas, columnas y nombres de columnas correspondientes. El extracto de la tabla muestra hasta cinco columnas y tres filas, en función del número de celdas de la tabla que sean relevantes para la consulta y del número de columnas disponibles en la tabla original. La celda más relevante se muestra en el extracto de la tabla, junto con las siguientes celdas más relevantes.

La respuesta incluye el bucket de confianza (`MEDIUM`, `HIGH` y `VERY_HIGH`) para mostrar la relevancia de la respuesta de la tabla para la consulta. Si el valor de una celda de la tabla tiene

una confianza VERY_HIGH, se convierte en la “respuesta principal” y se resalta. En el caso de los valores de las celdas de la tabla con una confianza HIGH, aparecen resaltados. En el caso de los valores de las celdas de la tabla con una confianza MEDIUM, no aparecen resaltados. La confianza general para la respuesta de la tabla se devuelve en la respuesta. Por ejemplo, si una tabla contiene principalmente celdas de la tabla con confianza HIGH, la confianza general que se devuelve en la respuesta de la tabla es de confianza HIGH.

De forma predeterminada, a las tablas no se les asigna un mayor nivel de importancia ni más peso que a otros componentes de un documento. Dentro de un documento, si una tabla es solo ligeramente relevante para una consulta, pero hay un párrafo muy relevante, Amazon Kendra devuelve un extracto del párrafo. Los resultados de búsqueda muestran el contenido que proporciona la mejor respuesta posible y la información más útil, en el mismo documento o en otros documentos. Si la confianza de una tabla es inferior a MEDIUM, el extracto de la tabla no se devuelve en la respuesta.

Para utilizar la búsqueda tabular en un índice existente, debe volver a indexar el contenido.

Amazon Kendra La búsqueda tabular admite sinónimos (incluidos los personalizados). Amazon Kendra solo admite documentos en inglés con tablas HTML que estén dentro de la etiqueta de tabla.

El siguiente ejemplo muestra un extracto de tabla incluido en el resultado de una consulta. Para ver un ejemplo de JSON con respuestas a consultas, incluidos extractos de tablas, consulte Respuestas a las consultas y tipos de respuestas.

Python

```
import boto3
import pprint

kendra = boto3.client("kendra")

# Provide the index ID
index_id = <index-id>
# Provide the query text
query = "search string"

response = kendra.query(
        QueryText = query,
        IndexId = index_id)

print("\nSearch results for query: " + query + "\n")
```

```
for query_result in response["ResultItems"]:

    print("-------------------")
    print("Type: " + str(query_result["Type"]))
    print("Type: " + str(query_result["Format"]))

    if query_result["Type"]=="ANSWER" and query_result["Format"]=="TABLE":
        answer_table = query_result["TableExcerpt"]
        print(answer_table)

    if query_result["Type"]=="ANSWER" and query_result["Format"]=="TEXT":
        answer_text = query_result["DocumentExcerpt"]
        print(answer_text)

    if query_result["Type"]=="QUESTION_ANSWER":
        question_answer_text = query_result["DocumentExcerpt"]["Text"]
        print(question_answer_text)

    if query_result["Type"]=="DOCUMENT":
        if "DocumentTitle" in query_result:
            document_title = query_result["DocumentTitle"]["Text"]
            print("Title: " + document_title)
        document_text = query_result["DocumentExcerpt"]["Text"]
        print(document_text)

    print("------------------\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResponse;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String query = "search string";
        String indexId = "index-id";
```

```
        QueryRequest queryRequest = QueryRequest
            .builder()
            .queryText(query)
            .indexId(indexId)
            .build();

        QueryResponse queryResponse = kendra.query(queryRequest);

        System.out.println(String.format("\nSearch results for query: %s", query));
        for(QueryResultItem item: queryResponse.resultItems()) {
            System.out.println("----------------------");
            System.out.println(String.format("Type: %s", item.type()));
            System.out.println(String.format("Format: %s", item.format()));

            switch(item.format()) {
                case TABLE:
                    String answerTable = item.TableExcerpt();
                    System.out.println(answerTable);
                    break;
            }

            switch(item.format()) {
                case TEXT:
                    String answerText = item.DocumentExcerpt();
                    System.out.println(answerText);
                    break;
            }

            switch(item.type()) {
                case QUESTION_ANSWER:
                    String questionAnswerText = item.documentExcerpt().text();
                    System.out.println(questionAnswerText);
                    break;
                case DOCUMENT:
                    String documentTitle = item.documentTitle().text();
                    System.out.println(String.format("Title: %s", documentTitle));
                    String documentExcerpt = item.documentExcerpt().text();
                    System.out.println(String.format("Excerpt: %s",
documentExcerpt));
                    break;
                default:
                    System.out.println(String.format("Unknown query result type:
%s", item.type()));
```

```
            }

            System.out.println("----------------------\n");
        }
    }
}
```

Sugerencias de consulta

Las sugerencias de consulta de Amazon Kendra pueden ayudar a los usuarios a escribir sus consultas de búsqueda más rápido y a guiar su búsqueda.

Amazon Kendra sugiere consultas relevantes para los usuarios en función de una de las siguientes opciones:

- Consultas populares en el historial de consultas o en el registro de consultas
- El contenido de los campos o atributos del documento

Puede configurar su preferencia de uso del historial de consultas o los campos del documento configurando `SuggestionTypes` como `QUERY` o `DOCUMENT_ATTRIBUTES` y llamando a GetQuerySuggestions. De forma predeterminada, Amazon Kendra utiliza el historial de consultas para basar las sugerencias. Si tanto el historial de consultas como los campos del documento están activados cuando llamas UpdateQuerySuggestionsConfigy no has establecido tu `SuggestionTypes` preferencia para usar los campos del documento, Amazon Kendra utiliza el historial de consultas.

Si utiliza la consola, puede basar las sugerencias de consulta en el historial de consultas o en los campos del documento. Primero debe seleccionar su índice y, a continuación, seleccionar Sugerencias de consulta en la sección Enriquecimientos del menú de navegación. A continuación, seleccione Configurar sugerencias de consulta. Tras configurar las sugerencias de consulta, accederá a una consola de búsqueda en la que podrá seleccionar el Historial de consultas o los Campos del documento en el panel derecho e introducir una consulta de búsqueda en la barra de búsqueda.

De forma predeterminada, las sugerencias de consulta que utilizan el historial de consultas y los campos del documento se activan sin coste adicional. Puede desactivar este tipo de sugerencias de consulta en cualquier momento mediante la API `UpdateQuerySuggestionsConfig`. Para

desactivar las sugerencias de consulta basadas en el historial de consultas, configure `Mode` en `DISABLED` al llamar a `UpdateQuerySuggestionsConfig`. Para desactivar las sugerencias de consulta basadas en los campos del documento, configure `AttributeSuggestionsMode` en `INACTIVE` en la configuración de los campos del documento y, a continuación, llame a `UpdateQuerySuggestionsConfig>`. Si usa la consola, puede desactivar las sugerencias de consulta en la Configuración de las sugerencias de consulta.

Las sugerencias de consulta no distinguen mayúsculas de minúsculas. Amazon Kendra convierte el prefijo de la consulta y la consulta sugerida a minúsculas, omite todas las comillas simples y dobles y reemplaza varios espacios en blanco por un solo espacio. Amazon Kendra coincide con todos los demás caracteres especiales tal como están. Amazon Kendra no muestra ninguna sugerencia si un usuario escribe menos de dos caracteres o más de 60 caracteres.

Temas

- Sugerencias de consulta mediante el historial de consultas
- Sugerencias de consulta mediante los campos del documento
- Bloquear determinadas consultas o contenidos de los campos del documento para que no usen en las sugerencias

Sugerencias de consulta mediante el historial de consultas

Temas

- Configuración para seleccionar consultas para sugerencias
- Borrar las sugerencias pero conservar el historial de consultas
- No hay sugerencias disponibles

Puede optar por sugerir consultas relevantes para sus usuarios en función de las consultas más frecuentes del historial de consultas o del registro de consultas. Amazon Kendra utiliza todas las consultas que buscan los usuarios y aprende de ellas para hacer sugerencias a los usuarios. Amazon Kendra sugiere consultas populares a los usuarios cuando empiezan a escribirlas. Amazon Kendra sugiere una consulta si el prefijo o los primeros caracteres de la consulta coinciden con lo que el usuario empieza a escribir en su consulta.

Por ejemplo, un usuario empieza a escribir la consulta “próximos eventos”. Amazon Kendra ha aprendido del historial de consultas que muchos usuarios han buscado “próximos eventos de 2050” muchas veces. El usuario ve aparecer “próximos eventos de 2050” directamente debajo de la barra

de búsqueda, y completa automáticamente su consulta de búsqueda. El usuario selecciona esta sugerencia de consulta y en los resultados de búsqueda aparece el documento “Nuevos eventos: qué pasará en 2050”.

Puede especificar cómo Amazon Kendra selecciona las consultas aptas para sugerirlas a sus usuarios. Por ejemplo, puede especificar que una sugerencia de consulta debe haber sido buscada por al menos 10 usuarios únicos (el valor predeterminado es tres), debe haber sido buscada en los últimos 30 días y no debe contener palabras o frases de su lista de bloqueados. Amazon Kendra requiere que la consulta tenga al menos un resultado de búsqueda y que contenga al menos una palabra de más de cuatro caracteres.

Configuración para seleccionar consultas para sugerencias

Puede configurar los siguientes ajustes para seleccionar consultas para sugerencias mediante la API UpdateQuerySuggestionsConfig:

- Modo: las sugerencias de consulta que utilizan el historial de consultas están `ENABLED` o son `LEARN_ONLY`. Amazon Kendra activa las sugerencias de consulta de forma predeterminada. `LEARN_ONLY` desactiva las sugerencias de consulta. Si está desactivada, Amazon Kendra sigue aprendiendo las sugerencias, pero no hace sugerencias de consulta a los usuarios.
- Periodo de tiempo del registro de consultas: la antigüedad de las consultas en el periodo de tiempo del registro de consultas. El periodo de tiempo es un valor entero para el número de días desde el día actual hasta los días anteriores.
- Consultas sin información del usuario: debe configurarlo como `TRUE` para incluir todas las consultas, o como `FALSE` para incluir solo las consultas con información del usuario. Puede usar esta configuración si la aplicación de búsqueda incluye información del usuario, como el ID de usuario, cuando un usuario realiza una consulta. De forma predeterminada, esta configuración no filtra las consultas si no hay información del usuario específica asociada a las consultas. Sin embargo, puede usar esta configuración para hacer sugerencias basadas únicamente en las consultas que incluyan información del usuario.
- Usuarios únicos: el número mínimo de usuarios únicos que deben haber buscado una consulta para que sea apta para sugerirla a sus usuarios. Este número es un valor entero.
- Recuento de consultas: el número mínimo de veces que se debe haber buscado una consulta para que sea apta para sugerirla a sus usuarios. Este número es un valor entero.

Estos ajustes afectan a la forma en que se seleccionan las consultas como consultas populares para sugerirlas a sus usuarios. La forma en que debe ajustar la configuración depende de sus necesidades específicas, por ejemplo:

- Si sus usuarios suelen buscar una vez al mes de media, puede establecer el número de días en el periodo de tiempo del registro de consultas en 30 días. Al usar esa configuración, capturará la mayoría de las consultas recientes de sus usuarios antes de que queden desactualizadas en el periodo de tiempo.
- Si solo un número reducido de consultas incluye información del usuario y no desea sugerir consultas basadas en un tamaño de muestra pequeño, puede configurar las consultas para que incluyan a todos los usuarios.
- Si define las consultas populares como las que han buscado al menos 10 usuarios únicos y las que se han buscado al menos 100 veces, debe establecer los usuarios únicos en 10 y el recuento de consultas en 100.

> **Warning**
>
> Es posible que los cambios en la configuración no se apliquen de forma inmediata. Puede realizar un seguimiento de los cambios en la configuración mediante la API DescribeQuerySuggestionsConfig. El tiempo que tarda en surtir efecto la configuración actualizada depende de las actualizaciones que realice y del número de consultas de búsqueda de su índice. Amazon Kendra actualiza automáticamente las sugerencias cada 24 horas, después de cambiar una configuración o después de aplicar una lista de bloqueo.

CLI

Para recuperar las sugerencias de consulta

```
aws kendra get-query-suggestions \
 --index-id index-id \
 --query-text "query-text" \
 --suggestion-types '["QUERY"]' \
 --max-suggestions-count 1 // If you want to limit the number of suggestions
```

Para actualizar las sugerencias de consulta

Por ejemplo, para cambiar el periodo de tiempo del registro de consultas y el número mínimo de veces que se debe haber buscado una consulta:

```
aws kendra update-query-suggestions-config \
 --index-id index-id \
 --query-log-look-back-window-in-days 30 \
 --minimum-query-count 100
```

Python

Para recuperar las sugerencias de consulta

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Get query suggestions.")

# Provide the index ID
index_id = "index-id"

# Provide the query text
query_text = "query"

# Provide the query suggestions type
query_suggestions_type = "QUERY"


# If you want to limit the number of suggestions
num_suggestions = 1

try:
    query_suggestions_response = kendra.get_query_suggestions(
        IndexId = index_id,
        QueryText = query_text,
        SuggestionTypes = query_suggestions_type,
        MaxSuggestionsCount = num_suggestions
    )

    # Print out the suggestions you received
    if ("Suggestions" in query_suggestions_response.keys()) {
        for (suggestion: query_suggestions_response["Suggestions"]) {
```

```
            print(suggestion["Value"]["Text"]["Text"]);
        }
    }

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Para actualizar las sugerencias de consulta

Por ejemplo, para cambiar el periodo de tiempo del registro de consultas y el número mínimo de veces que se debe haber buscado una consulta:

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Updating query suggestions settings/configuration for an index.")

# Provide the index ID
index_id = "index-id"

# Configure the settings you want to update
minimum_query_count = 100
query_log_look_back_window_in_days = 30

try:
    kendra.update_query_suggestions_config(
        IndexId = index_id,
        MinimumQueryCount = minimum_query_count,
        QueryLogLookBackWindowInDays = query_log_look_back_window_in_days
    )

    print("Wait for Amazon Kendra to update the query suggestions.")

    while True:
        # Get query suggestions description of settings/configuration
        query_sugg_config_response = kendra.describe_query_suggestions_config(
            IndexId = index_id
```

```
        )

        # If status is not UPDATING, then quit
        status = query_sugg_config_response["Status"]
        print(" Updating query suggestions config. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Borrar las sugerencias pero conservar el historial de consultas

Puede borrar las sugerencias de consultas mediante la API ClearQuerySuggestions. Al borrar las sugerencias, solo se eliminan las sugerencias de consulta existentes, no las consultas del historial de consultas. Al borrar las sugerencias, Amazon Kendra aprende las nuevas en función de las consultas nuevas que se hayan agregado al registro de consultas desde el momento en que las eliminaste.

CLI

Para borrar las sugerencias de consulta

```
aws  kendra clear-query-suggestions \
 --index-id index-id
```

Python

Para borrar las sugerencias de consulta

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Clearing out query suggestions for an index.")

# Provide the index ID
index_id = "index-id"
```

```
try:
    kendra.clear_query_suggestions(
        IndexId = index_id
    )

    # Confirm last cleared date-time and that there are no suggestions
    query_sugg_config_response = kendra.describe_query_suggestions_config(
        IndexId = index_id
    )
    print("Query Suggestions last cleared at: " +
 str(query_sugg_config_response["LastClearTime"]));
    print("Number of suggestions available from the time of clearing: " +
 str(query_sugg_config_response["TotalSuggestionsCount"]));

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

No hay sugerencias disponibles

Si no ve sugerencias para una consulta, puede deberse a uno de los siguientes motivos:

- No hay suficientes consultas en el índice de las Amazon Kendra que pueda aprender.
- La configuración de las sugerencias de consulta es demasiado estricta, por lo que la mayoría de las consultas se excluyen de las sugerencias.
- Ha aprobado las sugerencias recientemente y Amazon Kendra todavía necesita tiempo para acumular nuevas consultas y obtener nuevas sugerencias.

Puede comprobar la configuración actual mediante la API DescribeQuerySuggestionsConfig.

Sugerencias de consulta mediante los campos del documento

Temas

- Configuración para seleccionar campos para sugerencias
- Control de usuarios en los campos del documento

Puede optar por sugerir consultas relevantes para sus usuarios basadas en el contenido de los campos del documento. En lugar de utilizar el historial de consultas para sugerir otras consultas relevantes y populares, puede utilizar la información contenida en un campo de documento que sea útil para completar automáticamente la consulta. Amazon Kendra busca contenido relevante en los campos configurados como la consulta del usuario `Suggestable` y que se alinee estrechamente con ella. A continuación, Amazon Kendra sugiere este contenido al usuario cuando empiece a escribir la consulta.

Por ejemplo, si especificas el campo de título en el que basar las sugerencias y un usuario empieza a escribir la consulta «Cómo Amazon Ken... ', se podría sugerir el título más relevante, «Cómo Amazon Kendra funciona», para completar automáticamente la búsqueda. El usuario ve que aparece «Cómo Amazon Kendra funciona» directamente debajo de la barra de búsqueda, lo que completa automáticamente su consulta de búsqueda. El usuario selecciona esta sugerencia de consulta y aparece el documento «Cómo Amazon Kendra funciona» en los resultados de la búsqueda.

Puede utilizar el contenido de cualquier campo del documento de tipo `String` y `StringList` para sugerir una consulta configurando el campo como `Suggestable` como parte de la configuración de campos para las sugerencias de consulta. También puede utilizar una lista de bloqueo para que los usuarios no vean los campos de documentos sugeridos que contienen determinadas palabras o frases. Puede usar una única lista de bloqueo. La lista de bloqueo se aplica si configura las sugerencias de consulta para que utilicen tanto el historial de consultas como los campos del documento.

Configuración para seleccionar campos para sugerencias

Puede configurar los siguientes ajustes para seleccionar los campos del documento para las sugerencias, utilizando AttributeSuggestionsConfig y llamando a la API UpdateQuerySuggestionsConfig para actualizar los ajustes en el índice:

- Modo de sugerencias de campos o atributos: las sugerencias de consulta que utilizan los campos del documento pueden estar `ACTIVE` o `INACTIVE`. Amazon Kendra activa las sugerencias de consulta de forma predeterminada.
- Campos o atributos que se pueden sugerir: los nombres de los campos o las claves de campo en los que basar las sugerencias. Estos campos deben estar configurados como `TRUE` para `Suggestable`, como parte de la configuración de los campos. Puede anular la configuración de los campos en la consulta y, al mismo tiempo, mantener la configuración en el índice. Utilice la GetQuerySuggestionsAPI para realizar cambios `AttributeSuggestionConfig` en el nivel de

consulta. Esta configuración de la consulta puede resultar útil para probar rápidamente el uso de diferentes campos del documento sin tener que actualizar la configuración a nivel de índice.

- Campos o atributos adicionales: los campos adicionales que desea incluir en la respuesta para una sugerencia de consulta. Estos campos se utilizan para proporcionar información adicional en la respuesta; sin embargo, no se utilizan como base para las sugerencias.

> **Warning**
>
> Es posible que los cambios en la configuración no se apliquen de forma inmediata. Puede realizar un seguimiento de los cambios en la configuración mediante la API DescribeQuerySuggestionsConfig. El tiempo que tarda en surtir efecto la configuración actualizada depende de las actualizaciones que realices. Amazon Kendra actualiza automáticamente las sugerencias cada 24 horas, después de cambiar una configuración o después de aplicar una lista de bloqueados.

CLI

Para recuperar las sugerencias de consulta y anular la configuración de los campos del documento en la consulta en lugar de tener que cambiar la configuración en el índice.

```
aws kendra get-query-suggestions \
 --index-id index-id \
 --query-text "query-text" \
 --suggestion-types '["DOCUMENT_ATTRIBUTES"]' \
 --attribute-suggestions-config '{"SuggestionAttributes":'["field/attribute key
 1", "field/attribute key 2"]', "AdditionalResponseAttributes":'["response field/
attribute key 1", "response field/attribute key 2"]'}' \
 --max-suggestions-count 1 // If you want to limit the number of suggestions
```

Para actualizar las sugerencias de consulta

Por ejemplo, para cambiar la configuración de los campos del documento en el índice:

```
aws kendra update-query-suggestions-config \
 --index-id index-id \
 --attribute-suggestions-config '{"SuggestableConfigList": '[{"SuggestableConfig":
 "_document_title", "Suggestable": true}]', "AttributeSuggestionsMode": "ACTIVE"}'
```

Python

Para recuperar las sugerencias de consulta y anular la configuración de los campos del documento en la consulta en lugar de tener que cambiar la configuración en el índice.

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Get query suggestions.")

# Provide the index ID
index_id = "index-id"

# Provide the query text
query_text = "query"

# Provide the query suggestions type
query_suggestions_type = "DOCUMENT_ATTRIBUTES"

# Override fields/attributes configuration at query level
configuration = {"SuggestionAttributes":
    '["field/attribute key 1", "field/attribute key 2"]',
      "AdditionalResponseAttributes":
          '["response field/attribute key 1", "response field/attribute key 2"]'
          }

# If you want to limit the number of suggestions
num_suggestions = 1

try:
    query_suggestions_response = kendra.get_query_suggestions(
        IndexId = index_id,
        QueryText = query_text,
        SuggestionTypes = [query_suggestions_type],
        AttributeSuggestionsConfig = configuration,
        MaxSuggestionsCount = num_suggestions
    )

    # Print out the suggestions you received
    if ("Suggestions" in query_suggestions_response.keys()) {
        for (suggestion: query_suggestions_response["Suggestions"]) {
            print(suggestion["Value"]["Text"]["Text"]);
```

```
        }
    }

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Para actualizar las sugerencias de consulta

Por ejemplo, para cambiar la configuración de los campos del documento en el índice:

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Updating query suggestions settings/configuration for an index.")

# Provide the index ID
index_id = "index-id"

# Configure the settings you want to update at the index level
configuration = {"SuggestableConfigList":
    '[{"SuggestableConfig": "_document_title", "Suggestable": true}]',
       "AttributeSuggestionsMode": "ACTIVE"
       }

try:
    kendra.update_query_suggestions_config(
        IndexId = index_id,
        AttributeSuggestionsConfig = configuration
    )

    print("Wait for Amazon Kendra to update the query suggestions.")

    while True:
        # Get query suggestions description of settings/configuration
        query_sugg_config_response = kendra.describe_query_suggestions_config(
            IndexId = index_id
        )
```

```
        # If status is not UPDATING, then quit
        status = query_sugg_config_response["Status"]
        print(" Updating query suggestions config. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Control de usuarios en los campos del documento

Puede aplicar un filtrado por contexto de usuario a los campos del documento en los que desee basar las sugerencias de consulta. De este modo se filtra la información de los campos del documento en función del acceso del usuario o de su grupo a los documentos. Por ejemplo, un becario busca en el portal de la empresa y no tiene acceso a un documento empresarial de alto secreto. Por lo tanto, las consultas sugeridas basadas en el título del documento de alto secreto o en cualquier otro campo que pueda sugerirse no se muestran al becario.

Puede indexar sus documentos con una lista de control de acceso (ACL), que defina qué usuarios y grupos tienen asignado acceso a qué documentos. A continuación, puede aplicar un filtrado por contexto de usuario a los campos de sus documentos para las sugerencias de consulta. El filtrado por contexto de usuario que está configurado actualmente para su índice es el mismo filtrado por contexto de usuario que se aplica a la configuración de los campos del documento para las sugerencias de consultas. El filtrado por contexto de usuario forma parte de la configuración de los campos del documento. Debe usar AttributeSuggestionsGetConfig y llamar a GetQuerySuggestions.

Bloquear determinadas consultas o contenidos de los campos del documento para que no usen en las sugerencias

Una lista de bloqueo Amazon Kendra impide sugerir determinadas consultas a los usuarios. Una lista de bloqueo es una lista de palabras o frases que quieres excluir de las sugerencias de consultas. Amazon Kendra excluye las consultas que contienen una coincidencia exacta de las palabras o frases de la lista de bloqueados.

Puede usar una lista de bloqueo para protegerse de las palabras o frases ofensivas que suelen aparecer en su historial de consultas o en los campos de los documentos y que Amazon Kendra

podría seleccionar como sugerencias. Una lista de bloqueo también puede Amazon Kendra impedir que se sugieran consultas que contengan información que no esté lista para publicarse o anunciarse públicamente. Por ejemplo, pongamos el caso de que sus usuarios consultan con frecuencia sobre el próximo lanzamiento de un posible producto nuevo. Sin embargo, no quiere sugerir el producto porque no está preparado para lanzarlo. Puede bloquear las consultas que contengan el nombre y la información del producto para que no aparezcan en las sugerencias.

Puede crear una lista de bloqueo para las consultas mediante la API CreateQuerySuggestionsBlockList. Para ello debe colocar cada palabra o frase bloqueadas en una línea diferente de un archivo de texto. A continuación, carga el archivo de texto en su bucket de Amazon S3 y proporciona la ruta o ubicación del archivo Amazon S3. Amazon Kendra actualmente solo admite la creación de una lista de bloqueos.

Puedes reemplazar el archivo de texto de las palabras y frases bloqueadas en tu Amazon S3 lista. Para actualizar la lista de bloqueados Amazon Kendra, usa la UpdateQuerySuggestionsBlockListAPI.

Use la API DescribeQuerySuggestionsBlockList para obtener el estado de su lista de bloqueo. `DescribeQuerySuggestionsBlockList` también puede proporcionarle otra información útil, como la siguiente:

- Cuándo se actualizó su lista de bloqueo por última vez
- Cuántas palabras o frases hay en su lista de bloqueo actual
- Mensajes de error útiles al crear una lista de bloqueo

También puede usar la API ListQuerySuggestionsBlockLists para obtener una lista de resúmenes de listas de bloqueo para un índice.

Para eliminar tu lista de bloqueados, usa la DeleteQuerySuggestionsBlockListAPI.

Es posible que las actualizaciones de la lista de bloqueo no surtan efecto de inmediato. Puede realizar un seguimiento de las actualizaciones mediante la API `DescribeQuerySuggestionsBlockList`.

CLI

Para crear una lista de bloqueo

```
aws kendra create-query-suggestions-block-list \
 --index-id index-id \
 --name "block-list-name" \
```

```
 --description "block-list-description" \
 --source-s3-path "Bucket=bucket-name,Key=query-suggestions/block_list.txt" \
 --role-arn role-arn
```

Para actualizar una lista de bloqueo

```
aws kendra update-query-suggestions-block-list \
 --index-id index-id \
 --name "new-block-list-name" \
 --description "new-block-list-description" \
 --source-s3-path "Bucket=bucket-name,Key=query-suggestions/new_block_list.txt" \
 --role-arn role-arn
```

Para eliminar una lista de bloqueo

```
aws kendra delete-query-suggestions-block-list \
 --index-id index-id \
 --id block-list-id
```

Python

Para crear una lista de bloqueo

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create a query suggestions block list.")

# Provide a name for the block list
block_list_name = "block-list-name"
# Provide an optional description for the block list
block_list_description = "block-list-description"
# Provide the IAM role ARN required for query suggestions block lists
block_list_role_arn = "role-arn"

# Provide the index ID
index_id = "index-id"
```

```
s3_bucket_name = "bucket-name"
s3_key = "query-suggestions/block_list.txt"
source_s3_path = {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    block_list_response = kendra.create_query_suggestions_block_list(
        Description = block_list_description,
        Name = block_list_name,
        RoleArn = block_list_role_arn,
        IndexId = index_id,
        SourceS3Path = source_s3_path
    )

    print(block_list_response)

    block_list_id = block_list_response["Id"]

    print("Wait for Amazon Kendra to create the block list.")

    while True:
        # Get block list description
        block_list_description = kendra.describe_query_suggestions_block_list(
            Id = block_list_id,
            IndexId = index_id
        )
        # If status is not CREATING, then quit
        status = block_list_description["Status"]
        print("Creating block list. Status: " + status)
        if status != "CREATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Para actualizar una lista de bloqueo

```
import boto3
```

```
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Update a block list for query suggestions.")

# Provide the block list name you want to update
block_list_name = "new-block-list-name"
# Provide the block list description you want to update
block_list_description = "new-block-list-description"
# Provide the IAM role ARN required for query suggestions block lists
block_list_role_arn = "role-arn"

# Provide the block list ID
block_list_id = "block-list-id"
# Provide the index ID
index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "query-suggestions/new_block_list.txt"
source_s3_path = {
'Bucket': s3_bucket_name,
'Key': s3_key
}

try:
    kendra.update_query_suggestions_block_list(
        Id = block_list_id,
        IndexId = index_id,
        Description = block_list_description,
        Name = block_list_name,
        RoleArn = block_list_role_arn,
        SourceS3Path = source_s3_path
    )

    print("Wait for Amazon Kendra to update the block list.")

    while True:
        # Get block list description
        block_list_description = kendra.describe_query_suggestions_block_list(
            Id = block_list_id,
            IndexId = index_id
```

```
        )
        # If status is not UPDATING, then the update has finished
        status = block_list_description["Status"]
        print("Updating block list. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

except ClientError as e:
print("%s" % e)

print("Program ends.")
```

Para eliminar una lista de bloqueo

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Delete a block list for query suggestions.")

# provide the block list ID
query_suggestions_block_list_id = "query-suggestions-block-list-id"
# Provide the index ID
index_id = "index-id"

try:
    kendra.delete_query_suggestions_block_list(
        Id = query_suggestions_block_list_id,
        IndexId = index_id
    )

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Corrector ortográfico de las consultas

El corrector ortográfico de Amazon Kendra sugiere correcciones ortográficas para las consultas. De este modo, puede ayudarle a reducir al mínimo el número de casos de resultados de búsqueda nulos y a devolver resultados relevantes. Es posible que sus usuarios no reciban ningún resultado de búsqueda debido a consultas mal escritas sin resultados coincidentes o que no devuelvan documentos. También es posible que sus usuarios reciban resultados de búsqueda irrelevantes debido a consultas mal escritas.

El corrector ortográfico está diseñado para sugerir correcciones a las palabras mal escritas en función de las palabras que aparecen en sus documentos indexados y del grado de coincidencia entre una palabra corregida y una palabra mal escrita. Por ejemplo, si la palabra “estados” aparece en sus documentos indexados, podría coincidir estrechamente con la palabra “estatos” mal escrita en la consulta “estatos financieros de fin de año”.

El corrector ortográfico devuelve las palabras deseadas o corregidas que sustituyen a las palabras mal escritas en el texto original de la consulta. Por ejemplo, “impementar la búsqueda de Kendre” podría devolver “implementar la búsqueda de Kendra”. También puede usar las ubicaciones de desplazamiento proporcionadas en la API para resaltar o poner en cursiva las palabras corregidas devueltas en una consulta de su aplicación frontend. En la consola, las palabras corregidas aparecen resaltadas o en cursiva de forma predeterminada. Por ejemplo, “implementar la búsqueda de Kendra”.

En el caso de los términos especializados o específicos de la empresa que aparecen en los documentos indexados, el corrector ortográfico no los malinterpreta como errores ortográficos en la consulta. Por ejemplo, “amazon macie” no se corrige por “amazon nadie”.

En el caso de las palabras separadas con guiones, como “lectura-escritura”, el corrector ortográfico las trata como palabras individuales para sugerir correcciones. Por ejemplo, la corrección sugerida para “letcura-escritura” podría ser “lectura-escritura”.

Para los tipos de respuestas a consultas `DOCUMENT` y `QUESTION_ANSWER`, el corrector ortográfico sugiere correcciones para las palabras mal escritas en función de las palabras del cuerpo del documento. El cuerpo del documento es más fiable que el título a la hora de sugerir correcciones que coincidan estrechamente con las palabras mal escritas. Para los tipos de respuestas a consultas `ANSWER`, el corrector ortográfico sugiere correcciones en función de las palabras del documento de preguntas y respuestas predeterminado del índice.

Puedes activar el corrector ortográfico utilizando el SpellCorrectionConfigurationobjeto. Debe configurar `IncludeQuerySpellCheckSuggestions` como `TRUE`. El corrector ortográfico está activado de forma predeterminada en la consola. Está integrado en la consola de forma predeterminada.

El corrector ortográfico también puede sugerir correcciones ortográficas para consultas en varios idiomas, no solo en inglés. Para ver una lista de los idiomas compatibles con el corrector ortográfico, consulte Idiomas admitidos por Amazon Kendra.

Uso del corrector ortográfico de las consultas con los límites predeterminados

El corrector ortográfico está diseñado con ciertos límites o valores predeterminados. La siguiente lista presenta los límites actuales que se aplican al activar las sugerencias de corrección ortográfica.

- No se pueden devolver las correcciones ortográficas sugeridas para las palabras que tengan menos de tres caracteres o más de 30 caracteres. Para permitir más de 30 caracteres o menos de tres caracteres, póngase en contacto con Soporte.
- Las correcciones ortográficas sugeridas no pueden restringir las sugerencias basadas en el control de acceso de los usuarios o de su lista de control de acceso para el filtrado por contexto de usuario. Las correcciones ortográficas se basan en todas las palabras de los documentos indexados, estén restringidas a determinados usuarios o no. Si quiere evitar que determinadas palabras aparezcan en las correcciones ortográficas sugeridas para las consultas, no active `SpellCorrectionConfiguration`.
- No se pueden devolver correcciones ortográficas sugeridas para palabras que contienen números. Por ejemplo, “Cómo consultar documentos indexa2”.
- Las correcciones ortográficas sugeridas no pueden utilizar palabras que no aparezcan en los documentos indexados.
- Las correcciones ortográficas sugeridas no pueden utilizar palabras con una aparición menor al 0,01 % en los documentos indexados. Para cambiar el umbral del 0,01 %, póngase en contacto con Soporte.

Filtrado y búsqueda por facetas

Puede mejorar los resultados de búsqueda o la respuesta desde la API de consulta mediante filtros. Los filtros restringen los documentos de la respuesta a aquellos que se corresponden directamente

con la consulta. Para crear sugerencias de búsqueda por facetas, use los operadores booleanos para excluir los atributos específicos del documento de la respuesta o los documentos que no cumplan criterios específicos. Puede especificar facetas mediante el parámetro `Facets` de la API `Query`.

Para buscar documentos con Amazon Kendra los que ha indexado Amazon Lex, utilice AMAZON.KendraSearchIntent. Para ver un ejemplo de configuración Amazon Kendra con Amazon Lex, consulte Creación de un bot de preguntas frecuentes para un Amazon Kendra índice. También puede proporcionar un filtro para la respuesta utilizando AttributeFilter. Este es el filtro de consulta en JSON cuando se configura `AMAZON.KendraSearchIntent`. Para proporcionar un filtro de atributos al configurar una intención de búsqueda en la consola, vaya al editor de intenciones y elija la consulta de Amazon Kendra para proporcionar un filtro de consulta en JSON. Para obtener más información acerca de `AMAZON.KendraSearchIntent`, consulte la Guía de documentación de Amazon Lex.

Facetas

Las facetas son vistas limitadas de un conjunto de resultados de búsqueda. Por ejemplo, puede proporcionar resultados de búsqueda para ciudades de todo el mundo, donde los documentos se filtran por una ciudad específica a la que están asociados. O bien, puede crear facetas para mostrar los resultados de un autor específico.

Puede utilizar un atributo del documento o un campo de metadatos asociado a un documento como faceta para que los usuarios puedan buscar por categorías o valores dentro de esa faceta. También puede mostrar facetas anidadas en los resultados de búsqueda para que los usuarios puedan buscar no solo por categoría o campo, sino también por subcategoría o subcampo.

En el siguiente ejemplo se muestra cómo obtener información de facetas para el atributo personalizado “City”.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        Facets = [
            {
                "DocumentAttributeKey" : "City"
            }
        ]
        )
```

Puede utilizar facetas anidadas para restringir aún más la búsqueda. Por ejemplo, el atributo o faceta del documento “City” incluye un valor denominado “Seattle”. Además, el atributo o faceta "CityRegion" del documento incluye los valores «Norte» y «Sur» para los documentos asignados a «Seattle». Puede mostrar las facetas anidadas con sus recuentos en los resultados de búsqueda, de modo que los documentos se puedan buscar no solo por ciudad sino también por región dentro de una ciudad.

Tenga en cuenta que las facetas anidadas pueden afectar a la latencia de las consultas. Por regla general, cuantas más facetas anidadas utilice, mayor será el impacto potencial en la latencia. Otros factores que afectan a la latencia son el tamaño medio de los documentos indexados, el tamaño del índice, las consultas muy complejas y la carga general del índice de Amazon Kendra .

El siguiente ejemplo muestra cómo obtener información sobre las facetas del atributo personalizado «CityRegion», como una faceta anidada dentro de «Ciudad».

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        Facets = [
            {
                "DocumentAttributeKey" : "City",
                "Facets": [
                    {
                        "DocumentAttributeKey" : "CityRegion"
                    }
                ]
            }
        ]
        )
```

La información de las facetas, como el recuento de documentos, se devuelve en la matriz de respuestas de `FacetResults`. El contenido se utiliza para mostrar sugerencias de búsqueda por facetas en la aplicación. Por ejemplo, si el atributo del documento “City” contiene la ciudad a la que se podría aplicar la búsqueda, utilice esa información para mostrar una lista de las búsquedas de ciudades. Los usuarios pueden elegir una ciudad para filtrar los resultados de búsqueda. Para realizar la búsqueda por facetas, llame a la API de consulta y utilice el atributo de documento elegido para filtrar los resultados.

Puede mostrar hasta 10 valores de faceta por faceta de una consulta y solo una faceta anidada dentro de una faceta. Si desea aumentar estos límites, póngase en contacto con Soporte. Si desea

limitar el número de valores de facetas por faceta a menos de 10, puede especificarlo en el objeto `Facet`.

En el siguiente ejemplo de respuesta de JSON, se muestran las facetas limitadas al atributo del documento “City”. La respuesta incluye el recuento de documentos para el valor de la faceta.

```
{
    'FacetResults': [
        {
            'DocumentAttributeKey': 'City',
            'DocumentAttributeValueCountPairs': [
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Dubai'
                    }
                },
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Seattle'
                    }
                },
                {
                    'Count': 1,
                    'DocumentAttributeValue': {
                        'StringValue': 'Paris'
                    }
                }
            ]
        }
    ]
```

También puede mostrar la información de facetas de una faceta anidada, como una región dentro de una ciudad, para filtrar aún más los resultados de búsqueda.

En el siguiente ejemplo de respuesta de JSON, se muestran las facetas relacionadas con el atributo del documento «CityRegion», como una faceta anidada dentro de «Ciudad». La respuesta incluye el recuento de documentos para los valores de la faceta anidada.

```
{
    'FacetResults': [
```

```
        {
            'DocumentAttributeKey': 'City',
            'DocumentAttributeValueCountPairs': [
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Dubai'
                    },
                    'FacetResults': [
                        {
                            'DocumentAttributeKey': 'CityRegion',
                            'DocumentAttributeValueCountPairs': [
                                {
                                    'Count': 2,
                                    'DocumentAttributeValue': {
                                        'StringValue': 'Bur Dubai'
                                    }
                                },
                                {
                                    'Count': 1,
                                    'DocumentAttributeValue': {
                                        'StringValue': 'Deira'
                                    }
                                }
                             ]
                        }
                    ]
                },
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Seattle'
                    },
                    'FacetResults': [
                        {
                            'DocumentAttributeKey': 'CityRegion',
                            'DocumentAttributeValueCountPairs': [
                                {
                                    'Count': 1,
                                    'DocumentAttributeValue': {
                                        'StringValue': 'North'
                                    }
                                },
                                {
```

```
                                'Count': 2,
                                'DocumentAttributeValue': {
                                    'StringValue': 'South'
                                }
                            }
                        ]
                    }
                ]
            },
            {
                'Count': 1,
                'DocumentAttributeValue': {
                    'StringValue': 'Paris'
                },
                'FacetResults': [
                    {
                        'DocumentAttributeKey': 'CityRegion',
                        'DocumentAttributeValueCountPairs': [
                            {
                                'Count': 1,
                                'DocumentAttributeValue': {
                                    'StringValue': 'City center'
                                }
                            }
                        ]
                    }
                ]
            }
        }
    ]
}
```

Cuando utiliza un campo de lista de cadenas para crear facetas, los resultados por facetas devueltos se basan en el contenido de la lista de cadenas. Por ejemplo, si tiene un campo de lista de cadenas que contiene dos elementos, uno con la lista “teckel” y “perro salchicha” y otro con el valor “husky”, obtendrá `FacetResults` con tres facetas.

Para obtener más información, consulte Respuestas a las consultas y tipos de respuestas.

Utilizar atributos del documento para filtrar los resultados de búsqueda

De forma predeterminada, `Query` devuelve todos los resultados de búsqueda. Para filtrar las respuestas, puede realizar operaciones lógicas en los atributos del documento. Por ejemplo, si

solo desea documentos para una ciudad específica, puede filtrar por los atributos de documento personalizados “City” y “State”. Se utiliza AttributeFilterpara crear una operación booleana en los filtros que se proporcionan.

Se puede usar la mayoría de los atributos para filtrar las respuestas de todos los tipos de respuestas. Sin embargo, el atributo `_excerpt_page_number` solo se aplica a los tipos de respuestas `ANSWER` cuando se filtran las respuestas.

El siguiente ejemplo muestra cómo realizar una operación lógica AND filtrando por una ciudad específica, Seattle, y un estado, Washington.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {'AndAllFilters':
            [
                {"EqualsTo": {"Key": "City","Value": {"StringValue": "Seattle"}}},
                {"EqualsTo": {"Key": "State","Value": {"StringValue": "Washington"}}}
            ]
            }
        )
```

El siguiente ejemplo muestra cómo realizar una operación lógica OR cuando alguna de las claves `Fileformat`, `Author` o `SourceURI` coincide con los valores especificados.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {'OrAllFilters':
            [
                {"EqualsTo": {"Key": "Fileformat","Value": {"StringValue":
 "AUTO_DETECT"}}},
                {"EqualsTo": {"Key": "Author","Value": {"StringValue": "Ana
 Carolina"}}},
                {"EqualsTo": {"Key": "SourceURI","Value": {"StringValue": "https://
aws.amazonaws.com/234234242342"}}}
            ]
            }
        )
```

En el caso de los campos `StringList`, debe utilizar los filtros de atributos `ContainsAny` o `ContainsAll` para devolver los documentos con la cadena especificada. El siguiente ejemplo

muestra cómo devolver todos los documentos que tienen los valores “Seattle” o “Portland” en su atributo personalizado `Locations`.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {
                "ContainsAny": { "Key": "Locations", "Value": { "StringListValue":
 [ "Seattle", "Portland"] }}
            }
        )
```

Filtrar los atributos de cada documento en los resultados de búsqueda

Amazon Kendra devuelve los atributos de cada documento de los resultados de la búsqueda. Puede filtrar determinados atributos del documento que desee incluir en la respuesta como parte de los resultados de búsqueda. De forma predeterminada, todos los atributos del documento asignados a un documento se devuelven en la respuesta.

En el siguiente ejemplo, solo los atributos del documento `_source_uri` y `_author` se incluyen en la respuesta para un documento.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        RequestedDocumentAttributes = ["_source_uri", "_author"]
        )
```

Filtrar por contexto de usuario

Puede filtrar los resultados de búsqueda de un usuario según el acceso del usuario o de su grupo a los documentos. Puede usar un token de usuario, un ID de usuario o un atributo de usuario para filtrar los documentos. Amazon Kendra también puede asignar los usuarios a sus grupos. Puede optar por utilizar AWS IAM Identity Center como origen o almacén de identidades.

El filtrado por contexto de usuario es un tipo de búsqueda personalizada con la ventaja de controlar el acceso a los documentos. Por ejemplo, no todos los equipos que buscan información en el portal corporativo deben acceder a los documentos de alto secreto de la empresa, ni estos documentos son relevantes para todos los usuarios. Solo los usuarios o grupos de equipos específicos que

tengan acceso a documentos de alto secreto deberían ver estos documentos en sus resultados de búsqueda.

Cuando se indexa un documento Amazon Kendra, se incorpora la lista de control de acceso (ACL) correspondiente para la mayoría de los documentos. La ACL especifica a qué nombres de usuario y de grupo se les permite o deniega el acceso al documento. Los documentos sin una ACL son documentos públicos.

Amazon Kendra puede extraer la información de usuario o grupo asociada a cada documento para la mayoría de las fuentes de datos. Por ejemplo, un documento de Quip puede incluir una lista "compartida" de usuarios seleccionados que tienen acceso al documento. Si utiliza un bucket de S3 como origen de datos, debe proporcionar un archivo JSON para su ACL e incluir la ruta de S3 a este archivo como parte de la configuración del origen de datos. Si agrega documentos directamente a un índice, especifica la ACL en el objeto principal como parte del objeto de documento de la BatchPutDocumentAPI.

Puede usar la CreateAccessControlConfigurationAPI para volver a configurar su control de acceso a nivel de documento existente sin tener que volver a indexar todos los documentos. Por ejemplo, su índice contiene documentos empresariales de alto secreto a los que solo deben acceder determinados empleados o usuarios. Uno de estos usuarios deja la empresa o pasa a un equipo al que se le debería impedir el acceso a los documentos de alto secreto. El usuario sigue teniendo acceso a los documentos de alto secreto porque tenía acceso a ellos cuando los documentos estaban indexados anteriormente. Puede crear una configuración de control de acceso específica para el usuario con acceso denegado. Más adelante, puede actualizar la configuración de control de acceso para permitirle el acceso en caso de que el usuario regrese a la empresa y vuelva a unirse al equipo "de alto secreto". Puede volver a configurar el control de acceso a sus documentos a medida que cambian las circunstancias.

Para aplicar tu configuración de control de acceso a determinados documentos, llamas a la BatchPutDocumentAPI con el objeto `AccessControlConfigurationId` incluido en el documento. Si utiliza un bucket de S3 como fuente de datos, lo actualiza `.metadata.json` con la fuente de datos `AccessControlConfigurationId` y la sincroniza. Amazon Kendra Actualmente, solo admite la configuración de control de acceso para las fuentes de datos y los documentos de S3 indexados mediante la `BatchPutDocument` API.

Filtrado por token de usuario

Al consultar un índice, puede usar un token de usuario para filtrar los resultados de búsqueda según el acceso del usuario o de su grupo a los documentos. Al realizar una consulta, Amazon

Kendra extrae y valida el token, extrae y comprueba la información del usuario y del grupo y ejecuta la consulta. Se devuelven todos los documentos a los que el usuario tiene acceso, incluidos los documentos públicos. Para obtener más información, consulte Control de acceso de usuarios basado en tokens.

Debe proporcionar el token de usuario en el UserContextobjeto y pasarlo a la API de consultas.

El siguiente ejemplo muestra cómo incluir un token de usuario.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserToken = {
        Token = "token"
    })
```

Puede asignar usuarios a grupos. Al utilizar el filtrado por contexto de usuario, no es necesario incluir todos los grupos a los que pertenece un usuario al realizar la consulta. Con la PutPrincipalMappingAPI, puedes asignar usuarios a sus grupos. Si no desea utilizar la API `PutPrincipalMapping`, debe proporcionar el nombre de usuario y todos los grupos a los que pertenece el usuario cuando realice una consulta. También puede obtener los niveles de acceso de los grupos y usuarios de la fuente de identidad de su centro de identidad de IAM mediante el UserGroupResolutionConfigurationobjeto.

Filtrado por ID de usuario y grupo

Al consultar un índice, puede usar el ID de usuario y el grupo para filtrar los resultados de búsqueda según el acceso del usuario o de su grupo a los documentos. Al realizar una consulta, Amazon Kendra comprueba la información del usuario y del grupo y ejecuta la consulta. Se devuelven todos los documentos relevantes para la consulta a los que el usuario tiene acceso, incluidos los documentos públicos.

También puede filtrar los resultados de búsqueda por los orígenes de datos a las que tienen acceso los usuarios y los grupos. Especificar un origen de datos resulta útil si un grupo está vinculado a varios orígenes de datos, pero solo desea que el grupo acceda a los documentos de un origen de datos determinado. Por ejemplo, pongamos que los grupos “Investigación”, “Ingeniería” y “Ventas y marketing” están todos vinculados a los documentos de la empresa almacenados en los orígenes de datos Confluence y Salesforce. Sin embargo, solo el equipo de “Ventas y marketing” necesita acceder a los documentos relacionados con los clientes almacenados en Salesforce. De este modo,

cuando los usuarios de ventas y marketing busquen documentos relacionados con los clientes, podrán ver los documentos de Salesforce en sus resultados. Los usuarios que no trabajan en ventas y marketing no ven los documentos de Salesforce en sus resultados de búsqueda.

Debe proporcionar la información sobre el usuario, los grupos y las fuentes de datos en el UserContextobjeto y transferirla a la API de consultas. El ID de usuario y la lista de grupos y orígenes de datos deben coincidir con el nombre que especifique en el objeto Principal para identificar al usuario, los grupos y los orígenes de datos. Con el objeto `Principal`, puede añadir un usuario, un grupo o un origen de datos a una lista de permisos o de denegaciones para acceder a un documento.

Debe proporcionar uno de los siguientes datos:

- Información de usuarios y grupos, e información (opcional) sobre los orígenes de datos.
- Solo la información del usuario si asigna sus usuarios a grupos y fuentes de datos mediante la PutPrincipalMappingAPI. También puede obtener los niveles de acceso de los grupos y usuarios de la fuente de identidad de su centro de identidad de IAM mediante el UserGroupResolutionConfigurationobjeto.

Si esta información no está incluida en la consulta, Amazon Kendra devuelve todos los documentos. Si proporciona esta información, solo devolverá los documentos con los ID de usuario, grupos y orígenes de datos que coinciden.

El siguiente ejemplo muestra cómo incluir los ID de usuario, los grupos y los orígenes de datos.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserId = {
        UserId = "user1"
    },
    Groups = {
        Groups = ["Sales and Marketing"]
    },
    DataSourceGroups = {
        DataSourceGroups = [{"DataSourceId" : "SalesforceCustomerDocsGroup", "GroupId":
 "Sales and Marketing"}]
    })
```

Filtrado por atributo de usuario

Al consultar un índice, puede usar los atributos integrados `_user_id` y `_group_id` para filtrar los resultados de búsqueda según el acceso del usuario y de su grupo a los documentos. Puede configurar hasta 100 identificadores de grupo. Al realizar una consulta, Amazon Kendra comprueba la información del usuario y del grupo y ejecuta la consulta. Se devuelven todos los documentos relevantes para la consulta a los que el usuario tiene acceso, incluidos los documentos públicos.

Debe proporcionar los atributos de usuario y grupo en el AttributeFilterobjeto y pasarlos a la API de consultas.

El siguiente ejemplo muestra una solicitud que filtra la respuesta a la consulta en función del ID de usuario y de los grupos “HR” e “IT” a los que pertenece el usuario. La consulta devolverá cualquier documento que contenga el usuario o los grupos “HR” o “IT” en su la lista de permitidos. Si el usuario o alguno de los grupos figuran en la lista de denegados de un documento, ese documento no se devolverá.

```
response = kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {
            "OrAllFilters": [
                {
                    "EqualsTo": {
                        "Key": "_user_id",
                        "Value": {
                            "StringValue": "user1"
                        }
                     }
                },
                {
                    "EqualsTo": {
                        "Key": "_group_ids",
                        "Value": {
                            "StringListValue": ["HR", "IT"]
                        }
                    }
                }
            ]
        }
        )
```

También puede especificar a qué origen de datos puede acceder un grupo en el objeto `Principal`.

> **Note**
>
> El filtrado por contexto de usuario no es un control de autenticación o autorización del contenido. No autentica a los usuarios ni a los grupos enviados a la API `Query`. Es responsabilidad de su aplicación garantizar que la información de usuarios y grupos enviada a la API `Query` esté autenticada y autorizada.

Existe una implementación del filtrado por contexto de usuario para cada origen de datos. En la siguiente sección se describe cada implementación.

Temas

- Filtrado por contexto de usuario para los documentos añadidos directamente a un índice
- Filtrado por contexto de usuario para las preguntas más frecuentes
- Filtrado por contexto de usuario para los orígenes de datos

Filtrado por contexto de usuario para los documentos añadidos directamente a un índice

Al añadir documentos directamente a un índice mediante la BatchPutDocumentAPI, Amazon Kendra obtiene información de usuarios y grupos del `AccessControlList` campo del documento. Debe proporcionar una lista de control de acceso (ACL) para los documentos y la ACL se incorpora con los documentos.

Debe especificar la ACL en el objeto Principal como parte del objeto Documento de la API `BatchPutDocument`. Debe proporcionar la siguiente información:

- El acceso que debe tener el usuario o el grupo. Puede decir `ALLOW` o `DENY`.
- El tipo de entidad. Puede decir `USER` o `GROUP`.
- Nombre del usuario o grupo.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para las preguntas más frecuentes

Al añadir una pregunta frecuente a un índice, Amazon Kendra obtiene información sobre el usuario y el grupo del `AccessControlList` objeto o campo del archivo JSON de preguntas frecuentes. También puede usar un archivo CSV de preguntas frecuentes con campos o atributos personalizados para el control de acceso.

Debe proporcionar la siguiente información:

- El acceso que debe tener el usuario o el grupo. Puede decir `ALLOW` o `DENY`.
- El tipo de entidad. Puede decir `USER` o `GROUP`.
- Nombre del usuario o grupo.

Para obtener más información, consulte Archivos de preguntas frecuentes.

Filtrado por contexto de usuario para los orígenes de datos

Amazon Kendra también rastrea la información de la lista de control de acceso (ACL) de usuarios y grupos desde los conectores de fuentes de datos compatibles. Esto resulta útil para el filtrado por contexto de usuario, en el que se filtran los resultados de búsqueda en función del acceso del usuario o de su grupo a los documentos.

Temas

- Filtrado por contexto de usuario para los orígenes de datos de Adobe Experience Manager
- Filtrado por contexto de usuario para los orígenes de datos de Alfresco
- Filtrado de contexto de usuario para fuentes de datos Aurora (MySQL)
- Filtrado por contexto de usuario para los orígenes de datos de Aurora (PostgreSQL)
- Filtrado del contexto de usuario para Amazon FSx las fuentes de datos
- Filtrado por contexto de usuario para los orígenes de datos de bases de datos
- Filtrado por contexto de usuario para orígenes de datos de Amazon RDS (Microsoft SQL Server).
- Filtrado de contexto de usuario para fuentes de datos Amazon RDS (MySQL)
- Filtrado de contexto de usuario para fuentes de datos Amazon RDS (Oracle)
- Filtrado por contexto de usuario para los orígenes de datos de Amazon RDS (PostgreSQL)
- Filtrado por contexto de usuario para los orígenes de datos de Amazon S3
- Filtrado de contexto de usuario para fuentes Amazon WorkDocs de datos

- Filtrado por contexto de usuario para los orígenes de datos de Box
- Filtrado por contexto de usuario para los orígenes de datos de Confluence
- Filtrado por contexto de usuario para los orígenes de datos de Dropbox
- Filtrado por contexto de usuario para los orígenes de datos de Drupal
- Filtrado del contexto de usuario para las fuentes de datos GitHub
- Filtrado por contexto de usuario para los orígenes de datos de Gmail
- Filtrado por contexto de usuario para los orígenes de datos de Google Drive
- Filtrado por contexto de usuario para los orígenes de datos de IBM DB2
- Filtrado por contexto de usuario para los orígenes de datos de Jira
- Filtrado por contexto de usuario para orígenes de datos de Microsoft Exchange
- Filtrado de contexto de usuario para fuentes OneDrive de datos de Microsoft
- Filtrado de contexto de usuario para fuentes de datos de Microsoft OneDrive v2.0
- Filtrado de contexto de usuario para fuentes SharePoint de datos de Microsoft
- Filtrado por contexto de usuario para orígenes de datos de Microsoft SQL Server.
- Filtrado por contexto de usuario para orígenes de datos de Microsoft Teams
- Filtrado por contexto de usuario para orígenes de datos de Microsoft Yammer
- Filtrado por contexto de usuario para los orígenes de datos de MySQL
- Filtrado por contexto de usuario para los orígenes de datos de Oracle Database
- Filtrado por contexto de usuario para los orígenes de datos de PostgreSQL
- Filtrado por contexto de usuario para los orígenes de datos de Quip
- Filtrado por contexto de usuario para los orígenes de datos de Salesforce
- Filtrado por contexto de usuario para los orígenes de datos de ServiceNow
- Filtrado por contexto de usuario para los orígenes de datos de Slack
- Filtrado por contexto de usuario para los orígenes de datos de Zendesk

Filtrado por contexto de usuario para los orígenes de datos de Adobe Experience Manager

Cuando utiliza una fuente de datos de Adobe Experience Manager, Amazon Kendra obtiene la información de usuario y grupo de la instancia de Adobe Experience Manager.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`: los ID de grupo existen en el contenido de Adobe Experience Manager, donde hay permisos de acceso establecidos. Se asignan a partir de los nombres de los grupos en Adobe Experience Manager.
- `_user_id`: los ID de usuario existen en el contenido de Adobe Experience Manager, donde hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID en Adobe Experience Manager.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Alfresco

Cuando utiliza una fuente de datos de Alfresco, Amazon Kendra obtiene la información de usuario y grupo de la instancia de Alfresco.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`: los ID de grupo existen en Alfresco en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los nombres de sistema de los grupos (no de los nombres de visualización) de Alfresco.
- `_user_id`: los ID de usuario existen en Alfresco en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID en Alfresco.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado de contexto de usuario para fuentes de datos Aurora (MySQL)

Cuando utiliza una fuente de datos Aurora (MySQL), Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Una fuente de datos de base de datos Aurora (MySQL) tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.

- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado por contexto de usuario para los orígenes de datos de Aurora (PostgreSQL)

Cuando utiliza una fuente de datos Aurora (PostgreSQL) Amazon Kendra , obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Una fuente de datos de base de datos Aurora (PostgreSQL) tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado del contexto de usuario para Amazon FSx las fuentes de datos

Cuando utiliza una fuente de Amazon FSx datos, Amazon Kendra obtiene información de usuarios y grupos del servicio de directorio de la Amazon FSx instancia.

Los ID de Amazon FSx grupo y usuario se mapean de la siguiente manera:

- `_group_ids`: los ID de grupo existen en Amazon FSx en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los nombres de los grupos del sistema en el servicio de directorio de. Amazon FSx
- `_user_id`—Los ID de usuario existen en los archivos Amazon FSx en los que hay permisos de acceso establecidos. Se asignan a partir de los nombres de usuario del sistema en el servicio de directorio de. Amazon FSx

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de bases de datos

Cuando utiliza una fuente de datos de base de datos, por ejemplo Amazon Aurora PostgreSQL, Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla de fuentes.

Esta columna se especifica en el AclConfigurationobjeto como parte del DatabaseConfigurationobjeto de la CreateDataSourceAPI.

Un origen de datos de bases de datos tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado por contexto de usuario para orígenes de datos de Amazon RDS (Microsoft SQL Server).

Cuando utiliza una fuente de datos Amazon RDS (Microsoft SQL Server), Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Una fuente de datos de base de datos Amazon RDS (Microsoft SQL Server) tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado de contexto de usuario para fuentes de datos Amazon RDS (MySQL)

Cuando utiliza una fuente de datos Amazon RDS (MySQL), Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Una fuente de datos de base de datos Amazon RDS (MySQL) tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado de contexto de usuario para fuentes de datos Amazon RDS (Oracle)

Cuando utiliza una fuente de datos Amazon RDS (Oracle), Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Una fuente de datos de base de datos Amazon RDS (Oracle) tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado por contexto de usuario para los orígenes de datos de Amazon RDS (PostgreSQL)

Cuando utiliza una fuente de datos Amazon RDS (PostgreSQL) Amazon Kendra , obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Una fuente de datos de base de datos Amazon RDS (PostgreSQL) tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado por contexto de usuario para los orígenes de datos de Amazon S3

El filtrado por contexto de usuario se añade a un documento de una fuente de Amazon S3 datos mediante un archivo de metadatos asociado al documento. Debe añadir la información al campo `AccessControlList` del documento JSON. Para obtener más información sobre cómo añadir metadatos a los documentos indexados desde un origen de datos de Amazon S3 , consulte Metadatos de documentos de S3.

Debe proporcionar tres datos:

- El acceso que debe tener la entidad. Puede decir `ALLOW` o `DENY`.
- El tipo de entidad. Puede decir `USER` o `GROUP`.
- El nombre de la entidad.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado de contexto de usuario para fuentes Amazon WorkDocs de datos

Cuando utiliza una fuente de Amazon WorkDocs datos, Amazon Kendra obtiene información de usuarios y grupos de la Amazon WorkDocs instancia.

Los ID de Amazon WorkDocs grupo y usuario se mapean de la siguiente manera:

- `_group_ids`—Los ID de grupo existen en los archivos Amazon WorkDocs en los que hay permisos de acceso establecidos. Se asignan a partir de los nombres de los grupos incluidos en ellos. Amazon WorkDocs
- `_user_id`—Los ID de usuario existen en los archivos Amazon WorkDocs en los que hay permisos de acceso establecidos. Se mapean a partir de los nombres de usuario que aparecen en. Amazon WorkDocs

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Box

Cuando utiliza una fuente de datos de Box, Amazon Kendra obtiene información de usuarios y grupos de la instancia de Box.

Los ID de grupo y usuario de Box se asignan de la siguiente manera:

- `_group_ids`: los ID de grupo existen en Box en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los nombres de los grupos en Box.
- `_user_id`: los ID de usuario existen en Box en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID de usuario en Box.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Confluence

Cuando utilizas una fuente de datos de Confluence, Amazon Kendra obtiene información de usuarios y grupos de la instancia de Confluence.

Debe configurar el acceso de los usuarios y grupos a los espacios mediante la página de permisos de los espacios. Para las páginas y los blogs, debe utilizar la página de restricciones. Para obtener más información sobre los permisos de espacio, consulte Información general de los permisos de espacio en el sitio web de soporte de Confluence. Para obtener más información sobre las restricciones de páginas y blogs, consulte Información general de las restricciones de espacio en el sitio web de soporte de Confluence.

Los nombres de grupo y usuario de Confluence se asignan de la siguiente manera:

- `_group_ids`: los nombres de grupo están presentes en los espacios, páginas y blogs donde hay restricciones. Se asignan a partir del nombre del grupo en Confluence. Los nombres de grupo siempre están en minúscula.

- `_user_id`: los nombres de usuario están presentes en el espacio, página o blog donde hay restricciones. Se asignan en función del tipo de instancia de Confluence que utilice.

  Para el conector de Confluence v1.0

  - Servidor: `_user_id` es el nombre de usuario. El nombre de usuario siempre está en minúscula.
  - Nube: `_user_id` es el ID de cuenta del usuario.

  Para el conector de Confluence v2.0

  - Servidor: `_user_id` es el nombre de usuario. El nombre de usuario siempre está en minúscula.
  - Nube: `_user_id` es el ID de correo electrónico del usuario.

> **Important**
>
> Para que el filtrado por contexto de usuario funcione correctamente en su conector de Confluence, debe asegurarse de que la visibilidad de un usuario al que se le ha concedido acceso a una página de Confluence esté configurada como Cualquiera. Para obtener más información, consulte Configurar la visibilidad del correo electrónico en la documentación para desarrolladores de Atlassian.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Dropbox

Cuando utilizas una fuente de datos de Dropbox, Amazon Kendra obtiene la información del usuario y del grupo de la instancia de Dropbox.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`: los ID de grupo existen en Dropbox en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los nombres de los grupos en Dropbox.
- `_user_id`: los ID de usuario existen en Dropbox en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID en Dropbox.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Drupal

Cuando utilizas una fuente de datos de Drupal, Amazon Kendra obtiene la información del usuario y del grupo de la instancia de Drupal.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`: los ID de grupo existen en Drupal en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los nombres de los grupos en Drupal.
- `_user_id`: los ID de usuario existen en Drupal en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID en Drupal.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado del contexto de usuario para las fuentes de datos GitHub

Cuando utiliza una fuente de GitHub datos, Amazon Kendra obtiene información de usuario de la GitHub instancia.

Los ID GitHub de usuario se mapean de la siguiente manera:

- `_user_id`—Los ID de usuario existen en los archivos GitHub en los que hay permisos de acceso establecidos. Se mapean a partir de los correos electrónicos de los usuarios como identificadores. GitHub

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Gmail

Cuando utilizas una fuente de datos de Gmail, Amazon Kendra obtiene la información del usuario de la instancia de Gmail.

Los ID de usuario se asignan de la siguiente manera:

- `_user_id`: los ID de usuario existen en Gmail en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID en Gmail.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Google Drive

Un origen de datos de Google Workspace Drive devuelve información de usuarios y grupos para usuarios y grupos de Google Drive. La pertenencia a grupos y dominios se asigna al campo del índice `_group_ids`. El nombre de usuario de Google Drive se asigna al campo `_user_id`.

Si proporciona una o más direcciones de correo electrónico de usuario en la API `Query`, solo se devuelven los documentos que se hayan compartido con esas direcciones de correo electrónico. El siguiente parámetro `AttributeFilter` solo devuelve los documentos compartidos con “martha@example.com”.

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_user_id",
```

```
                "Value": {
                    "StringValue": "martha@example.com"
                }
            }
        }
```

Si proporciona una o más direcciones de correo electrónico de grupos en la consulta, solo se devolverán los documentos compartidos con los grupos. El siguiente parámetro `AttributeFilter` solo devuelve los documentos compartidos con el grupo “hr@example.com”.

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_group_ids",
                   "Value": {
                       "StringListValue": ["hr@example.com"]
                   }
               }
           }
```

Si proporciona el dominio en la consulta, se devolverán todos los documentos compartidos con el dominio. El siguiente parámetro `AttributeFilter` devuelve los documentos compartidos con el dominio “example.com”.

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_group_ids",
                   "Value": {
                       "StringListValue": ["example.com"]
                   }
               }
           }
```

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de IBM DB2

Cuando utiliza una fuente de datos de IBM DB2, Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Un origen de datos de una base de datos de IBM DB2 tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado por contexto de usuario para los orígenes de datos de Jira

Cuando utilizas una fuente de datos de Jira, Amazon Kendra obtiene información de usuarios y grupos de la instancia de Jira.

Los ID de usuario de Jira se asignan de la siguiente manera:

- `_user_id`: los ID de usuario existen en Jira en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID de usuario en Jira.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para orígenes de datos de Microsoft Exchange

Cuando utiliza una fuente de datos de Microsoft Exchange, Amazon Kendra obtiene la información del usuario de la instancia de Microsoft Exchange.

Los ID de usuario de Microsoft Exchange se asignan de la siguiente manera:

- `_user_id`—Los ID de usuario existen en los permisos de Microsoft Exchange para que los usuarios accedan a determinado contenido. Se asignan a partir de los nombres de usuario como identificadores en Microsoft Exchange.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado de contexto de usuario para fuentes OneDrive de datos de Microsoft

Amazon Kendra recupera información de usuarios y grupos de Microsoft OneDrive cuando indexa los documentos del sitio. La información de usuario y grupo se toma del SharePoint sitio de Microsoft subyacente que aloja OneDrive.

Cuando utilice un OneDrive usuario o un grupo para filtrar los resultados de la búsqueda, calcule el identificador de la siguiente manera:

1. Obtenga el nombre del sitio. Por ejemplo, `https://host.onmicrosoft.com/sites/siteName`.
2. Tome el hash MD5 del nombre del sitio. Por ejemplo, `430a6b90503eef95c89295c8999c7981`.
3. Cree el ID del correo electrónico del usuario o el ID del grupo concatenando el hash MD5 con una barra vertical (|) y el ID. Por ejemplo, si el nombre de un grupo es "localGroupName«, el ID del grupo sería:

   `"430a6b90503eef95c89295c8999c7981 | localGroupName"`

   > **Note**
   >
   > Incluye un espacio antes y después de la barra vertical. La barra vertical se utiliza para identificarse `localGroupName` con su hash MD5.

   Para el nombre de usuario “someone@host.onmicrosoft.com”, el ID de usuario sería el siguiente:

   `"430a6b90503eef95c89295c8999c7981 | someone@host.onmicrosoft.com"`

Envía el ID de usuario o grupo Amazon Kendra como `_group_id` atributo `_user_id` o cuando llames a la API de consulta. Por ejemplo, el AWS CLI comando que usa un grupo para filtrar los resultados de la búsqueda tiene el siguiente aspecto:

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "430a6b90503eef95c89295c8999c7981 |
 localGroupName"}
                  }}'
```

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado de contexto de usuario para fuentes de datos de Microsoft OneDrive v2.0

Una fuente de datos de Microsoft OneDrive v2.0 devuelve información de secciones y páginas de las entidades de la lista de control de OneDrive acceso (ACL). Amazon Kendra usa el dominio OneDrive arrendatario para conectarse a la OneDrive instancia y, a continuación, puede filtrar los resultados de la búsqueda en función del acceso de los usuarios o grupos a las secciones y los nombres de los archivos.

En el caso de los objetos estándar, los `_user_id` y `_group_id` se utilizan de la siguiente manera:

- `_user_id`— Su ID OneDrive de correo electrónico de usuario de Microsoft está asignado al `_user_id` campo.
- `_group_id`— El correo electrónico de su OneDrive grupo de Microsoft está asignado al `_group_id` campo.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado de contexto de usuario para fuentes SharePoint de datos de Microsoft

Amazon Kendra recupera la información de usuarios y grupos de Microsoft SharePoint cuando indexa los documentos del sitio. Para filtrar los resultados de la búsqueda en función del acceso de los usuarios o grupos, proporciona información sobre los usuarios y los grupos cuando llames a la `Query` API.

Para filtrar mediante un nombre de usuario, utilice la dirección de correo electrónico del usuario. Por ejemplo, johnstiles@example.com.

Cuando utilices un SharePoint grupo para filtrar los resultados de la búsqueda, calcula el ID del grupo de la siguiente manera:

Para grupos locales

1. Obtenga el nombre del sitio. Por ejemplo, `https://host.onmicrosoft.com/sites/siteName.`
2. Tome el hash SHA256 del nombre del sitio. Por ejemplo, `430a6b90503eef95c89295c8999c7981`.
3. Cree el ID del grupo concatenando el hash SHA256 con una barra vertical (|) y el nombre del grupo. Por ejemplo, si el nombre del grupo es "localGroupName«, el ID del grupo sería:

```
"430a6b90503eef95c89295c8999c7981 | localGroupName"
```

> **Note**
>
> Incluye un espacio antes y después de la barra vertical. La barra vertical se utiliza para identificarse `localGroupName` con su hash SHA256.

Envía el ID de grupo Amazon Kendra como `_group_id` atributo cuando llames a la API de consulta. Por ejemplo, el AWS CLI comando tiene este aspecto:

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "430a6b90503eef95c89295c8999c7981 |
 localGroupName"}
                  }}'
```

Para grupos de AD

1. Usa el ID de grupo de AD para configurar el filtrado de los resultados de búsqueda.

Envía el ID de grupo Amazon Kendra como `_group_id` atributo cuando llames a la API de consultas. Por ejemplo, el AWS CLI comando tiene este aspecto:

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "AD group"}
                  }}'
```

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para orígenes de datos de Microsoft SQL Server.

Cuando utiliza una fuente de datos de Microsoft SQL Server, Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Un origen de datos de la base de datos de Microsoft SQL Server tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado por contexto de usuario para orígenes de datos de Microsoft Teams

Amazon Kendra recupera la información de usuario de Microsoft Teams cuando indexa los documentos. La información del usuario se toma de la instancia de Microsoft Teams subyacente.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para orígenes de datos de Microsoft Yammer

Amazon Kendra recupera la información del usuario de Microsoft Yammer cuando indexa los documentos. La información de usuario y grupo se toma de la instancia subyacente de Microsoft Yammer.

Los ID de usuario de Microsoft Yammer se asignan de la siguiente manera:

- `_email_id`— El ID de correo electrónico de Microsoft asignado al `_user_id` campo.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de MySQL

Cuando utiliza una fuente de datos MySQL, Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Un origen de datos de la base de datos de MySQL tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado por contexto de usuario para los orígenes de datos de Oracle Database

Cuando utiliza una fuente de datos de Oracle Database, Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Un origen de datos de la base de datos de Oracle Database tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado por contexto de usuario para los orígenes de datos de PostgreSQL

Cuando utiliza una fuente de datos de PostgreSQL Amazon Kendra , obtiene información de usuarios y grupos de una columna de la tabla de fuentes. Esta columna se especifica en la consola o se utiliza el TemplateConfigurationobjeto como parte de la CreateDataSourceAPI.

Un origen de datos de la base de datos de PostgreSQL tiene las siguientes limitaciones:

- Solo puede especificar una lista de permisos para un origen de datos de la base de datos. No puede especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permisos.

- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado por contexto de usuario para los orígenes de datos de Quip

Cuando utiliza una fuente de datos de Quip, Amazon Kendra obtiene la información de usuario de la instancia de Quip.

Los ID de usuario de Quip se asignan de la siguiente manera:

- `_user_id`: los ID de usuario existen en Quip en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID en Quip.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Salesforce

Un origen de datos de Salesforce devuelve información de usuarios y grupos a partir de las entidades de la lista de control de acceso (ACL) de Salesforce. Puede aplicar el filtrado por contexto de usuario a los objetos estándar o las fuentes de chat de Salesforce. El filtrado por contexto de usuario no está disponible para los artículos de conocimiento de Salesforce.

Si asigna cualquier campo de Salesforce a los campos de título y cuerpo del documento de Amazon Kendra, Amazon Kendra utilizará los datos de los campos de título y cuerpo del documento en las respuestas de búsqueda.

En el caso de los objetos estándar, los `_user_id` y `_group_ids` se utilizan de la siguiente manera:

- `_user_id`: el nombre de usuario del usuario de Salesforce.
- `_group_ids`—
  - Nombre de `Profile` de Salesforce
  - Nombre de `Group` de Salesforce
  - Nombre de `UserRole` de Salesforce
  - Nombre de `PermissionSet` de Salesforce

Para las fuentes de chat, los `_user_id` y `_group_ids` se utilizan de la siguiente manera:

- `_user_id`: el nombre de usuario del usuario de Salesforce. Solo está disponible si el elemento está publicado en la fuente del usuario.
- `_group_ids`: los ID de grupo se utilizan de la siguiente manera. Solo está disponible si el elemento de la fuente se publica en un grupo de colaboración o chat.
  - El nombre del grupo de colaboración o chat.
  - Si el grupo es público, `PUBLIC:ALL`.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de ServiceNow

El filtrado del contexto de usuario solo ServiceNow se admite en la TemplateConfiguration API y ServiceNow en la versión 2.0 de Connector. ServiceNowConfigurationLa API y ServiceNow Connector v1.0 no admiten el filtrado del contexto de los usuarios.

Cuando usa una fuente de ServiceNow datos, Amazon Kendra obtiene la información de usuario y grupo de la ServiceNow instancia.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`—Los ID de grupo existen ServiceNow en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los nombres de los roles de `sys_ids` in. ServiceNow
- `_user_id`—Los ID de usuario existen en los archivos ServiceNow en los que hay permisos de acceso establecidos. Se mapean a partir de los correos electrónicos de los usuarios como identificadores. ServiceNow

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Slack

Cuando utilizas una fuente de datos de Slack, Amazon Kendra obtiene la información del usuario de la instancia de Slack.

Los ID de usuario de Slack se asignan de la siguiente manera:

- `_user_id`: los ID de usuario existen en Slack en los mensajes y canales en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID en Slack.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Filtrado por contexto de usuario para los orígenes de datos de Zendesk

Cuando se utiliza una fuente de datos de Zendesk, Amazon Kendra obtiene la información del usuario y del grupo de la instancia de Zendesk.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`: los ID de grupo existen en los tickets y artículos de Zendesk en los que hay permisos de acceso establecidos. Se asignan a partir de los nombres de los grupos en Zendesk.
- `_user_id`: los ID de grupo existen en los tickets y artículos de Zendesk en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como los ID en Zendesk.

Puede añadir hasta 200 entradas en el campo `AccessControlList`.

Respuestas a las consultas y tipos de respuestas

Amazon Kendra admite diferentes respuestas a consultas y tipos de respuesta.

Respuestas a las consultas

Llamar a la API de consulta devuelve información sobre los resultados de una búsqueda. Los resultados se encuentran en una matriz de QueryResultItemobjetos (`ResultItems`). Cada `QueryResultItem` incluye un resumen del resultado. También incluye los atributos del documento asociados al resultado de la consulta.

Información del resumen

La información del resumen varía según el tipo de resultado. En cada caso, incluye el texto del documento que coincide con el término de búsqueda. También incluye información resaltada que puede usar para resaltar el texto de la búsqueda en el resultado de la aplicación. Por ejemplo, si la búsqueda es ¿Cuál es la altura de la Space Needle?, la información del resumen incluye la ubicación en el texto de las palabras altura y space needle. Para obtener más información sobre los tipos de respuestas, consulte Respuestas a las consultas y tipos de respuestas.

Atributos del documento

Cada resultado contiene los atributos del documento que coincide con una consulta. Algunos de los atributos están predefinidos como `DocumentId`, `DocumentTitle`, y `DocumentUri`. Otros son atributos personalizados que puede definir. Puede usar los atributos del documento para filtrar la respuesta de la API `Query`. Por ejemplo, es posible que solo desee los documentos escritos por un autor específico o la versión específica de un documento. Para obtener más información, consulte Filtrado y búsqueda por facetas. Debe especificar los atributos del documento al añadir documentos a un índice. Para obtener más información, consulte Campos o atributos personalizados.

El siguiente ejemplo muestra un código JSON para el resultado de una consulta. Observe los atributos del documento en `DocumentAttributes` y`AdditionalAttributes`.

```
{
    "QueryId": "query-id",
    "ResultItems": [
        {
            "Id": "result-id",
            "Type": "ANSWER",
            "AdditionalAttributes": [
                {
                    "Key": "AnswerText",
                    "ValueType": "TEXT_WITH_HIGHLIGHTS_VALUE",
                    "Value": {
                        "TextWithHighlightsValue": {
                            "Text": "text",
                            "Highlights": [
                                {
                                    "BeginOffset": 55,
                                    "EndOffset": 90,
                                    "TopAnswer": false
                                }
                            ]
                        }
                    }
                }
            ],
            "DocumentId": "document-id",
            "DocumentTitle": {
                "Text": "title"
            },
            "DocumentExcerpt": {
                "Text": "text",
                "Highlights": [
```

```
                {
                    "BeginOffset": 0,
                    "EndOffset": 300,
                    "TopAnswer": false
                }
            ]
        },
        "DocumentURI": "uri",
        "DocumentAttributes": [],
        "ScoreAttributes": "score",
        "FeedbackToken": "token"
    },
    {
        "Id": "result-id",
        "Type": "ANSWER",
        "Format": "TABLE",
        "DocumentId": "document-id",
        "DocumentTitle": {
            "Text": "title"
        },
        "TableExcerpt": {
            "Rows": [{
                "Cells": [{
                    "Header": true,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }, {
                    "Header": true,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }, {
                    "Header": true,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }, {
                    "Header": true,
                    "Highlighted": false,
                    "TopAnswer": false,
                    "Value": "value"
                }]
            }, {
```

```
                    "Cells": [{
                        "Header": false,
                        "Highlighted": false,
                        "TopAnswer": false,
                        "Value": "value"
                    }, {
                        "Header": false,
                        "Highlighted": false,
                        "TopAnswer": false,
                        "Value": "value"
                    }, {
                        "Header": false,
                        "Highlighted": true,
                        "TopAnswer": true,
                        "Value": "value"
                    }, {
                        "Header": false,
                        "Highlighted": false,
                        "TopAnswer": false,
                        "Value": "value"
                    ]}
                }],
                    "TotalNumberofRows": number
    },
            "DocumentURI": "uri",
            "ScoreAttributes": "score",
            "FeedbackToken": "token"
        },
        {
            "Id": "result-id",
            "Type": "DOCUMENT",
            "AdditionalAttributes": [],
            "DocumentId": "document-id",
            "DocumentTitle": {
                "Text": "title",
                "Highlights": []
            },
            "DocumentExcerpt": {
                "Text": "text",
                "Highlights": [
                    {
                        "BeginOffset": 74,
                        "EndOffset": 77,
                        "TopAnswer": false
```

```
                    }
                ]
            },
            "DocumentURI": "uri",
            "DocumentAttributes": [
                {
                    "Key": "_source_uri",
                    "Value": {
                        "StringValue": "uri"
                    }
                }
            ],
            "ScoreAttributes": "score",
            "FeedbackToken": "token",
        }
    ],
    "FacetResults": [],
    "TotalNumberOfResults": number
}
```

Tipos de respuestas

Amazon Kendra devuelve tres tipos de respuesta a la consulta.

- Respuesta (incluye respuestas de tabla)
- Documento
- Pregunta y respuesta

El tipo de respuesta se devuelve en el campo de `Type` respuesta del QueryResultItemobjeto.

Respuesta

Amazon Kendra detectó una o más respuestas a una pregunta en la respuesta. Una trivialidad es la respuesta a una pregunta sobre quién, qué, cuándo o dónde. Por ejemplo: ¿Dónde está el centro de servicios más cercano a mi ubicación? Amazon Kendra devuelve el texto del índice que mejor coincide con la consulta. El texto está en el campo `AnswerText` y contiene información destacada sobre el término de búsqueda en el texto de la respuesta. `AnswerText` incluye el extracto completo del documento con el texto resaltado, mientras que `DocumentExcerpt` incluye el extracto del documento truncado (290 caracteres) con el texto resaltado.

Amazon Kendra solo devuelve una respuesta por documento, y esa es la respuesta con mayor confianza. Para obtener varias respuestas de un documento, debe dividir el documento en varios documentos.


```
{
    'AnswerText': {
        'TextWithHighlights': [
            {
                'BeginOffset': 271,
                'EndOffset': 279,
                'TopAnswer': False
            },
            {
                'BeginOffset': 481,
                'EndOffset': 489,
                'TopAnswer': False
            },
            {
                'BeginOffset': 547,
                'EndOffset': 555,
                'TopAnswer': False
            },
            {
                'BeginOffset': 764,
                'EndOffset': 772,
                'TopAnswer': False
            }
        ],
        'Text': 'Asynchronousoperationscan\n''alsoprocess
\n''documentsthatareinPDF''format.UsingPDFformatfilesallowsyoutoprocess''multi-
page\n''documents.\n''Forinformationabouthow''AmazonTextractrepresents
\n''documentsasBlockobjects,
        ''seeDocumentsandBlockObjects.
\n''\n''\n''\n''Forinformationaboutdocument''limits,
        seeLimitsinAmazonTextract.
\n''\n''\n''\n''TheAmazonTextractsynchronous''operationscanprocessdocumentsstoredinanAmazon
\n''S3Bucketoryoucanpass''base64encodedimagebytes.\n''Formoreinformation,

 see''CallingAmazonTextractSynchronousOperations.''Asynchronousoperationsrequireinputdocuments
\n''tobesuppliedinanAmazon''S3Bucket.'
    },
    'DocumentExcerpt': {
        'Highlights': [
```

```
            {
                'BeginOffset': 0,
                'EndOffset': 300,
                'TopAnswer': False
            }
        ],
        'Text': 'Asynchronousoperationscan\n''alsoprocess
\n''documentsthatareinPDF''format.UsingPDFformatfilesallowsyoutoprocess''multi-page
\n''documents.\n''ForinformationabouthowAmazon''Textractrepresents\n'''
    },
    'Type': 'ANSWER'
}
```

Documento

Amazon Kendra devuelve los documentos clasificados de los que coinciden con el término de búsqueda. La clasificación se basa en la confianza que Amazon Kendra se deposita en la precisión del resultado de la búsqueda. La información sobre el documento coincidente se devuelve en el QueryResultItem. Incluye el título del documento. El extracto incluye información destacada para el texto de búsqueda y la sección de texto coincidente del documento. El URI de los documentos coincidentes se encuentra en el atributo del documento `SourceURI`. En el siguiente ejemplo de JSON se muestra el resumen de un documento coincidente.

```
{
    'DocumentTitle': {
        'Highlights': [
            {
                'BeginOffset': 7,
                'EndOffset': 15,
                'TopAnswer': False
            },
            {
                'BeginOffset': 97,
                'EndOffset': 105,
                'TopAnswer': False
            }
        ],
        'Text': 'AmazonTextractAPIPermissions: Actions,
        \n''Permissions,
        andResourcesReference-''AmazonTextract'
    },
    'DocumentExcerpt': {
```

```
        'Highlights': [
            {
                'BeginOffset': 68,
                'EndOffset': 76,
                'TopAnswer': False
            },
            {
                'BeginOffset': 121,
                'EndOffset': 129,
                'TopAnswer': False
            }
        ],
        'Text': '...LoggingandMonitoring\tMonitoring
\n''\tCloudWatchMetricsforAmazonTextract
\n''\tLoggingAmazonTextractAPICallswithAWSCloudTrail\n''\tAPIReference\tActions
\tAnalyzeDocument\n''\tDetectDocumentText\n''\tGetDocumentAnalysis...'
    },
    'Type': 'DOCUMENT'
}
```

Pregunta y respuesta

Se devuelve una respuesta de pregunta y respuesta cuando se hace Amazon Kendra coincidir una pregunta con una de las preguntas frecuentes del índice. La respuesta incluye la pregunta y la respuesta coincidentes en el QueryResultItemcampo. También incluye información destacada sobre los términos de consulta detectados en la cadena de la consulta. El siguiente ejemplo de JSON muestra una respuesta de pregunta y respuesta. Observe que la respuesta incluye el texto de la pregunta.

```
{
    'AnswerText': {
        'TextWithHighlights': [

        ],
        'Text': '605feet'
    },
    'DocumentExcerpt': {
        'Highlights': [
            {
                'BeginOffset': 0,
                'EndOffset': 8,
                'TopAnswer': False
```

```
            }
        ],
        'Text': '605feet'
    },
    'Type': 'QUESTION_ANSWER',
    'QuestionText': {
        'Highlights': [
            {
                'BeginOffset': 12,
                'EndOffset': 18,
                'TopAnswer': False
            },
            {
                'BeginOffset': 26,
                'EndOffset': 31,
                'TopAnswer': False
            },
            {
                'BeginOffset': 32,
                'EndOffset': 38,
                'TopAnswer': False
            }
        ],
        'Text': 'whatistheheightoftheSpaceNeedle?'
    }
}
```

Para obtener información sobre cómo añadir el texto de una pregunta y una respuesta a un índice, consulte Añadir preguntas frecuentes a un índice.

Ajustar y ordenar las respuestas

Puede modificar el efecto de un campo o atributo en la relevancia de las búsquedas ajustando su relevancia. También puede ordenar los resultados de búsqueda por un atributo o campo determinado.

Temas

- Ajustar las respuestas
- Ordenar las respuestas

Ajustar las respuestas

Puede modificar el efecto de un campo o atributo en la relevancia de las búsquedas ajustando su relevancia. Para probar rápidamente el ajuste de relevancia, use la API de consulta para aplicar las configuraciones de ajuste en la consulta. Así puede ver los distintos resultados de búsqueda que obtiene de las distintas configuraciones. La consola no permite ajustar la relevancia en la consulta. También puede ajustar los campos o atributos del tipo `StringList` solo en el índice. Para más información, consulte Ajuste de la relevancia de la búsqueda.

De forma predeterminada, las respuestas a las consultas se ordenan según la puntuación de relevancia que se Amazon Kendra determina para cada resultado de la respuesta.

Puede ajustar los resultados de cualquier atributo o campo integrado o personalizado de los siguientes tipos:

- Valor de fecha
- Valor largo
- Valor de cadena

No se pueden ordenar los atributos del siguiente tipo:

- Valores de listas de cadenas

Clasifique y ajuste los resultados de los documentos (AWS SDK)

Establezca el parámetro `Searchable` como true para priorizar la configuración de los metadatos del documento.

Para ajustar un atributo en una consulta, defina el parámetro `DocumentRelevanceOverrideConfigurations` de la API `Query` y especifique el nombre del atributo que desea ajustar.

En el siguiente ejemplo de JSON, se muestra un objeto `DocumentRelevanceOverrideConfigurations` que anula el ajuste del atributo denominado “department” en el índice.

```
"DocumentRelevanceOverrideConfigurations" : [
    "Name": "department",
```

```
    "Relevance": {
        "Importance": 1,
        "ValueImportanceMap": {
            "IT": 3,
            "HR": 7
        }
    }
]
```

Ordenar las respuestas

Amazon Kendra utiliza el campo o atributo de ordenación como parte de los criterios de los documentos devueltos por la consulta. Por ejemplo, es posible que los resultados devueltos por una consulta ordenada por “_created_at” no contengan los mismos resultados que una consulta ordenada por “_version”.

De forma predeterminada, las respuestas a las consultas se ordenan según la puntuación de relevancia que se Amazon Kendra determina para cada resultado de la respuesta. Para cambiar el orden de clasificación, haga que un atributo del documento se pueda ordenar y, a continuación, Amazon Kendra configúrelo para usar ese atributo para ordenar las respuestas.

Puede ordenar los resultados por cualquier atributo o campo integrado o personalizado de los siguientes tipos:

- Valor de fecha
- Valor largo
- Valor de cadena

No se pueden ordenar los atributos del siguiente tipo:

- Valores de listas de cadenas

Puede ordenar los resultados por uno o más atributos del documento en cada consulta. Las consultas devuelven 100 resultados. Si hay menos de 100 documentos con el atributo de ordenación establecido, los documentos sin un valor para el atributo de ordenación se devuelven al final de los resultados, ordenados por su relevancia para la consulta.

Para ordenar los resultados de los documentos (AWS SDK)

1. Para usar la UpdateIndexAPI para ordenar un atributo, defina `true` el `Sortable` parámetro en. En el siguiente ejemplo de JSON se utiliza `DocumentMetadataConfigurationUpdates` para añadir al índice un atributo denominado “Department” y hacer que se pueda ordenar.

```
"DocumentMetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE",
       "Search": {
           "Sortable": "true"
       }
   }
]
```

2. Para usar un atributo que se puede ordenar en una consulta, defina el parámetro `SortingConfiguration` de la API de consulta. Especifique el nombre del atributo según el cual se van a ordenar los resultados y si se deben ordenar en la respuesta en orden ascendente o descendente.

   El siguiente ejemplo de JSON muestra el parámetro `SortingConfiguration` que se usa para ordenar los resultados de una consulta por el atributo “Department” en orden ascendente.

```
   "SortingConfiguration": {
      "DocumentAttributeKey": "Department",
      "SortOrder": "ASC"
   }
```

3. Para usar más de un atributo que se puede ordenar en una consulta, defina el parámetro `SortingConfigurations` de la API de consulta. Puede configurar hasta 3 campos para que Amazon Kendra los use para ordenar los resultados. También puede especificar si los resultados se deben ordenar en orden ascendente o descendente. Puede aumentar la cuota de campos de ordenación.

   Si no proporciona una configuración de clasificación, los resultados se ordenan según la relevancia que Amazon Kendra determine el resultado. En caso de empate en la ordenación de los resultados, estos se ordenan por relevancia.

El siguiente ejemplo de JSON muestra el parámetro `SortingConfigurations` que se usa para ordenar los resultados de una consulta por los atributos “Name” y “Price” en orden ascendente.

```
"CollapseConfiguration" : {
    "DocumentAttributeKey": "Name",
    "SortingConfigurations": [
        {
            "DocumentAttributeKey": "Price",
            "SortOrder": "ASC"
        }
    ],
    "MissingAttributeKeyStrategy": "IGNORE"
}
```

Para ordenar los resultados de los documentos (consola)

> **Note**
>
> La ordenación por varios atributos no es compatible actualmente con la AWS Management Console.

1. Para hacer que un atributo se pueda ordenar en la consola, elija Ordenable en la definición del atributo. Puede hacer que un atributo se pueda ordenar al crearlo o bien modificarlo más adelante.
2. Para ordenar la respuesta a una consulta en la consola, elija el atributo para ordenar la respuesta en el menú Ordenar. En la lista solo aparecen los atributos que se marcaron como ordenables durante la configuración del origen de datos.

Contraer o expandir los resultados de la consulta

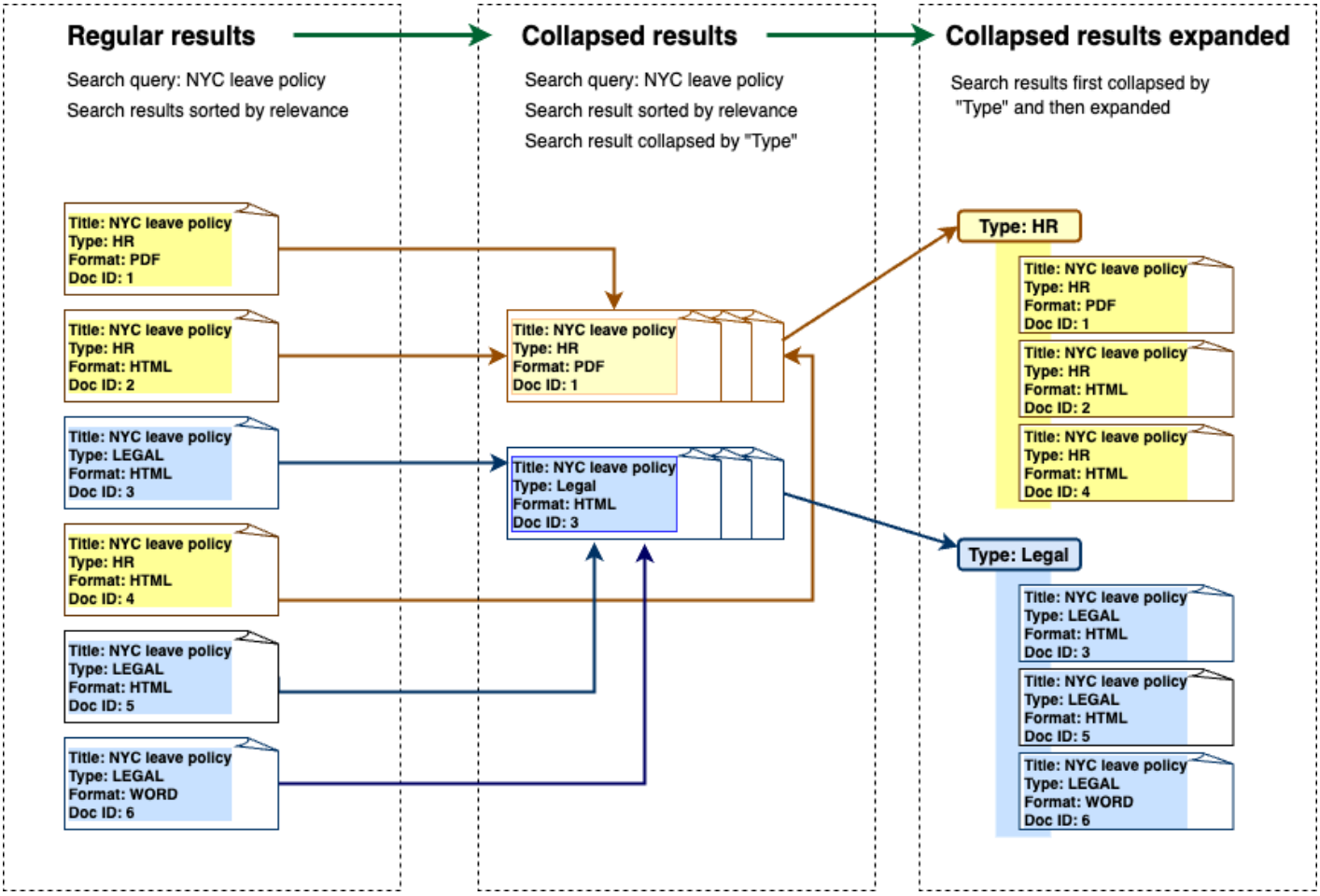
Cuando se conecta Amazon Kendra a sus datos, rastrea los atributos de los metadatos del documento (como `_document_title_created_at`, y) y `_document_id` utiliza estos atributos o campos para proporcionar funciones de búsqueda avanzada durante el tiempo de consulta.

La característica de contraer y expandir los resultados de las consultas de Amazon Kendra le permite agrupar los resultados de búsqueda mediante un atributo de documento común y mostrarlos, contraídos o parcialmente expandidos, en un documento principal designado.

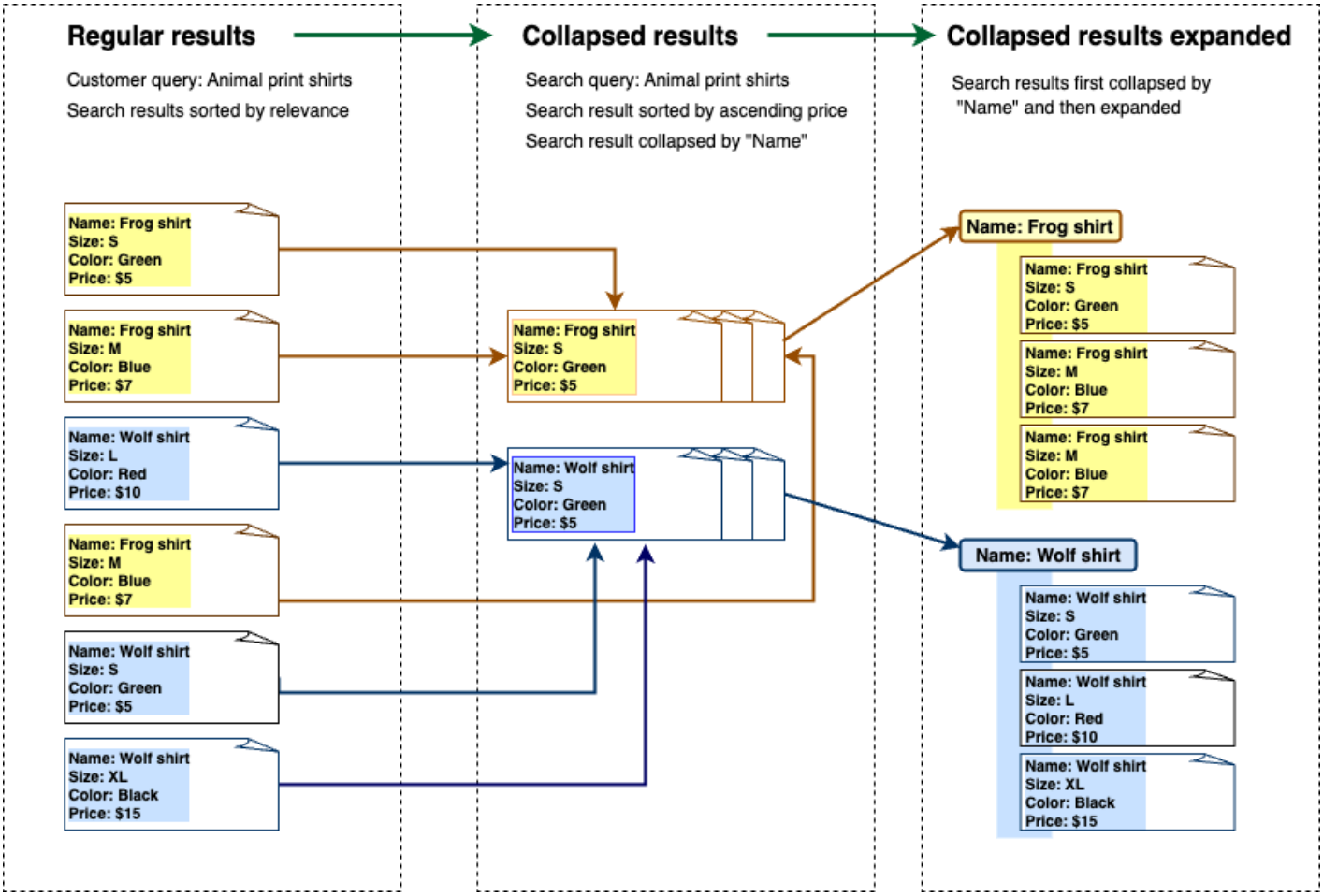
> **Note**
>
> La característica de contraer y expandir los resultados de las consultas solo está disponible actualmente a través de la API Amazon Kendra.

Esto resulta útil en las siguientes situaciones de búsqueda:

- Cuando existen varias versiones del contenido en los documentos de su índice. Cuando su usuario final consulte el índice, su intención es que vea la versión más relevante del documento con los duplicados ocultos o contraídos. Por ejemplo, si su índice contiene varias versiones de un documento denominado «Política de permisos de la ciudad de Nueva York», puede optar por agrupar los documentos para los grupos específicos «Recursos humanos» y «Legal» mediante el atributo/campo «Tipo».



- Su índice contiene varios documentos con información única sobre un tipo de elemento u objeto, como el inventario de un producto, por ejemplo. Para capturar y ordenar la información de los elementos de forma cómoda, su intención es que los usuarios finales accedan a todos los documentos enlazados por un elemento u objeto como un resultado de búsqueda. En el siguiente ejemplo, si un cliente busca «camisas con estampado animal», se muestran los resultados agrupados por nombre y ordenados por orden de precios ascendente.



Contraer los resultados

Para agrupar documentos similares o relacionados, debes especificar el atributo por el que deseas comprimirlos (por ejemplo, puedes comprimirlos o agruparlos). `_category` Para ello, llama a la API de consultas y usa el CollapseConfigurationobjeto para especificar el objeto sobre el que se va `DocumentAttributeKey` a contraer. La `DocumentAttributeKey` controla el campo por el que se contraerán los resultados de búsqueda. Los campos clave de atributos admitidos incluyen `String` y `Number`. Los tipos `String list` y `Date` tipo no son compatibles.

Elegir un documento principal mediante la ordenación

Para configurar el documento principal para que se muestre en un grupo contraído, utilice el `SortingConfigurations` parámetro siguiente CollapseConfiguration. Por ejemplo, para obtener la versión más reciente de un documento, ordenaría cada grupo contraído por`_version`. Puede especificar hasta 3 atributos o campos por los que ordenar y una ordenación para cada atributo o

campo utilizando `SortingConfigurations`. Puede solicitar un aumento de cuota para el número de atributos de ordenación.

De forma predeterminada, Amazon Kendra ordena las respuestas a las consultas según la puntuación de relevancia que determina para cada resultado de la respuesta. Para cambiar el orden de clasificación predeterminado, haga que los atributos del documento se puedan ordenar y, a continuación, Amazon Kendra configúrelos para usar estos atributos para ordenar las respuestas. Para obtener más información, consulte Ordenar las respuestas.

No hay una estrategia clave para el documento

Si el documento no tiene un valor de atributo de contracción, Amazon Kendra ofrece tres opciones de personalización:

- Seleccione que se `COLLAPSE` todos los documentos con valores nulos o sin valores en un grupo. Esta configuración es la predeterminada.
- Seleccione que se `IGNORE` los documentos con valores nulos o sin valores. Los documentos ignorados no aparecerán en los resultados de la consulta.
- Seleccione que se `EXPAND` cada documento con un valor nulo o sin valor en su propio grupo.

Expandir los resultados

Puede elegir si los grupos de resultados de búsqueda contraídos se expanden utilizando el `Expand` parámetro del CollapseConfigurationobjeto. Los resultados expandidos mantienen el mismo orden que se ha utilizado para seleccionar el documento principal del grupo.

Para configurar el número de grupos de resultados de búsqueda contraídos que se van a expandir, utilice el `MaxResultItemstoExpand` parámetro del ExpandConfigurationobjeto. Si establece este valor como 10, por ejemplo, solo los primeros 10 de los 100 grupos de resultados tendrán la funcionalidad de expansión.

Para configurar el número de resultados expandidos que se mostrarán por documento principal contraído, utilice el parámetro `MaxExpandResultsPerItem`. Por ejemplo, si establece este valor como 3, se mostrarán como máximo 3 resultados por grupo contraído.

Interacciones con otras Amazon Kendra funciones

- Al contraer y expandir los resultados no se modifica el número de facetas ni se ve afectado el número total de resultados mostrados.

- Los resultados de búsqueda destacados de Amazon Kendra no se contraerán aunque tengan el mismo valor de campo que el campo para contraer que haya configurado.
- La contracción y la expansión de los resultados solo se aplican a los resultados de tipo`DOCUMENT`.

Ajuste de la relevancia de las búsquedas

Amazon Kendra las consultas producen resultados de búsqueda clasificados por su relevancia. Todos los campos o atributos del índice en los que se pueden realizar búsquedas tienen efecto en esta clasificación.

Puede modificar el efecto de un campo o atributo en la relevancia de las búsquedas ajustando su relevancia. Puede ajustar la relevancia de las búsquedas manualmente a nivel de índice, donde se establecen las configuraciones de ajuste para el índice, o a nivel de consulta, anulando las configuraciones establecidas a nivel de índice.

Cuando se ajusta la relevancia, se da prioridad a los resultados en la respuesta cuando la consulta incluye términos que coinciden con el campo o atributo. También se especifica el nivel de prioridad que se da al documento cuando hay una coincidencia. El ajuste de relevancia no implica incluir un documento en la respuesta Amazon Kendra a la consulta, sino que es solo uno de los factores que se Amazon Kendra utilizan para determinar la relevancia de un documento.

Puede dar prioridad a campos o atributos específicos en su índice para asignar más importancia a respuestas específicas. Por ejemplo, cuando alguien busca “¿Cuándo tiene lugar re:Invent?” podría aumentar la relevancia de la frescura de los documentos `_last_update_at` sobre el terreno. O bien, en un índice de informes de investigación, puede priorizar un origen de datos específico en el campo “source”.

También puede priorizar los documentos en función de los votos o el número de visualizaciones, algo habitual en los foros y otras bases de conocimientos de soporte. Puede combinar prioridades, por ejemplo, para priorizar los documentos con más visualizaciones y los más recientes.

El nivel de prioridad que se da un documento se establece mediante el parámetro `Importance`. Cuanto mayor sea el parámetro `Importance`, mayor relevancia le dará el campo o atributo al documento. Al ajustar el índice o al ajustar a nivel de consulta, aumente el valor del parámetro `Importance` en pequeños incrementos hasta obtener el efecto deseado. Para determinar si está mejorando los resultados de búsqueda, realice la búsqueda y compare los resultados con consultas anteriores.

Puede especificar los atributos de fecha, número o cadena para ajustar un índice o ajustar a nivel de consulta. Solo puede ajustar a nivel de índice los campos o atributos del tipo `StringList`. Cada campo o atributo tiene criterios específicos para determinar cuándo prioriza un resultado.

- Campos o atributos de fecha: hay tres criterios específicos para los campos de fecha, que son `Duration`, `Freshness` y `RankOrder`.
  - `Duration` especifica el periodo de tiempo al que se aplica la prioridad. Por ejemplo, si establece el período de tiempo en 86 400 segundos (es decir, un día), la prioridad comienza a disminuir al cabo de un día. Cuanto mayor sea la importancia, más rápido se reduce el efecto de la prioridad.
  - `Freshness` determina qué tan reciente es un documento cuando se aplica a un campo o atributo. Si aplica `Freshness` al campo de la fecha de creación o de la fecha de la última actualización, un documento creado o actualizado más recientemente se considera “más reciente” que otro documento anterior. Por ejemplo, si el documento 1 se creó el 14 de noviembre y el documento 2 se creó el 5 de noviembre, el documento 1 es “más reciente” que el documento 2. Pero si el documento 1 se actualizó por última vez el 14 de noviembre y el documento 2 se actualizó por última vez el 20 de noviembre, el documento 2 es “más reciente” que el documento 1. Cuanto más reciente sea el documento, más prioridad se le dará. Solo puede tener un campo `Freshness` en el índice.
  - `RankOrder` puede dar prioridad en orden ascendente o descendente. Si especifica `ASCENDING`, las fechas posteriores tienen prioridad. Si especifica `DESCENDING`, las fechas anteriores tienen prioridad.
- Campos o atributos numéricos: en el caso de los campos o atributos numéricos, puede especificar el orden de clasificación que Amazon Kendra debe utilizarse para determinar la relevancia del campo o atributo. Si especifica `ASCENDING`, se da prioridad a los números más altos. Si especifica `DESCENDING`, los números más bajos tienen prioridad.
- Campos o atributos de cadena: en el caso de los campos o atributos de cadena, puede crear categorías de un campo para dar a cada categoría una prioridad diferente. Por ejemplo, si prioriza un campo o atributo llamado “Department”, puede dar una prioridad diferente a los documentos de “HR” que a los documentos de “Legal”. Puede priorizar un campo o atributo del tipo `String`. Solo puede priorizar los campos `StringList` a nivel de índice.

Ajuste de la relevancia a nivel de índice

Para ajustar la relevancia de un campo o atributo a nivel de índice, utilice la consola para configurar los detalles del índice o la UpdateIndexAPI.

En el siguiente ejemplo, se establece el `_last_updated_at` campo como `Freshness` campo de un documento.

```
"DocumentMetadataConfigurationUpdates" : [
    {
        "Name": "_last_updated_at",
        "Type": "DATE_VALUE",
        "Relevance": {
            "Freshness": TRUE,
            "Importance": 2
        }
    }
]
```

En el siguiente ejemplo se aplica una importancia diferente a las distintas categorías del campo “department”.

```
"DocumentMetadataConfigurationUpdates" : [
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 3,
                "Legal": 1
            }
        }
    }
]
```

Ajuste de la relevancia a nivel de consulta

Puede ajustar la relevancia de un campo o atributo a nivel de consulta mediante la API Query.

La consola no permite ajustar la relevancia a nivel de consulta.

Ajustar a nivel de consulta puede acelerar el proceso de comprobación del ajuste de la relevancia, ya que no es necesario actualizar manualmente las configuraciones de ajuste del índice para cada prueba. Puede ajustar la relevancia de un documento introduciendo configuraciones de ajuste en la consulta. Así puede ver los distintos resultados que obtiene de las distintas configuraciones. La configuración introducida en la consulta anula la configuración establecida a nivel de índice.

El siguiente ejemplo anula la importancia que se da al campo “department” y a cada categoría de departamento establecida a nivel de índice del ejemplo anterior. Cuando un usuario introduce su consulta de búsqueda, el campo “department” tiene un nivel de importancia razonable y el departamento “Legal” tiene más importancia que el departamento “HR”.

```
"DocumentRelevanceOverrideConfigurations" : [
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 2,
                "Legal": 8
            }
        }
    }
]
```

Obtener información con análisis de búsqueda

Puedes usar Amazon Kendra Search Analytics para obtener información sobre cómo tu aplicación de búsqueda está ayudando a los usuarios a encontrar información con éxito o sin éxito.

Amazon Kendra Los análisis proporcionan una visión general de la forma en que los usuarios interactúan con la aplicación de búsqueda y de la eficacia de la configuración de la aplicación de búsqueda. Puede ver los datos de las métricas mediante la GetSnapshotsAPI o seleccionando Analytics en el panel de navegación de la consola.

Puede renderizar los datos generados por `GetSnapshots` en su propio panel de control personalizado. O puede utilizar el panel de métricas proporcionado en la consola, que incluye gráficos visuales. Con un panel visual, puedes buscar tendencias o patrones en el comportamiento de los usuarios a lo largo del tiempo o detectar problemas con la configuración de tu aplicación de búsqueda. Por ejemplo, un gráfico lineal que muestre un número constante de consultas por día y un aumento constante podría indicar un aumento de la adopción y el uso. Por otro lado, una caída abrupta podría indicar que hay un problema que debe investigarse.

Puede utilizar las métricas para establecer conexiones entre distintos puntos de datos y resolver problemas relacionados con la forma en que sus usuarios buscan información o descubren oportunidades de negocio. Por ejemplo, el documento “¿Cómo funciona la IA?” es el documento en el que más se ha hecho clic en los resultados de búsqueda y la consulta más buscada es “¿Cómo funciona el machine learning?”. Esto le informa sobre los términos y el idioma preferidos que utilizan sus usuarios. Puede integrar estos términos en sus documentos o utilizar sinónimos personalizados para estos términos para que los usuarios puedan realizar búsquedas en sus documentos con mayor facilidad.

Métricas de búsqueda

Hay 10 métricas para analizar el rendimiento de la aplicación de búsqueda o la información que buscan los usuarios. Para recuperar los datos de las métricas, especifique el nombre de cadena de los datos de las métricas que desee recuperar cuando llame a `GetSnapshots`.

También debe proporcionar un intervalo de tiempo o una ventana de tiempo para ver los datos de las métricas. El intervalo de tiempo utiliza la zona horaria de su índice. Puede ver los datos en las siguientes ventanas de tiempo:

- `THIS_WEEK`: La semana actual, que comienza el domingo y termina el día anterior a la fecha actual.
- `ONE_WEEK_AGO`: La semana anterior, comienza el domingo y termina el sábado siguiente.
- `TWO_WEEKS_AGO`: La semana anterior a la anterior, comenzando el domingo y terminando el sábado siguiente.
- `THIS_MONTH`: El mes en curso, comenzando el primer día del mes y terminando el día anterior a la fecha actual.
- `ONE_MONTH_AGO`: El mes anterior, comienza el primer día del mes y termina el último día del mes.
- `TWO_MONTHS_AGO`: El mes anterior al mes anterior, comenzando el primer día del mes y terminando el último día del mes.

En la consola, las ventanas horarias admitidas son Esta semana, Semana anterior, Este mes, Mes anterior.

Tasa de clics

La proporción de consultas que llevan a que se haga clic en un documento en los resultados de la búsqueda. Esto le ayuda a comprender si la configuración de la aplicación de búsqueda ayuda a los usuarios a encontrar información relevante para sus consultas. En el caso de las consultas que devuelven respuestas instantáneas, es posible que los usuarios no necesiten hacer clic en un documento para obtener más información. Para obtener más información, consulte the section called “Tasa de respuesta instantánea”. Debe llamar SubmitFeedbackpara asegurarse de que se recopilan los comentarios sobre los clics.

Para recuperar datos sobre el porcentaje de clics mediante la API `GetSnapshots`, especifique el `metricType` como `AGG_QUERY_DOC_METRICS`. También puede ver esta métrica en la consola seleccionando Análisis en el panel de navegación.

Tasa de clics cero

Proporción de consultas que dan lugar a cero clics en los resultados de búsqueda. Esto le ayuda a comprender las lagunas de su contenido que proporcionan resultados de búsqueda irrelevantes. En el caso de las consultas que devuelven respuestas instantáneas, es posible que los usuarios no necesiten hacer clic en un documento para obtener más información. Para obtener más información, consulte the section called “Tasa de respuesta instantánea”. Además, la configuración de búsqueda,

como el ajuste de las configuraciones, podría afectar a la forma en que se muestran los documentos en los resultados de la búsqueda.

Para recuperar datos sobre la tasa de clics cero mediante la API `GetSnapshots`, especifique el `metricType` como `AGG_QUERY_DOC_METRICS`. También puede ver esta métrica en la consola seleccionando Análisis en el panel de navegación.

Tasa de resultados de búsqueda cero

La proporción de consultas que conducen a cero resultados de búsqueda. Esto le ayuda a comprender las lagunas de su contenido que no proporcionan resultados de búsqueda relevantes.

Para recuperar datos sobre la tasa cero de resultados de búsqueda utilizando la API `GetSnapshots`, especifique el `metricType` como `AGG_QUERY_DOC_METRICS`. También puede ver esta métrica en la consola seleccionando Análisis en el panel de navegación.

Tasa de respuesta instantánea

La proporción de consultas con una respuesta instantánea o una pregunta frecuente devueltas. Esto le ayuda a comprender el papel de las respuestas instantáneas a la hora de proporcionar información.

Para recuperar datos sobre la tasa de respuesta instantánea mediante la API `GetSnapshots`, especifique el `metricType` como `AGG_QUERY_DOC_METRICS`. También puede ver esta métrica en la consola seleccionando Análisis en el panel de navegación.

Consultas principales

Las 100 consultas más buscadas por los usuarios. Esto le ayuda a entender qué consultas son populares y qué tipo de información les interesa más a sus usuarios.

Las métricas incluyen el número de veces que se busca en la consulta, la proporción de clics en un documento, la proporción de no clics en un documento, la profundidad media de clics en los resultados de la búsqueda de la consulta, la proporción de respuestas instantáneas para la consulta y la confianza media de los primeros 10 resultados de búsqueda de una consulta.

Para recuperar datos sobre las consultas principales mediante la API `GetSnapshots`, especifique el `metricType` como `QUERIES_BY_COUNT`. También puede ver esta métrica en la consola si selecciona Análisis en el panel de navegación de la consola y, a continuación, selecciona Consultas principales en las listas de consultas.

Consultas principales con cero clics

Las 100 consultas principales que no conducen a ningún clic en los resultados de búsqueda. Esto le ayuda a comprender cualquier laguna en el contenido, ya sea que falten documentos relevantes para algunas consultas o que la configuración de la aplicación de búsqueda devuelva resultados de búsqueda irrelevantes. En el caso de las consultas que devuelven respuestas instantáneas, es posible que los usuarios no necesiten hacer clic en un documento para obtener más información. Para obtener más información, consulte the section called “Tasa de respuesta instantánea”.

Las métricas incluyen el número de veces que la consulta genera cero clics, la proporción de cero clics para la consulta, la proporción de respuestas instantáneas para la consulta y la confianza media de los primeros 10 resultados de búsqueda de una consulta.

Para recuperar datos sobre las consultas con cero clics mediante la API `GetSnapshots`, especifique el `metricType` como `QUERIES_BY_ZERO_CLICK_RATE`. También puede ver esta métrica en la consola si selecciona Análisis en el panel de navegación de la consola y, a continuación, selecciona Consultas con cero clics en las listas de consultas.

Consultas principales con cero resultados de búsqueda

Las 100 consultas principales que conducen a cero resultados de búsqueda. Esto le ayuda a comprender las lagunas en su contenido, cuando no hay documentos relevantes para algunas consultas. O bien, los usuarios pueden realizar consultas con términos especializados que, posiblemente, no arrojen resultados de búsqueda, lo que te pedirá que crees sinónimos personalizados para solucionar este problema.

Las métricas incluyen el número de veces que la consulta arroja cero resultados de búsqueda, la proporción de resultados de búsqueda cero para la consulta y la proporción de veces que se busca en la consulta en comparación con todas las consultas.

Para recuperar datos sobre las consultas principales con cero resultados de búsqueda mediante la API `GetSnapshots`, especifique el `metricType` como `QUERIES_BY_ZERO_RESULT_RATE`. También puede ver esta métrica en la consola si selecciona Análisis en el panel de navegación de la consola y, a continuación, selecciona Consultas con cero resultados en las listas de consultas.

La mayoría de las veces has hecho clic en los documentos

Los 100 documentos en los que más se ha hecho clic en los resultados de búsqueda. Esto le ayuda a entender qué documentos o resultados de búsqueda son más relevantes para sus usuarios cuando buscan información.

Las métricas incluyen el número de veces que se hace clic en el documento, el número de “me gusta” que recibe un documento por parte de los usuarios (con el visto bueno hacia arriba) y el número de “no me gusta” que recibe un documento de los usuarios (con el visto bueno hacia abajo).

Para recuperar los datos de los documentos sobre los que se ha hecho clic mediante la API `GetSnapshots`, especifique el `metricType` como `DOCS_BY_CLICK_COUNT`. También puede ver esta métrica en la consola seleccionando Análisis en el panel de navegación de la consola y, a continuación, seleccionando los documentos en los que se ha hecho más clic en las listas de consultas.

Consultas totales

El número total de consultas realizadas por sus usuarios. Esto le ayuda a comprender el grado de interacción de sus usuarios con la aplicación de búsqueda.

Para recuperar datos sobre las consultas totales mediante la API `GetSnapshots`, especifique el `metricType` como `AGG_QUERY_DOC_METRICS`. También puede ver esta métrica en la consola seleccionando Análisis en el panel de navegación.

Documentos totales

El número total de documentos de su índice. Esto le ayuda a comparar el tamaño del índice con el número total de consultas para comprobar si hay un número adecuado de documentos para el volumen de consultas.

Para recuperar datos sobre el total de documentos mediante la API `GetSnapshots`, especifique el `metricType` como `AGG_QUERY_DOC_METRICS`. También puede ver esta métrica en la consola seleccionando Análisis en el panel de navegación.

Ejemplo de recuperación de datos métricos

El siguiente código es un ejemplo de cómo recuperar datos de las consultas principales del mes anterior.

Console

Para recuperar las consultas principales del mes anterior

1. En el panel de navegación izquierdo, en Índices, seleccione su índice y, a continuación, seleccione Análisis.

2. En la página de Análisis, seleccione el botón Esta semana para cambiar el período de recuperación de los datos a Mes anterior.
3. En la página de Análisis, en Listas de consultas, seleccione Consultas principales.

CLI

Para recuperar las consultas principales del mes anterior

```
aws kendra get-snapshots \
--index-id index-id \
--interval "ONE_MONTH_AGO" \
--metric-type "QUERIES_BY_COUNT"
```

Python

Para recuperar las consultas principales del mes anterior

```
import boto3

kendra = boto3.client("kendra")

index_id = "index-id"
interval = "ONE_MONTH_AGO"
metric_type = "QUERIES_BY_COUNT"

snapshots_response = kendra.get_snapshots(
    IndexId = index_id,
    Interval = interval,
    MetricType = metric_type
)

print("Top queries data: " + snapshots_response["snapshotsData"])
```

Java

Para recuperar las consultas principales del mes anterior

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.GetSnapshotsRequest;
import software.amazon.awssdk.services.kendra.model.GetSnapshotsResponse;
```

```
public class TopQueriesExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indexId = "indexID";
        String interval = "ONE_MONTH_AGO";
        String metricType = "QUERIES_BY_COUNT";

        GetSnapshotsRequest getSnapshotsRequest = GetSnapshotsRequest
            .builder()
            .indexId(indexId)
            .interval(interval)
            .metricType(metricType)
            .build();

        GetSnapshotsResponse getSnapshotsResponse =
 kendra.getSnapshots(GetSnapshotsRequest);

        System.out.println(String.format("Top queries data: ",
 getSnapshotsResponse.snapshotsData()))
```

Desde métricas hasta información procesable

La información procesable es información significativa que se extrae de datos sin procesar y se utiliza para guiar sus acciones o decisiones. Para extraer significado de las métricas y utilizarlas para obtener información procesable, es importante no solo analizarlas de forma aislada, sino también establecer conexiones entre ellas.

Por ejemplo, la consulta principal con cero clics es “¿Qué regiones están disponibles actualmente?”. Sin embargo, también tiene una tasa de respuesta instantánea del 100 por ciento. Esto sugiere que los usuarios reciban la respuesta a esta pregunta sin necesidad de hacer clic en un resultado de búsqueda o en un documento que proporcione información sobre las regiones disponibles. Si se fijara únicamente en los clics nulos, no obtendría la historia completa y posiblemente sacaría conclusiones erróneas sobre el éxito de la configuración de su aplicación de búsqueda a la hora de gestionar esta consulta.

Otro ejemplo de información útil es descubrir una oportunidad de negocio. Las empresas suelen buscar oportunidades para aumentar sus clientes mediante el análisis de las métricas de búsqueda. El documento en el que más se hace clic es “Regiones disponibles”. Además, la mayoría de las

consultas más buscadas están relacionadas con preguntas sobre la disponibilidad de productos en la región oceánica, con una tasa de respuesta instantánea del 100 por ciento y una alta tasa de clics para obtener más información sobre las regiones disponibles como parte de la respuesta. Esto sugiere que hay interés y demanda por tu producto o servicio en esta región.

Visualización y generación de informes sobre los análisis de búsqueda

Hay cinco métricas que incluyen datos de tendencias para que puedas visualizar y buscar tendencias o patrones a lo largo del tiempo. Si utiliza la consola, se proporcionan gráficos de los datos de tendencias. Si utiliza las API, podrá recuperar los datos de tendencias para crear sus propios gráficos o visualizaciones. La mayoría de los gráficos de la consola trazan los puntos de datos diarios a lo largo de la ventana temporal elegida.

La consola proporciona un panel de control de las métricas en el que puede seleccionar el gráfico y la lista principal que le interese ver. Puede exportar las métricas que se muestran en su panel de control en formato CSV seleccionando Exportar en la página de inicio de Análisis. Puede incluir estos informes en sus documentos o presentaciones empresariales.

Puede ver las siguientes métricas:

Gráfico de consultas totales

Un gráfico lineal del número de consultas emitidas por día. El gráfico le ayuda a visualizar los patrones de participación diaria de los usuarios. Algunos ejemplos incluyen un aumento o disminución constante de la participación de los usuarios, o una caída drástica a 0 consultas debido a un bloqueo de la aplicación de búsqueda o a problemas con el sitio web.

Si utiliza la API, puedes recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede usar los datos para crear sus propios gráficos o usar los gráficos que se proporcionan en la consola.

Gráfico de tasa de clics

Un gráfico lineal de las proporciones de clics por día. El gráfico le ayuda a visualizar los patrones de la tasa de clics diaria. Algunos ejemplos incluyen un aumento o disminución constante de la tasa de clics, o una disminución de las respuestas instantáneas, lo que podría influir en un aumento de los clics.

Si utiliza la API, puedes recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede usar los datos para crear sus propios gráficos o usar los gráficos que se proporcionan en la consola.

Gráfico de tasa de clics cero

Un gráfico lineal de la proporción de cero clics por día. El gráfico le ayuda a visualizar patrones en la tasa diaria de clics cero. Algunos ejemplos incluyen un aumento o disminución constante de la tasa de clics nulos, o un aumento de las respuestas instantáneas que posiblemente influya en un aumento de los clics nulos.

Si utiliza la API, puedes recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede usar los datos para crear sus propios gráficos o usar los gráficos que se proporcionan en la consola.

Gráfico de tasa de resultados de búsqueda cero

Un gráfico lineal de la proporción de resultados de búsqueda nulos por día. El gráfico le ayuda a visualizar patrones en la tasa diaria de resultados de búsqueda cero. Algunos ejemplos incluyen un aumento o una disminución constantes de la tasa de resultados de búsqueda cero, o una disminución pronunciada en el número de documentos del índice, lo que podría influir en un aumento de los resultados de búsqueda nulos.

Si utiliza la API, puedes recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede usar los datos para crear sus propios gráficos o usar los gráficos que se proporcionan en la consola.

Gráfico de tasa de respuesta instantánea

Un gráfico lineal de la proporción de consultas con respuesta instantánea o FAQ devueltas. El gráfico le ayuda a visualizar patrones en la tasa diaria de respuestas instantáneas. Algunos ejemplos incluyen el aumento o la disminución constante de las consultas de tipo pregunta-respuesta, o una disminución de los clics que posiblemente influya en un aumento de las respuestas instantáneas.

Si utiliza la API, puedes recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede usar los datos para crear sus propios gráficos o usar los gráficos que se proporcionan en la consola.

Envío de valoraciones para el aprendizaje incremental

Amazon Kendra utiliza el aprendizaje incremental para mejorar los resultados de búsqueda. Mediante las valoraciones de las consultas, el aprendizaje incremental mejora los algoritmos de clasificación y optimiza los resultados de búsqueda para lograr una mayor precisión.

Por ejemplo, supongamos que los usuarios buscan la frase “beneficios de atención médica”. Si los usuarios eligen siempre el segundo resultado de la lista, con el tiempo, Amazon Kendra prioriza ese resultado como el primero. El aumento disminuye con el paso del tiempo, por lo que si los usuarios dejan de seleccionar un resultado, al Amazon Kendra final lo eliminan y muestran otro resultado más popular. Esto ayuda a Amazon Kendra priorizar los resultados en función de la relevancia, la edad y el contenido.

El aprendizaje incremental está activado para todos los índices y para todos los tipos de documentos admitidos.

Amazon Kendra comienza a aprender en cuanto comentas tu opinión, aunque pueden pasar más de 24 horas hasta que veas los resultados de la valoración. Amazon Kendra te ofrece tres métodos para enviar comentarios: la AWS consola, una JavaScript biblioteca que puedes incluir en la página de resultados de búsqueda y una API que puedes usar.

Amazon Kendra acepta dos tipos de comentarios de los usuarios:

- Clics: información sobre qué resultados de las consultas eligen los usuarios. Las valoraciones incluyen el ID del resultado y la marca de tiempo Unix de la fecha y la hora en que se eligió el resultado de búsqueda.

  Para enviar valoraciones mediante clics, su aplicación debe recopilar información sobre los clics de las actividades de sus usuarios y, a continuación, enviarla a Amazon Kendra. Puedes recopilar información sobre los clics con la consola, la JavaScript biblioteca y la Amazon Kendra API.
- Relevancia: información sobre la relevancia de un resultado de búsqueda, que suelen proporcionar los usuarios. Las valoraciones contienen el ID del resultado y un indicador de relevancia (`RELEVANT` o `NOT_RELEVANT`). El usuario determina la información sobre la relevancia.

  Para enviar valoraciones de la relevancia, su aplicación debe proporcionar un mecanismo de valoraciones que permita al usuario elegir la relevancia adecuada para el resultado de una consulta y, a continuación, enviar esa información a Amazon Kendra. Solo puedes recopilar información relevante con la consola y la Amazon Kendra API.

Las valoraciones se utilizan mientras el índice está activo. Las valoraciones solo afectan al índice al que se envían y no se pueden usar en varios índices ni para cuentas diferentes.

Debe proporcionar un contexto de usuario adicional cuando consulte su Amazon Kendra índice. Al proporcionar un contexto de usuario, Amazon Kendra es capaz de saber si los comentarios provienen de un solo usuario o de varios usuarios y ajustar los resultados de la búsqueda en consecuencia.

Al proporcionar un contexto de usuario, las valoraciones sobre la consulta se asocian al usuario específico proporcionado en el contexto. Si no especifica un contexto de usuario, puede proporcionar un ID de visitante que se utilice para agrupar y agregar consultas.

Si no proporciona un contexto de usuario o un ID de visitante, las valoraciones son anónimas y se agregan a otras valoraciones anónimas.

El siguiente código muestra cómo incluir un contexto de usuario como un token o como un ID de visitante.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserToken = {
        Token = "token"
    })

    OR

    response = kendra.query(
    QueryText = query,
    IndexId = index,
    VisitorId = "visitor-id")
```

En el caso de las aplicaciones web, puede utilizar cookies, ubicaciones o usuarios del navegador para generar un ID de visitante para cada usuario.

En el caso de las consultas head (puestos en cabeza de la lista), que son las que tienen el mayor volumen de aparición, si se proporcionan las valoraciones mediante clics se obtiene suficiente información para mejorar la precisión general. En el caso de las consultas tail (puestos en la cola de la lista), que son las menos frecuentes, expertos en la materia deben enviar valoraciones sobre la relevancia de las consultas para mejorar su precisión.

Además de la consola, puedes usar uno de estos dos métodos: una JavaScript biblioteca o la SubmitFeedbackAPI. Debe usar solamente un método para recopilar las valoraciones. Para obtener los mejores resultados, debe enviar sus valoraciones en un plazo de 24 horas desde la realización de la consulta.

Temas

- Usar la Amazon Kendra JavaScript biblioteca para enviar comentarios
- Uso de la Amazon Kendra API para enviar comentarios

Usar la Amazon Kendra JavaScript biblioteca para enviar comentarios

Amazon Kendra proporciona una JavaScript biblioteca que puedes usar para añadir comentarios sobre los clics a tu página de resultados de búsqueda. Para utilizar la biblioteca debe insertar una etiqueta de cadena en su código cliente que muestre el resultado de búsqueda y, a continuación, añadir información a cada uno de los enlaces de los documentos de su lista de resultados. Cuando un usuario elige un enlace para ver un documento, se envía la información sobre los clics a Amazon Kendra.

La biblioteca funciona con navegadores compatibles con la JavaScript versión ES6/ES2015.

Paso 1: Inserta una etiqueta de script en tu aplicación de búsqueda Amazon Kendra

En el código de cliente que muestra los resultados de la Amazon Kendra búsqueda, inserta una <script>etiqueta y añade una referencia a la JavaScript biblioteca:

```
<script>
 (function(w, d, s, c, g, n) {
   if(!w[n]) {
     w[n] = w[n] || function () {
           (w[n].q = w[n].q || []).push(arguments);
     }
     w[n].st = new Date().getTime();
     w[n].ep = g;
     var e = document.createElement(s),
         j = document.getElementsByTagName(s)[0];
```

```
    e.async = 1;
    e.src = c;
    e.type = 'module';
    j.parentNode.insertBefore(e, j);
  }
 })(window, document, 'script',
 'library download URL',
 'feedback endpoint',
 'kendraFeedback');
</script>
```

El script descarga la JavaScript biblioteca de forma asíncrona desde una CDN Amazon Kendra alojada e inicializa una variable global denominada así `kendraFeedback` que permite establecer parámetros opcionales.

Sustituya *la URL de descarga de la biblioteca* y el *punto final de comentarios* por un identificador de la siguiente tabla en función de la región que aloja su índice. Amazon Kendra

Región	Descargar URL	Punto de conexión de valoraciones
us-east-1	https://d2zm0lpns956f8.cloudfront.net/ksf-v1.js	https://ujxwp5s92h.execute-api.us-east-1.amazonaws.com/prod/submit
us-east-2	https://d2crv7fufeg244.cloudfront.net/ksf-v1.js	https://i6h76zwzf3.execute-api.us-east-2.amazonaws.com/prod/submit
us-west-2	https://d2iezfpnpcoujy.cloudfront.net/ksf-v1.js	https://wg6nim909c.execute-api.us-west-2.amazonaws.com/prod/submit
ca-central-1	https://d1zbkfomowykaq.cloudfront.net/ksf-v1.js	https://budi8txevj.execute-api.ca-central-1.amazonaws.com/prod/submit
eu-west-1	https://d3gptlxtulu4us.cloudfront.net/ksf-v1.js	https://po2b11740b.execute-api.eu-west-1.amazonaws.com/prod/submit

Región	Descargar URL	Punto de conexión de valoraciones
ap-southeast-1	https://d1vvuam7g4taoe.cloudfront.net/ksf-v1	https://9je5uw7t5l.execute-api.ap-southeast-1.amazonaws.com/prod/submit
ap-southeast-2	https://dopqntoe6z0ce.cloudfront.net/ksf-v1.js	https://oovf4nvjj7.execute-api.ap-southeast-2.amazonaws.com/prod/submit
ap-south-1	https://d1ts9ouelsmk3g.cloudfront.net/ksf-v1.js	https://k1abnmd43b.execute-api.ap-south-1.amazonaws.com/prod/submit
ap-northeast-1	https://d3w0ybsa293kb4.cloudfront.net/ksf-v1.js	https://wg7rz0uzjh.execute-api.ap-northeast-1.amazonaws.com/prod/submit
eu-west-2	https://d1tsrujswld1d1.cloudfront.net/ksf-v1.js	https://qi7mct3x7f.execute-api.eu-west-2.amazonaws.com/prod/submit

Por ejemplo, si su índice está en Este de EE. UU. (Norte de Virginia), la *URL de descarga de biblioteca* es `https://d2zm0lpns956f8.cloudfront.net/ksf-v1.js` y el *punto de conexión de comentarios* es `https://ujxwp5s92h.execute-api.us-east-1.amazonaws.com/prod/submit`.

Hay dos ajustes opcionales que puede realizar para la Amazon Kendra JavaScript biblioteca:

- `disableCookies`— De forma predeterminada, Amazon Kendra establece una cookie que identifica de forma única al usuario. Configúrelo como `true` para deshabilitar la cookie.

  ```
  kendraFeedback('disableCookie', 'true | false');
  ```

  `searchDivClassName`: de forma predeterminada, Amazon Kendra supervisa los clics en todos los enlaces de su página de resultados de búsqueda. Configúrelo como un nombre de clase `<div>` para que los supervise solo en los enlaces de la clase especificada.

```
kendraFeedback('searchDivClassName', 'class name');
```

Paso 2: Añadir el token de valoración a los resultados de búsqueda

En su página de resultados, añada un atributo HTML llamado `data-kendra-token` a la etiqueta delimitadora o a la etiqueta div principal inmediata que contenga un enlace al documento desde la respuesta a la consulta. Por ejemplo:

```
<a href="document location" data-kendra-token="feedback token value"></a>
OR
<div data-url="document location" data-kendra-token="feedback token value"></div>
```

La respuesta a una consulta contiene un token en el campo `feedbackToken`. El token identifica la respuesta de forma única si el usuario la elige. Asigne el valor del token al atributo `data-kendra-token`. La Amazon Kendra JavaScript biblioteca busca este token cuando el usuario elige el resultado y lo envía a un Amazon Kendra punto final como comentario.

La Amazon Kendra JavaScript biblioteca solo envía el token de comentarios y otros metadatos, como la hora en que se eligió el resultado y un identificador de visitante único.

Paso 3: Probar la cadena de valoración

Para asegurarte de que la JavaScript biblioteca está configurada correctamente y de que envía los comentarios al punto final correcto, haz lo siguiente. En este ejemplo, se utiliza el navegador Chrome.

1. Abra las Herramientas para desarrolladores web en el navegador. En Chrome, abra el menú de Chrome en la esquina superior derecha del navegador, seleccione Más herramientas y, a continuación, Herramientas para desarrolladores.
2. Asegúrese de que no haya errores relacionados con la Amazon Kendra JavaScript biblioteca en la pestaña de la consola.
3. Realice una búsqueda y elija cualquier resultado. En el panel de Herramientas para desarrolladores, elija la pestaña Red. Debería ver una solicitud enviada al punto de conexión de valoración, el token del resultado y un estado 200 OK.

Uso de la Amazon Kendra API para enviar comentarios

Para usar la Amazon Kendra API para enviar comentarios sobre consultas, usa la SubmitFeedbackAPI. Para identificar la consulta, debes proporcionar el ID de índice del índice al que se aplica la consulta y el ID de consulta devuelto en la respuesta de la API de consultas.

En el siguiente ejemplo se muestra cómo enviar valoraciones mediante clics y sobre la relevancia con la API Amazon Kendra . Puede enviar varios conjuntos de valoraciones a través de las matrices `ClickFeedbackItems` y `RelevanceFeedbackItems`. En este ejemplo, se envía un solo elemento de valoración mediante clic y uno solo sobre la relevancia. El envío de valoraciones utiliza la hora actual.

Para enviar comentarios para una búsqueda (AWS SDK)

1. Puedes usar el siguiente código de ejemplo con los valores necesarios:

   a. `index id`: el ID del índice al que se aplica la consulta.

   b. `query id`—La consulta sobre la que quieres enviar comentarios.

   c. `result id`: el ID del resultado de la consulta sobre el que quieres enviar comentarios. La respuesta a la consulta contiene el ID del resultado.

   d. `relevance value`—Ya sea `RELEVANT` (el resultado de la consulta es relevante) o `NOT_RELEVANT` (el resultado de la consulta no es relevante).

   Python

```
import boto3
import time

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the query ID
query_id = "query-id"
# Provide the search result ID
result_id = "result-id"

# Configure the feedback item
feedback_item = {"ClickTime": int(time.time()),
```

```
    "ResultId":result_id}

# Configure the relevance value
relevance_value = "RELEVANT"
relevance_item = {"RelevanceValue": relevance_value,
    "ResultId": result_id
    }

response = kendra.submit_feedback(
    QueryId = query_id,
    IndexId = index_id,
    ClickFeedbackItems = [feedback_item],
    RelevanceFeedbackItems = [relevance_item]
)


print("Submitted feedback for query: " + query_id)
```

Java

```
package com.amazonaws.kendra;

import java.time.Instant;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.ClickFeedback;
import software.amazon.awssdk.services.kendra.model.RelevanceFeedback;
import software.amazon.awssdk.services.kendra.model.RelevanceType;
import software.amazon.awssdk.services.kendra.model.SubmitFeedbackRequest;
import software.amazon.awssdk.services.kendra.model.SubmitFeedbackResponse;

public class SubmitFeedbackExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        SubmitFeedbackRequest submitFeedbackRequest = SubmitFeedbackRequest
            .builder()
            .indexId("IndexId")
            .queryId("QueryId")
```

```
            .clickFeedbackItems(
                ClickFeedback
                .builder()
                .clickTime(Instant.now())
                .resultId("ResultId")
                .build())
            .relevanceFeedbackItems(
                RelevanceFeedback
                .builder()
                .relevanceValue(RelevanceType.RELEVANT)
                .resultId("ResultId")
                .build())
            .build();

        SubmitFeedbackResponse response =
 kendra.submitFeedback(submitFeedbackRequest);

        System.out.println("Feedback is submitted");
    }
}
```

2. Ejecute el código. Una vez enviada la valoración, el código muestra un mensaje.

Adición de sinónimos personalizados a un índice

Para añadir sinónimos personalizados a un índice, debe especificarlos en un archivo de tesauros. Puede incluir términos específicos de la empresa o especializados al Amazon Kendra usar sinónimos. Los sinónimos genéricos en inglés`leader, head`, como, están integrados Amazon Kendra y no deben incluirse en un archivo de tesauros, incluidos los sinónimos genéricos que utilizan guiones. Amazon Kendra admite sinónimos para todos los tipos de respuesta, incluidos los tipos de `DOCUMENT` respuesta `QUESTION_ANSWER` o `ANSWER` tipos de respuesta. Amazon Kendra actualmente no admite la adición de sinónimos marcados como palabras inútiles. Esto se incluirá en una versión futura.

Amazon Kendra establece correlaciones entre sinónimos. Por ejemplo, al usar el par de sinónimos`Dynamo, Amazon DynamoDB`, Amazon Kendra correlaciona Dynamo con. Amazon DynamoDB La pregunta “¿Qué es la dinamo?” a continuación, devuelve un documento como «¿Qué es Amazon DynamoDB?». Con los sinónimos, Amazon Kendra puede detectar la correlación más fácilmente.

El archivo de sinónimos es un archivo de texto almacenado en un Amazon S3 depósito. Consulte Adición de un tesauro a un índice.

El archivo de sinónimos utiliza el formato de sinónimos de Solr. Amazon Kendra tiene un límite en el número de tesauros por índice. Consulte las cuotas.

Los sinónimos pueden ser útiles en las siguientes situaciones:

- Términos especializados que no son sinónimos tradicionales en inglés, como `NLP, Natural Language Processing`.
- Sustantivos propios con asociaciones semánticas complejas. Estos son sustantivos que es poco probable que el público en general comprenda, por ejemplo, en el machine learning, `cost, loss, model performance`.
- Diferentes formas de nombres de productos, por ejemplo, `Elastic Compute Cloud, EC2`.
- Términos específicos de un dominio o de una empresa, como nombres de productos. Por ejemplo, `Route53, DNS`.

No utilice sinónimos en las siguientes situaciones:

- Sinónimos genéricos en inglés, como `leader, head`. Estos sinónimos no son específicos de un dominio y el uso de sinónimos en estos escenarios puede tener efectos no deseados.
- Errores tipográficos como `teh => the`.
- Variantes morfológicas como los plurales y posesivos de los sustantivos, la forma comparativa y superlativa de los adjetivos y el tiempo pasado, el participio pasado y la forma progresiva de los verbos. Un ejemplo de adjetivos comparativos y superlativos es `good, better, best`.
- Unigrama (palabra única): palabras que detienen palabras como `WHO`. Las palabras paralizantes de Unigram no están permitidas en el tesauro y se excluyen de la búsqueda. Por ejemplo, `WHO => World Health Organization` se rechaza. Sin embargo, puede usar `W.H.O.` como sinónimo y puede usar palabras vacías como parte de un sinónimo de varias palabras. Por ejemplo, `of` está permitido, pero `United States of America` no.

Los sinónimos personalizados facilitan la comprensión Amazon Kendra de la terminología específica de su empresa al ampliar las consultas para incluir los sinónimos específicos de su empresa. Si bien los sinónimos pueden mejorar la precisión de las búsquedas, es importante entender cómo afectan a la latencia para poder optimizarlos.

Una regla general para los sinónimos es: cuantos más términos de la consulta coincidan y se expandan con sinónimos, mayor será el impacto potencial en la latencia. Otros factores que afectan a la latencia son el tamaño medio de los documentos indexados, el tamaño del índice, los posibles filtros en los resultados de búsqueda y la carga total del índice. Amazon Kendra Las consultas que no coincidan con ningún sinónimo no se ven afectadas.

Una guía general sobre cómo los sinónimos afectan a la latencia:

Caso de uso	Aumento de la latencia*
Consultas típicas de lenguaje natural o palabras clave de 3 a 5 palabras cada una	Menos del 15 %
Un término de consulta se amplía a 3 sinónimos	
Índice de unos 500 000 documentos (con un promedio de 10,48 KB de texto extraído por documento) o 30 000 pares de preguntas frecuentes y preguntas	

* El rendimiento varía en función del uso específico de los sinónimos y las configuraciones del índice. Es mejor probar el rendimiento de las búsquedas para obtener puntos de referencia más precisos para tu caso de uso específico.

Si el tesauro es grande, tiene una tasa de expansión temporal alta y el aumento de la latencia no está dentro de los límites aceptables, puede probar una de las siguientes opciones o ambas:

- Recorte el tesauro para reducir la relación de expansión (número de sinónimos por término).
- Reduzca la cobertura general de los términos (número de líneas del tesauro).

Como alternativa, puede aumentar la capacidad de aprovisionamiento (unidades de almacenamiento virtuales) para compensar el aumento de la latencia.

Temas

- Crear un archivo de tesauro
- Adición de un tesauro a un índice
- Actualización de un tesauro
- Eliminar un tesauro
- Aspectos destacados en los resultados de búsqueda

Crear un archivo de tesauro

Un archivo de Amazon Kendra sinónimos es un archivo codificado en UTF-8 que contiene una lista de sinónimos en el formato de lista de sinónimos de Solr. El archivo.zip debe tener menos de 5 MB.

Hay dos formas de especificar las asignaciones de sinónimos:

- Los sinónimos bidireccionales se especifican como una lista de términos separados por comas. Si el usuario consulta alguno de los términos, se utilizarán todos los términos de la lista para buscar documentos, incluido el término original consultado.
- Los sinónimos unidireccionales se especifican como términos separados por el símbolo “**=>**” entre ellos para asignar los términos a sus sinónimos. Si el usuario consulta un término a la izquierda del símbolo “**=>**”, se asigna a un término de la derecha para buscar documentos utilizando el sinónimo. No se mapea al revés, por lo que es unidireccional.

Los sinónimos en sí distinguen mayúsculas de minúsculas, pero los términos a los que se asignan no distinguen mayúsculas de minúsculas. Por ejemplo, `ML => Machine Learning` significa que si su usuario consulta “ML” o “ml” o utiliza algún otro caso, se mapeará a "Machine Learning". Si tuviera que trazar este mapa a la inversa, `Machine Learning => ML`, entonces “Machine Learning” o “machine learning” o algún otro caso se asignaría a “ML”.

Un sinónimo no busca una coincidencia exacta en caracteres especiales. Por ejemplo, si busca dead-letter-queue ««, Amazon Kendra puede devolver documentos que coincidan con «cola de mensajes sin escribir» (sin guiones). Si los documentos contienen guiones, como dead-letter-queue ««, Amazon Kendra los procesa durante la búsqueda para eliminar los guiones. En el caso de los términos sinónimos genéricos en inglés que estén integrados en un archivo de tesauros Amazon Kendra y que no deban incluirse en él, Amazon Kendra puede buscar tanto en la versión del término con guión como en la versión sin guiones del término. Por ejemplo, si busca «tercero» y «tercero», obtendrá documentos que coincidan con cualquiera de Amazon Kendra las versiones de esos términos.

En el caso de los sinónimos que contienen palabras innecesarias o palabras de uso común, Amazon Kendra devuelve los documentos que coincidan con los términos, incluidas las palabras inútiles. Por ejemplo, puedes crear una regla de sinónimos para mapear las palabras «incorporación» e «incorporación». No puede utilizar únicamente palabras rápidas como sinónimos. Por ejemplo, si busca «activado», Amazon Kendra no podrá mostrar todos los documentos que contengan «activado».

Se ignoran algunas reglas de sinónimos. Por ejemplo, `a => b` es una regla, pero `a => a` se ignora y no cuenta como regla.

El número de términos es el número de términos únicos en el archivo de sinónimos. El siguiente archivo de ejemplo incluye los términos `AWS CodeStar` `MLMachine Learning`,`autoscaling group`,`ASG`, y más.

Hay una cantidad máxima de reglas de sinónimos por tesauro y una cantidad máxima de sinónimos por término. Para obtener más información, consulte Cuotas para Amazon Kendra.

En el siguiente ejemplo, se muestra un archivo de tesauros con reglas de sinónimos. Cada línea contiene una única regla de sinónimos. Se ignoran las líneas en blanco y los comentarios.

```
# Lines starting with pound are comments and blank lines are ignored.

# Synonym relationships can be defined as unidirectional or bidirectional
 relationships.
```

```
# Unidirection relationships are represented by any term sequence
# on the left hand side (LHS) of "=>" followed by synonyms on the right hand side (RHS)
CodeStar => AWS CodeStar
# This will map CodeStar to AWS CodeStar, but not vice-versa

# To map terms vice versa
ML => Machine Learning
Machine Learning => ML

# Multiple synonym relationships may be defined in one line as well by comma
 seperation.
autoscaling group, ASG => Auto Scaling group, autoscaling
# The above is equivalent to:
# autoscaling group => Auto Scaling group, autoscaling
# ASG => Auto Scaling group, autoscaling

# Bi-directional synonyms are comma separated terms with no "=>"
DNS, Route53, Route 53
# DNS, Route53, and Route 53 map to one another and are interchangeable at match time
# The above is equivalent to:
# DNS => Route53, Route 53
# Route53 => DNS, Route 53
# Route 53 => DNS, Route53

# Overlapping LHS terms will be merged
Beta => Alpha
Beta => Gamma
Beta, Delta
# is equivalent to:
# Beta => Alpha, Gamma, Delta
# Delta => Beta

# Each line contains a single synonym rule.
# Synonym rule count is the total number of lines defining synonym relationships
# Term count is the total number of unique terms for all rules.
# Comments and blanks lines do not count.
```

Adición de un tesauro a un índice

Los siguientes procedimientos muestran cómo agregar un archivo de sinónimos a un índice. Puede tardar de hasta 30 minutos en ver los efectos del archivo de sinónimos actualizado. Para más información sobre el archivo del tesauro, véase Crear un archivo de tesauro.

Console

Para agregar un diccionario de sinónimos

1. En el panel de navegación izquierdo, bajo el índice donde desea añadir una lista de sinónimos, su tesauro, seleccione Sinónimos.
2. En la página de Sinónimos, elija Añadir tesauro.
3. En Definir tesauro, asigne un nombre al tesauro y, si lo desea, una descripción.
4. En la configuración del tesauro, indique la Amazon S3 ruta al archivo del tesauro. El archivo debe tener un tamaño inferior a 5 MB.
5. Para el rol de IAM, seleccione un rol o seleccione Crear un nuevo rol y especifique un nombre de rol para crear un nuevo rol. Amazon Kendra utiliza este rol para acceder al Amazon S3 recurso en su nombre. El rol de IAM tiene el prefijo "AmazonKendra-».
6. Seleccione Guardar para guardar la configuración y añadir el tesauro. Una vez ingerido, el tesauro se activa y los sinónimos aparecen resaltados en los resultados. Puede tardar hasta 30 minutos en ver los efectos de su archivo de tesauro.

CLI

Para añadir un tesario a un índice con el, llame a: AWS CLI`create-thesaurus`

```
aws kendra create-thesaurus \
--index-id index-id \
--name "thesaurus-name" \
--description "thesaurus-description" \
--source-s3-path "Bucket=bucket-name,Key=thesaurus/synonyms.txt" \
--role-arn role-arn
```

Llame a `list-thesauri` para ver una lista de tesauros:

```
aws kendra list-thesauri \
```

```
--index-id index-id
```

Para ver los detalles de un tesauro, llame a `describe-thesaurus`:

```
aws kendra describe-thesaurus \
--index-id index-id \
--index-id thesaurus-id
```

Puede tardar hasta 30 minutos en ver los efectos de su archivo de tesauro.

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Create a thesaurus")

thesaurus_name = "thesaurus-name"
thesaurus_description = "thesaurus-description"
thesaurus_role_arn = "role-arn"

index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "thesaurus-file"
source_s3_path= {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    thesaurus_response = kendra.create_thesaurus(
        Description = thesaurus_description,
        Name = thesaurus_name,
        RoleArn = thesaurus_role_arn,
        IndexId = index_id,
        SourceS3Path = source_s3_path
    )
```

```
    pprint.pprint(thesaurus_response)

    thesaurus_id = thesaurus_response["Id"]

    print("Wait for Kendra to create the thesaurus.")

    while True:
        # Get thesaurus description
        thesaurus_description = kendra.describe_thesaurus(
            Id = thesaurus_id,
            IndexId = index_id
        )
        # If status is not CREATING quit
        status = thesaurus_description["Status"]
        print("Creating thesaurus. Status: " + status)
        if status != "CREATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.CreateThesaurusResponse;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusResponse;
import software.amazon.awssdk.services.kendra.model.S3Path;
import software.amazon.awssdk.services.kendra.model.ThesaurusStatus;

public class CreateThesaurusExample {

  public static void main(String[] args) throws InterruptedException {

    KendraClient kendra = KendraClient.builder().build();

    String thesaurusName = "thesaurus-name";
```

```
    String thesaurusDescription = "thesaurus-description";
    String thesaurusRoleArn = "role-arn";

    String s3BucketName = "bucket-name";
    String s3Key = "thesaurus-file";
    String indexId = "index-id";

    System.out.println(String.format("Creating a thesaurus named %s",
thesaurusName));
    CreateThesaurusRequest createThesaurusRequest = CreateThesaurusRequest
        .builder()
        .name(thesaurusName)
        .indexId(indexId)
        .description(thesaurusDescription)
        .roleArn(thesaurusRoleArn)
        .sourceS3Path(S3Path.builder()
            .bucket(s3BucketName)
            .key(s3Key)
            .build())
        .build();
    CreateThesaurusResponse createThesaurusResponse =
kendra.createThesaurus(createThesaurusRequest);
    System.out.println(String.format("Thesaurus response %s",
createThesaurusResponse));

    String thesaurusId = createThesaurusResponse.id();

    System.out.println(String.format("Waiting until the thesaurus with ID %s is
created.", thesaurusId));

    while (true) {
      DescribeThesaurusRequest describeThesaurusRequest =
DescribeThesaurusRequest.builder()
          .id(thesaurusId)
          .indexId(indexId)
          .build();
      DescribeThesaurusResponse describeThesaurusResponse =
kendra.describeThesaurus(describeThesaurusRequest);
      ThesaurusStatus status = describeThesaurusResponse.status();
      if (status != ThesaurusStatus.CREATING) {
        break;
      }

      TimeUnit.SECONDS.sleep(60);
```

```
    }

    System.out.println("Thesaurus creation is complete.");
  }
}
```

Actualización de un tesauro

Puede cambiar la configuración de un tesauro después de crearlo. Puede cambiar detalles como el nombre del tesauro y la información de IAM. También puede cambiar la ubicación de la ruta Amazon S3 del archivo de tesauros. Si cambia la ruta al archivo de tesauros, Amazon Kendra reemplaza el tesauro existente por el tesauro especificado en la ruta actualizada.

Puede tardar de hasta 30 minutos en ver los efectos del archivo de sinónimos actualizado.

> **Note**
>
> Si hay errores de validación o de sintaxis en el archivo del tesauro, se conserva el archivo del tesauro cargado anteriormente.

Los siguientes procedimientos muestran cómo modificar los detalles del tesauro.

Console

Modificación de los detalles del tesauro

1. En el panel de navegación izquierdo, en el índice que desee modificar, elija Sinónimos.
2. En la página Sinónimos, seleccione el tesauro que desee modificar y, a continuación, elija Editar.
3. En la página Actualizar el tesauro, actualice los detalles del tesauro.
4. (Opcional) Seleccione Cambiar la ruta del archivo del tesauro y, a continuación, especifique una Amazon S3 ruta al nuevo archivo del tesauro. El archivo de tesauros existente se sustituye por el archivo que especifique. Si no cambia la ruta, Amazon Kendra vuelve a cargar el tesauro desde la ruta existente.

   Si selecciona Conservar el archivo del tesauro actual, Amazon Kendra no se vuelve a cargar el archivo del tesauro.

5. Seleccione Guardar para guardar la configuración.

También puede volver a cargar el tesauro desde la ruta del tesauro existente.

Para volver a cargar un tesauro desde una ruta existente

1. En el panel de navegación izquierdo, en el índice que desee modificar, elija Sinónimos.
2. En la página Sinónimos, seleccione el tesauro que desea recargar y luego elija Actualizar.
3. En la página Recargar el archivo del tesauro, confirme que desea actualizar el archivo del tesauro.

CLI

Para actualizar un tesauro, llame a `update-thesaurus`:

```
aws kendra update-thesaurus \
--index-id index-id \
--name "thesaurus-name" \
--description "thesaurus-description" \
--source-s3-path "Bucket=bucket-name,Key=thesaurus/synonyms.txt" \
--role-arn role-arn
```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra = boto3.client("kendra")

print("Update a thesaurus")

thesaurus_name = "thesaurus-name"
thesaurus_description = "thesaurus-description"
thesaurus_role_arn = "role-arn"

thesaurus_id = "thesaurus-id"
index_id = "index-id"

s3_bucket_name = "bucket-name"
```

```
s3_key = "thesaurus-file"
source_s3_path= {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    kendra.update_thesaurus(
        Id = thesaurus_id,
        IndexId = index_id,
        Description = thesaurus_description,
        Name = thesaurus_name,
        RoleArn = thesaurus_role_arn,
        SourceS3Path = source_s3_path
    )

    print("Wait for Kendra to update the thesaurus.")

    while True:
        # Get thesaurus description
        thesaurus_description = kendra.describe_thesaurus(
            Id = thesaurus_id,
            IndexId = index_id
        )
        # If status is not UPDATING quit
        status = thesaurus_description["Status"]
        print("Updating thesaurus. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.UpdateThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusRequest;
```

```
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusResponse;
import software.amazon.awssdk.services.kendra.model.S3Path;
import software.amazon.awssdk.services.kendra.model.ThesaurusStatus;

public class UpdateThesaurusExample {

  public static void main(String[] args) throws InterruptedException {

    KendraClient kendra = KendraClient.builder().build();

    String thesaurusName = "thesaurus-name";
    String thesaurusDescription = "thesaurus-description";
    String thesaurusRoleArn = "role-arn";

    String s3BucketName = "bucket-name";
    String s3Key = "thesaurus-file";

    String thesaurusId = "thesaurus-id";
    String indexId = "index-id";

    UpdateThesaurusRequest updateThesaurusRequest = UpdateThesaurusRequest
        .builder()
        .id(thesaurusId)
        .indexId(indexId)
        .name(thesaurusName)
        .description(thesaurusDescription)
        .roleArn(thesaurusRoleArn)
        .sourceS3Path(S3Path.builder()
            .bucket(s3BucketName)
            .key(s3Key)
            .build())
        .build();
    kendra.updateThesaurus(updateThesaurusRequest);

    System.out.println(String.format("Waiting until the thesaurus with ID %s is
 updated.", thesaurusId));

    // a new source s3 path requires re-consumption by Kendra
    // and so can take as long as a Create Thesaurus operation
    while (true) {
      DescribeThesaurusRequest describeThesaurusRequest =
 DescribeThesaurusRequest.builder()
          .id(thesaurusId)
          .indexId(indexId)
```

```
            .build();
        DescribeThesaurusResponse describeThesaurusResponse =
 kendra.describeThesaurus(describeThesaurusRequest);
        ThesaurusStatus status = describeThesaurusResponse.status();
        if (status != ThesaurusStatus.UPDATING) {
          break;
        }

        TimeUnit.SECONDS.sleep(60);
      }

      System.out.println("Thesaurus update is complete.");
    }
}
```

Eliminar un tesauro

Los siguientes procedimientos muestran cómo eliminar un tesauro.

Console

1. En el panel de navegación izquierdo, en el índice que desee modificar, elija Sinónimos.
2. En la página Sinónimos, seleccione el tesauro que desee eliminar.
3. En la página Detalles del Tesauro, seleccione Eliminar y luego confirme para borrar.

CLI

Para eliminar un tesauro y colocarlo en un índice con el, llame a: AWS CLI`delete-thesaurus`

```
aws kendra delete-thesaurus \
--index-id index-id \
--id thesaurus-id
```

Python

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")
```

```
print("Delete a thesaurus")

thesaurus_id = "thesaurus-id"
index_id = "index-id"

try:
    kendra.delete_thesaurus(
        Id = thesaurus_id,
        IndexId = index_id
    )

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.DeleteThesaurusRequest;

public class DeleteThesaurusExample {

  public static void main(String[] args) throws InterruptedException {

    KendraClient kendra = KendraClient.builder().build();

    String thesaurusId = "thesaurus-id";
    String indexId = "index-id";

    DeleteThesaurusRequest updateThesaurusRequest = DeleteThesaurusRequest
        .builder()
        .id(thesaurusId)
        .indexId(indexId)
        .build();
    kendra.deleteThesaurus(updateThesaurusRequest);
  }
}
```

Aspectos destacados en los resultados de búsqueda

El resaltado de sinónimos está activado de forma predeterminada. La información destacada se incluye en los resultados de las consultas Amazon Kendra del SDK y la CLI. Si interactúa con el Amazon Kendra SDK o la CLI, determina cómo mostrar los resultados.

Los sinónimos destacados tendrán el tipo de resaltado `THESAURUS_SYNONYM`. Para obtener más información sobre las características principales, consulte el objeto Resaltar.

Tutorial: Creación de una solución de búsqueda inteligente y enriquecida con metadatos con Amazon Kendra

Este tutorial le muestra cómo crear una solución de búsqueda inteligente enriquecida con metadatos y basada en lenguaje natural para los datos de su empresa mediante Amazon Kendra, Amazon Comprehend, Amazon Simple Storage Service (S3), y AWS CloudShell.

Amazon Kendra es un servicio de búsqueda inteligente que puede crear un índice de búsqueda para sus repositorios de datos no estructurados en lenguaje natural. Para facilitar a sus clientes la búsqueda y el filtrado de las respuestas relevantes, puede utilizar Amazon Comprehend para extraer metadatos de sus datos e incorporarlos a su índice de búsqueda de Amazon Kendra.

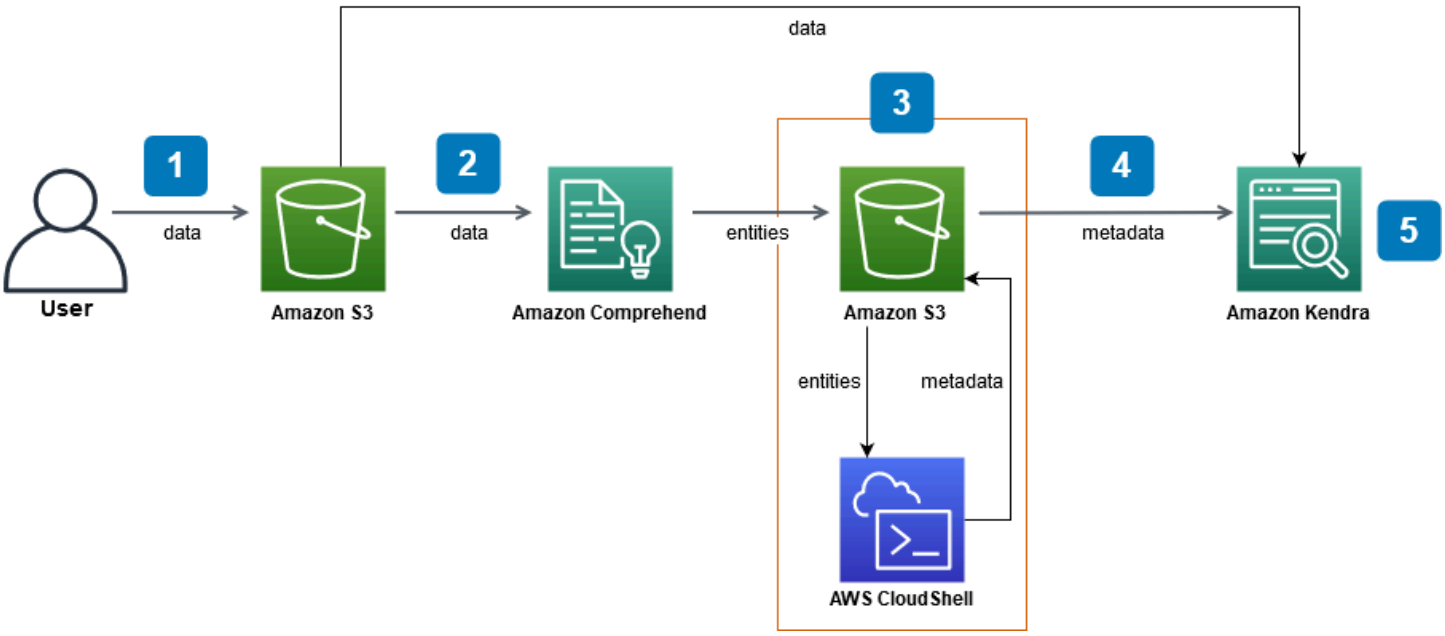
Amazon Comprehend es un servicio de procesamiento de lenguaje natural (NLP) que puede identificar entidades. Las entidades son referencias a personas, lugares, ubicaciones, organizaciones y objetos en sus datos.

En este tutorial, se utiliza un conjunto de datos de muestra de artículos de noticias para extraer entidades, convertirlas en metadatos e incorporarlas al índice de Amazon Kendra para realizar búsquedas. Los metadatos agregados le permiten filtrar los resultados de la búsqueda utilizando cualquier subconjunto de estas entidades y mejoran la precisión de la búsqueda. Al seguir este tutorial, aprenderá a crear una solución de búsqueda para los datos de su empresa sin necesidad de conocimientos especializados en machine learning.

En este tutorial, se muestra cómo crear una solución de búsqueda mediante los siguientes pasos:

1. Almacenamiento de un conjunto de datos de muestra de artículos de noticias en Amazon S3.
2. Uso de Amazon Comprehend para extraer entidades de sus datos.
3. Ejecutar un script de Python 3 para convertir las entidades al formato de metadatos de índice de Amazon Kendra y almacenar estos metadatos en S3.
4. Crear un índice de búsqueda de Amazon Kendra e ingerir los datos y los metadatos.
5. Consulta del índice de búsqueda.

El siguiente diagrama muestra el flujo de trabajo:



Tiempo estimado para completar este tutorial: 1 hora

Coste estimado: algunas de las acciones de este tutorial conllevan cargos en tu AWS cuenta. Para obtener más información sobre el coste de cada servicio, consulte las páginas de precios de Amazon S3, Amazon Comprehend, AWS CloudShell y Amazon Kendra.

Temas

- Requisitos previos
- Paso 1: Añadir documentos a Amazon S3
- Paso 2: Ejecutar un trabajo de análisis de entidades en Amazon Comprehend
- Paso 3: Formatear el resultado del análisis de entidades como metadatos de Amazon Kendra
- Paso 4: Creación de un índice de Amazon Kendra e ingesta de los metadatos
- Paso 5: Consulta del índice de Amazon Kendra
- Paso 6: Limpieza

Requisitos previos

Para completar este tutorial, necesita los siguientes recursos:

- Una AWS cuenta. Si no tienes una AWS cuenta, sigue los pasos que se indican en Configuración de Amazon Kendra para configurarla. AWS

- Un ordenador de desarrollo con Windows, macOS o Linux, para acceder a la consola de gestión de AWS . Para obtener más información, consulte Configuración de la consola AWS de administración.
- Un usuario de AWS Identity and Access Management (IAM). Para obtener información sobre cómo configurar un usuario y un grupo de IAM para su cuenta, consulte la sección Primeros pasos de la Guía del usuario de IAM.

  Si utiliza la AWS Command Line Interface, también debe adjuntar la siguiente política a su usuario de IAM para concederle los permisos básicos necesarios para completar este tutorial.

  Para más información, consulte Creación de políticas de IAM y Adición y eliminación de permisos de identidad de IAM.
- La lista de servicios regionales de AWS. Para reducir la latencia, debe elegir la región de AWS más cercana a su ubicación geográfica que sea compatible con Amazon Comprehend y Amazon Kendra.
- (Opcional) Un AWS Key Management Service. Si bien este tutorial no utiliza el cifrado, es posible que desee utilizar las mejores prácticas de cifrado para su caso de uso específico.
- (Opcional) Una Amazon Virtual Private Cloud. Aunque este tutorial no utiliza una VPC, es posible que desee utilizar las mejores prácticas de VPC para garantizar la seguridad de los datos para su caso de uso específico.

Paso 1: Añadir documentos a Amazon S3

Antes de ejecutar un trabajo de análisis de entidades de Amazon Comprehend en su conjunto de datos, debe crear un bucket de Amazon S3 para alojar los datos, los metadatos y el resultado del análisis de entidades de Amazon Comprehend.

Temas

- Descarga del conjunto de datos de muestra
- Creación de un bucket de Amazon S3
- Crear carpetas de datos y metadatos en su bucket de S3
- Cargar los datos de entrada.

Descarga del conjunto de datos de muestra

Antes de que Amazon Comprehend pueda ejecutar un trabajo de análisis de entidades en sus datos, debe descargar y extraer el conjunto de datos y cargarlo en un bucket de S3.

Para descargar y extraer el conjunto de datos (consola)

1. Descargue la carpeta tutorial-dataset.zip en su dispositivo.
2. Extraiga la carpeta `tutorial-dataset` para acceder a la carpeta `data`.

Para descargar y extraer el conjunto de datos (Terminal)

1. Para descargar el `tutorial-dataset`, ejecute el siguiente comando en una ventana de terminal:

   Linux

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   Donde:

   - *path/* es la ruta del archivo local a la ubicación en la que desea guardar la carpeta zip.

   macOS

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   Donde:

   - *path/* es la ruta del archivo local a la ubicación en la que desea guardar la carpeta zip.

   Windows

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

Donde:

- *path/* es la ruta del archivo local a la ubicación en la que desea guardar la carpeta zip.

2. Para extraer los datos de la carpeta zip, ejecute el siguiente comando en la ventana del terminal:

   Linux

   ```
   unzip path/tutorial-dataset.zip -d path/
   ```

   Donde:

   - *path/* es la ruta de archivo local a la carpeta zip guardada.

   macOS

   ```
   unzip path/tutorial-dataset.zip -d path/
   ```

   Donde:

   - *path/* es la ruta de archivo local a la carpeta zip guardada.

   Windows

   ```
   tar -xf path/tutorial-dataset.zip -C path/
   ```

   Donde:

   - *path/* es la ruta de archivo local a la carpeta zip guardada.

Al final de este paso, deberías tener los archivos extraídos en una carpeta descomprimida llamada `tutorial-dataset`. Esta carpeta contiene un archivo `README` con una atribución de código abierto de Apache 2.0 y una carpeta llamada `data` que contiene el conjunto de datos de este tutorial. El conjunto de datos consta de 100 archivos con extensiones `.story`.

Creación de un bucket de Amazon S3

Después de descargar y extraer la carpeta de datos de muestra, se almacena en un bucket de Amazon S3.

> **Important**
>
> El nombre de un bucket de Amazon S3 debe ser único en todas las AWS.

Para crear un bucket de S3 (Consola)

1. Inicie sesión en la consola de Amazon S3 AWS Management Console y ábrala en https://console.aws.amazon.com/s3/.
2. En Buckets, seleccione Crear bucket.
3. En Nombre del bucket, escriba un nombre único.
4. En Región, elija la AWS región en la que desee crear el bucket.

   > **Note**
   >
   > Debe elegir una región que admita Amazon Comprehend y Amazon Kendra. No puede cambiar la región de un bucket después de haberlo creado.

5. Mantenga la configuración predeterminada para Bloquear el acceso público para este bucket, el control de versiones del bucket y las etiquetas.
6. Para el cifrado predeterminado, seleccione Desactivar.
7. Mantenga la configuración predeterminada para la Configuración avanzada.
8. Revise la configuración del bucket y elija Crear bucket.

Para crear un bucket de S3 (AWS CLI)

1. Para crear un bucket de S3 use el comando create-bucket en la AWS CLI.

   Linux

   ```
   aws s3api create-bucket \
           --bucket DOC-EXAMPLE-BUCKET \
           --region aws-region \
   ```

```
        --create-bucket-configuration LocationConstraint=aws-region
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket,
- *aws-region* es la región en la que quiere crear su bucket.

macOS

```
aws s3api create-bucket \
        --bucket DOC-EXAMPLE-BUCKET \
        --region aws-region \
        --create-bucket-configuration LocationConstraint=aws-region
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket,
- *aws-region* es la región en la que quiere crear su bucket.

Windows

```
aws s3api create-bucket ^
        --bucket DOC-EXAMPLE-BUCKET ^
        --region aws-region ^
        --create-bucket-configuration LocationConstraint=aws-region
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket,
- *aws-region* es la región en la que quiere crear su bucket.

> **Note**
>
> Debe elegir una región que admita Amazon Comprehend y Amazon Kendra. No puede cambiar la región de un bucket después de haberlo creado.

2. Para asegurarse de que su bucket se creó correctamente, utilice el comando list:

Linux

```
aws s3 ls
```

macOS

```
aws s3 ls
```

Windows

```
aws s3 ls
```

Crear carpetas de datos y metadatos en su bucket de S3

Tras crear su bucket S3, cree carpetas de datos y metadatos en su interior.

Para crear carpetas en su bucket de S3 (Consola)

1. Abra la consola de Amazon S3 en https://console.aws.amazon.com/s3.
2. En Buckets, haga clic en el nombre de su bucket de la lista.
3. En la pestaña Objetos, elija Crear carpeta.
4. Para el nombre de la nueva carpeta, escriba **`data`**.
5. Para la configuración de cifrado de carpeta, elija Desactivar.
6. Elija Crear carpeta.
7. Repita los pasos 3 a 6 para crear otra carpeta para almacenar los metadatos de Amazon Kendra y asigne un nombre a la carpeta creada en el paso 4 **`metadata`**.

Para crear carpetas en su bucket de S3 (AWS CLI)

1. Para crear la carpeta `data` en su bucket de S3, utilice el comando put-object en la AWS CLI:

   Linux

   ```
   aws s3api put-object \
           --bucket DOC-EXAMPLE-BUCKET \
   ```

```
        --key data/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

macOS

```
aws s3api put-object \
        --bucket DOC-EXAMPLE-BUCKET \
        --key data/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

Windows

```
aws s3api put-object ^
        --bucket DOC-EXAMPLE-BUCKET ^
        --key data/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

2. Para crear la carpeta `metadata` en su bucket de S3, utilice el comando put-object en la AWS CLI:

Linux

```
aws s3api put-object \
        --bucket DOC-EXAMPLE-BUCKET \
        --key metadata/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

macOS

```
aws s3api put-object \
        --bucket DOC-EXAMPLE-BUCKET \
        --key metadata/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

Windows

```
aws s3api put-object ^
        --bucket DOC-EXAMPLE-BUCKET ^
        --key metadata/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

3. Para asegurarse de que sus carpetas se han creado correctamente, compruebe el contenido de su bucket utilizando el comando list:

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

Cargar los datos de entrada.

Tras crear las carpetas de datos y metadatos, debe cargar el conjunto de datos de muestra en la carpeta `data`.

Para cargar el conjunto de datos de muestra en la carpeta de datos (Consola)

1. Abra la consola de Amazon S3 en https://console.aws.amazon.com/s3.
2. En Buckets, haga clic en el nombre de su bucket de la lista y haga clic en `data`.
3. Elija Cargar y, a continuación, Agregar archivo.
4. En el cuadro de diálogo, navegue hasta la carpeta `data` incluida en la carpeta `tutorial-dataset` de su dispositivo local, seleccione todos los archivos y, a continuación, elija Abrir.
5. Mantenga la configuración predeterminada de Destino, Permisos y Propiedades.
6. Seleccione Cargar.

Para cargar el conjunto de datos de muestra en la carpeta de datos (AWS CLI)

1. Para cargar los datos de muestra en la carpeta `data`, utilice el comando copy en la AWS CLI:

   Linux

   ```
   aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
   ```

   Donde:

   - *path/* es la ruta del archivo a la carpeta `tutorial-dataset` de su dispositivo,
   - DOC-EXAMPLE-BUCKET es el nombre de su bucket.

macOS

```
aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
```

Donde:

- *path/* es la ruta del archivo a la carpeta `tutorial-dataset` de su dispositivo,
- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

Windows

```
aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
```

Donde:

- *path/* es la ruta del archivo a la carpeta `tutorial-dataset` de su dispositivo,
- DOC-EXAMPLE-BUCKET es el nombre de su bucket.

2. Para asegurarse de que sus archivos de conjuntos de datos se han cargado correctamente en su carpeta `data`, utilice el comando list en la AWS CLI:

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

Al final de este paso, dispondrá de un bucket de S3 con el conjunto de datos almacenado en la carpeta `data` y de una carpeta `metadata` vacía en la que se almacenarán los metadatos de Amazon Kendra.

Paso 2: Ejecutar un trabajo de análisis de entidades en Amazon Comprehend

Tras almacenar el conjunto de datos de muestra en su bucket de S3, ejecuta un trabajo de análisis de entidades de Amazon Comprehend para extraer entidades de sus documentos. Estas entidades formarán atributos personalizados de Amazon Kendra y le ayudarán a filtrar los resultados de búsqueda en su índice. Para más información, consulte Detectar entidades.

Temas

- Ejecución de un trabajo de análisis de entidades de Amazon Comprehend

Ejecución de un trabajo de análisis de entidades de Amazon Comprehend

Para extraer entidades de su conjunto de datos, ejecute un trabajo de análisis de entidades de Amazon Comprehend.

Si utiliza la AWS CLI en este paso, primero debe crear y adjuntar un rol y una política de AWS IAM para Amazon Comprehend y, a continuación, ejecutar un trabajo de análisis de entidades. Para ejecutar un trabajo de análisis de entidades en sus datos de muestra, Amazon Comprehend necesita:

- una función AWS Identity and Access Management (IAM) que la reconozca como una entidad de confianza

- una política de AWS IAM asociada a la función de IAM que le otorga permisos para acceder a su bucket de S3

Para obtener más información, consulte Cómo funciona Amazon Comprehend con IAM y las políticas basadas en la identidad de Amazon Comprehend.

Para ejecutar un trabajo de análisis de entidades de Amazon Comprehend (Consola)

1. Abra la consola Amazon Comprehend en https://console.aws.amazon.com/comprehend/.

   > **Important**
   >
   > Asegúrese de estar en la misma región en la que creó el bucket de Amazon S3. Si se encuentra en otra región, elija la AWS región en la que creó su bucket de S3 en el selector de regiones de la barra de navegación superior.

2. Elija Lanzar Amazon Comprehend.
3. En el panel de navegación izquierdo, elija Trabajos de análisis.
4. Seleccione Crear trabajo.
5. En la sección Configuración, realice lo siguiente:
   a. En Nombre, escriba **`data-entities-analysis`**.
   b. En Tipo de análisis, elija Entidades.
   c. En Idioma, elija Inglés.
   d. Mantenga desactivado el cifrado de trabajos.
6. En la sección Datos de entrada, realice lo siguiente:
   a. En Origen de datos, seleccione Mis documentos.
   b. Para la ubicación de S3, elija Examinar S3.
   c. En Elegir recursos, haga clic en el nombre de su bucket de la lista.
   d. Para Objetos, seleccione el botón de opción para `data` y seleccione Elegir.
   e. En Formato de entrada, elija Un documento por archivo.
7. En la sección Datos de salida, realice lo siguiente:
   a. Para la ubicación de S3, elija Examinar S3 y, a continuación, seleccione la casilla de opciones para su bucket en la lista de buckets y seleccione Elegir.

b. Mantenga desactivado el Cifrado.

8. En la sección Permisos de acceso, haga lo siguiente:

   a. En Rol de IAM, elija Crear un nuevo rol.

   b. Para ver los permisos de acceso, seleccione Buckets de S3 de entrada y salida.

   c. En Sufijo de nombre, escriba **`comprehend-role`**. Este rol proporciona acceso a su bucket de Amazon S3.

9. Mantenga la Configuración de la VPC predeterminada.

10. Seleccione Crear trabajo.

Para ejecutar un trabajo de análisis de entidades de Amazon Comprehend (AWS CLI)

1. Para crear y adjuntar un rol de IAM para Amazon Comprehend que lo reconozca como una entidad de confianza, haga lo siguiente:

   a. Guarde la siguiente política de confianza como un archivo JSON llamado `comprehend-trust-policy.json` en un editor de texto de su dispositivo local.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "comprehend.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

   b. Para crear un rol de IAM llamado `comprehend-role` y adjuntarle el archivo `comprehend-trust-policy.json` guardado, use el comando create-role:

      Linux

```
aws iam create-role \
          --role-name comprehend-role \
```

```
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Donde:

- *path/* es la ruta del archivo a `comprehend-trust-policy.json` en su dispositivo local.

macOS

```
aws iam create-role \
          --role-name comprehend-role \
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Donde:

- *path/* es la ruta del archivo a `comprehend-trust-policy.json` en su dispositivo local.

Windows

```
aws iam create-role ^
          --role-name comprehend-role ^
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Donde:

- *path/* es la ruta del archivo a `comprehend-trust-policy.json` en su dispositivo local.

c. Copie el nombre de recurso de Amazon (ARN) en el editor de texto y guárdelo de forma local como `comprehend-role-arn`.

> **Note**
>
> El ARN tiene un formato similar a *arn:aws:iam: :123456789012:role/comprehend-role*. Necesita el ARN que guardó como `comprehend-role-arn` para ejecutar el trabajo de análisis de Amazon Comprehend.

2. Para crear y adjuntar una política de IAM a su rol de IAM que le conceda permisos para acceder a su bucket de S3, haga lo siguiente:

   a. Guarde la siguiente política de confianza como un archivo JSON llamado `comprehend-S3-access-policy.json` en un editor de texto de su dispositivo local.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::DOC-EXAMPLE-BUCKET/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::DOC-EXAMPLE-BUCKET"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:PutObject"
            ],
            "Resource": [
                "arn:aws:s3:::DOC-EXAMPLE-BUCKET/*"
            ],
            "Effect": "Allow"
```

```
        }
    ]
}
```

b. Para crear una política de IAM llamada `comprehend-S3-access-policy` para acceder a su bucket de S3, utilice el comando create-policy:

Linux

```
aws iam create-policy \
          --policy-name comprehend-S3-access-policy \
          --policy-document file://path/comprehend-S3-access-policy.json
```

Donde:

- *path/* es la ruta del archivo a `comprehend-S3-access-policy.json` en su dispositivo local.

macOS

```
aws iam create-policy \
          --policy-name comprehend-S3-access-policy \
          --policy-document file://path/comprehend-S3-access-policy.json
```

Donde:

- *path/* es la ruta del archivo a `comprehend-S3-access-policy.json` en su dispositivo local.

Windows

```
aws iam create-policy ^
          --policy-name comprehend-S3-access-policy ^
          --policy-document file://path/comprehend-S3-access-policy.json
```

Donde:

- *path/* es la ruta del archivo a `comprehend-S3-access-policy.json` en su dispositivo local.

c. Copie el nombre de recurso de Amazon (ARN) en el editor de texto y guárdelo de forma local como `comprehend-S3-access-arn`.

> **Note**
>
> El ARN tiene un formato similar a *`arn:aws:iam: :123456789012:role/Comprehend-S3-Access-Policy`*. Necesita el ARN que guardó como `comprehend-S3-access-arn` para asociar el `comprehend-S3-access-policy` a su rol de IAM.

d. Para asociarlo `comprehend-S3-access-policy` a su función de IAM, utilice el attach-role-policycomando:

Linux

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name comprehend-role
```

Donde:

- *`policy-arn`* es el ARN que guardó como `comprehend-S3-access-arn`.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name comprehend-role
```

Donde:

- *`policy-arn`* es el ARN que guardó como `comprehend-S3-access-arn`.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name comprehend-role
```

Donde:

- *policy-arn* es el ARN que guardó como comprehend-S3-access-arn.

3. Para ejecutar un trabajo de análisis de entidades de Amazon Comprehend, utilice el start-entities-detection-jobcomando:

Linux

```
aws comprehend start-entities-detection-job \
        --input-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/
data/,InputFormat=ONE_DOC_PER_FILE \
        --output-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/ \
        --data-access-role-arn role-arn \
        --job-name data-entities-analysis \
        --language-code en \
        --region aws-region
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.
- *role-arn* es el ARN que guardó como comprehend-role-arn,
- *aws-region* es tu región. AWS

macOS

```
aws comprehend start-entities-detection-job \
        --input-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/
data/,InputFormat=ONE_DOC_PER_FILE \
        --output-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/ \
        --data-access-role-arn role-arn \
        --job-name data-entities-analysis \
        --language-code en \
        --region aws-region
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.
- *role-arn* es el ARN que guardó como comprehend-role-arn,

- *aws-region* es tu región. AWS

Windows

```
aws comprehend start-entities-detection-job ^
        --input-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/
data/,InputFormat=ONE_DOC_PER_FILE ^
        --output-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/ ^
        --data-access-role-arn role-arn ^
        --job-name data-entities-analysis ^
        --language-code en ^
        --region aws-region
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.
- *role-arn* es el ARN que guardó como `comprehend-role-arn`,
- *aws-region* es tu región. AWS

4. Copie el análisis de entidades `JobId` y guárdelo en un editor de texto como `comprehend-job-id`. El `JobId` le ayuda a realizar el seguimiento del estado de su trabajo de análisis de entidades.
5. Para realizar un seguimiento del progreso de su trabajo de análisis de entidades, utilice el describe-entities-detection-jobcomando:

   Linux

   ```
   aws comprehend describe-entities-detection-job \
           --job-id entities-job-id \
           --region aws-region
   ```

   Donde:

   - *entities-job-id*está guardadocomprehend-job-id,
   - *aws-region* es tu región. AWS

macOS

```
aws comprehend describe-entities-detection-job \
        --job-id entities-job-id \
        --region aws-region
```

Donde:

- *entities-job-id*es tu salvado, `comprehend-job-id`
- *aws-region* es tu región. AWS

Windows

```
aws comprehend describe-entities-detection-job ^
        --job-id entities-job-id ^
        --region aws-region
```

Donde:

- *entities-job-id*es tu salvado, `comprehend-job-id`
- *aws-region* es tu región. AWS

Puede tardar varios minutos en cambiar el `JobStatus` a `COMPLETED`.

Al final de este paso, Amazon Comprehend almacena los resultados del análisis de entidades como un archivo `output.tar.gz` comprimido dentro de una carpeta `output` generada automáticamente en su bucket de S3. Asegúrese de que el estado de su trabajo de análisis esté completo antes de pasar al siguiente paso.

Paso 3: Formatear el resultado del análisis de entidades como metadatos de Amazon Kendra

Para convertir las entidades extraídas por Amazon Comprehend al formato de metadatos requerido por un índice de Amazon Kendra, ejecute un script de Python 3. Los resultados de la conversión se almacenan en la carpeta `metadata` del bucket de Amazon S3.

Para obtener más información sobre el formato y la estructura de los metadatos de Amazon Kendra, consulte Metadatos de documentos de S3.

Temas

- Descargar y extraer el resultado de Amazon Comprehend
- Cargar la salida en el bucket de S3
- Conversión de la salida al formato de metadatos de Amazon Kendra
- Limpieza del bucket de Amazon S3

Descargar y extraer el resultado de Amazon Comprehend

Para formatear la salida del análisis de entidades de Amazon Comprehend, primero debe descargar el archivo `output.tar.gz` de análisis de entidades de Amazon Comprehend y extraer el archivo de análisis de entidades.

Para descargar y extraer el archivo de salida (Consola)

1. En el panel de navegación de la consola Amazon Comprehend, vaya a Trabajos de análisis.
2. Elija su trabajo de análisis de entidades `data-entities-analysis`.
3. En Salida, elija el enlace que aparece junto a la ubicación de los datos de salida. Esto lo redirige al archivo `output.tar.gz` de su bucket de S3.
4. En la página Información general, seleccione Descargar.

> **Tip**
>
> El resultado de todos los trabajos de análisis de Amazon Comprehend tiene el mismo nombre. Cambiar el nombre de su archivo le ayudará a rastrearlo más fácilmente.

5. Descomprime y extrae el archivo Amazon Comprehend descargado en tu dispositivo.

Para descargar y extraer el archivo de salida (AWS CLI)

1. Para acceder al nombre de la carpeta generada automáticamente por Amazon Comprehend en su bucket de S3 y que contiene los resultados del trabajo de análisis de entidades, utilice el comando: describe-entities-detection-job

Linux

```
aws comprehend describe-entities-detection-job \
          --job-id entities-job-id \
          --region aws-region
```

Donde:

- *entities-job-id*está guardado `comprehend-job-id` desde, the section called “Paso 2: Detectar entidades”
- *aws-region* es tu región. AWS

macOS

```
aws comprehend describe-entities-detection-job \
          --job-id entities-job-id \
          --region aws-region
```

Donde:

- *entities-job-id*está guardado `comprehend-job-id` desde, the section called “Paso 2: Detectar entidades”
- *aws-region* es tu región. AWS

Windows

```
aws comprehend describe-entities-detection-job ^
          --job-id entities-job-id ^
          --region aws-region
```

Donde:

- *entities-job-id*está guardado `comprehend-job-id` desde, the section called “Paso 2: Detectar entidades”
- *aws-region* es tu región. AWS

2. Del objeto `OutputDataConfig` de la descripción del trabajo de su entidad, copie y guarde el valor `S3Uri` como `comprehend-S3uri` en un editor de texto.

> **Note**
>
> *El `S3Uri` valor tiene un formato similar a `s3://DOC-EXAMPLE-BUCKET/...`* /output/output.tar.gz.

3. Para descargar el archivo de salida de las entidades, utilice el comando copy:

   Linux

   ```
   aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
   ```

   Donde:

   - *`s3://DOC-EXAMPLE-BUCKET/...`* /output/output.tar.gz es el `S3Uri` valor que guardastecomprehend-S3uri,
   - *`path/`* es el directorio local en el que desea guardar la salida.

   macOS

   ```
   aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
   ```

   Donde:

   - *`s3://DOC-EXAMPLE-BUCKET/...`* /output/output.tar.gz es el `S3Uri` valor que guardastecomprehend-S3uri,
   - *`path/`* es el directorio local en el que desea guardar la salida.

   Windows

   ```
   aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
   ```

   Donde:

- *s3://DOC-EXAMPLE-BUCKET/...* /output/output.tar.gz es el `S3Uri` valor que guardastecomprehend-`S3uri`,
- *path/* es el directorio local en el que desea guardar la salida.

4. Para extraer la salida de las entidades, ejecute el siguiente comando en una ventana de terminal:

   Linux

   ```
   tar -xf path/output.tar.gz -C path/
   ```

   Donde:

   - *path/* es la ruta del archivo `output.tar.gz` descargado en su dispositivo local.

   macOS

   ```
   tar -xf path/output.tar.gz -C path/
   ```

   Donde:

   - *path/* es la ruta del archivo `output.tar.gz` descargado en su dispositivo local.

   Windows

   ```
   tar -xf path/output.tar.gz -C path/
   ```

   Donde:

   - *path/* es la ruta del archivo `output.tar.gz` descargado en su dispositivo local.

Al final de este paso, deberías tener un archivo en tu dispositivo llamado `output` con una lista de las entidades identificadas por Amazon Comprehend.

Cargar la salida en el bucket de S3

Tras descargar y extraer el archivo de análisis de entidades de Amazon Comprehend, debe cargar el archivo `output` extraído en su bucket de Amazon S3.

Para cargar el archivo de salida extraído de Amazon Comprehend (Consola)

1. Abra la consola de Amazon S3 en https://console.aws.amazon.com/s3.
2. En la sección Buckets, haga clic en su bucket y, a continuación, elija Cargar.
3. En Archivos y carpetas, elija Añadir archivos.
4. En el cuadro de diálogo, navegue hasta el archivo `output` extraído en su dispositivo, selecciónelo y elija Abrir.
5. Mantenga la configuración predeterminada de Destino, Permisos y Propiedades.
6. Seleccione Cargar.

Para cargar el archivo de salida extraído de Amazon Comprehend (AWS CLI)

1. Para cargar el archivo `output` extraído a su bucket, utilice el comando copy:

   Linux

   ```
   aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
   ```

   Donde:

   - *path/* es la ruta del archivo local al archivo `output` extraído,
   - DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

   macOS

   ```
   aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
   ```

   Donde:

   - *path/* es la ruta del archivo local al archivo `output` extraído,
   - DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

Windows

```
aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
```

Donde:

- *path/* es la ruta del archivo local al archivo `output` extraído,
- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

2. Para asegurarse de que el archivo `output` se ha cargado correctamente en su bucket de S3, compruebe su contenido mediante el comando list:

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

Conversión de la salida al formato de metadatos de Amazon Kendra

Para convertir la salida de Amazon Comprehend en metadatos de Amazon Kendra, ejecute un script de Python 3. Si está utilizando la consola, utilice AWS CloudShell este paso.

Ejecución del script de Python 3 (consola)

1. Descargue el archivo comprimido converter.py.zip en su dispositivo.
2. Extraiga el archivo `converter.py` Python 3.
3. Inicie sesión en la consola AWS de administración y asegúrese de que su AWS región esté configurada en la misma región que su bucket de S3 y su trabajo de análisis de Amazon Comprehend.
4. Elija el AWS CloudShell icono o escríbalo AWS CloudShellen el cuadro de búsqueda de la barra de navegación superior para iniciar un entorno.

   > **Note**
   >
   > Cuando se AWS CloudShell abre por primera vez en una nueva ventana del navegador, aparece un panel de bienvenida con una lista de las funciones principales. El intérprete de comandos está listo para la interacción después de cerrar este panel y aparece el símbolo del sistema.

5. Una vez que el terminal esté preparado, seleccione Acciones en el panel de navegación y, a continuación, seleccione Cargar archivo en el menú.
6. En el cuadro de diálogo que se abre, seleccione Seleccionar archivo y, a continuación, elige el archivo `converter.py` de Python 3 descargado de su dispositivo. Seleccione Cargar.
7. En el AWS CloudShell entorno, introduzca el siguiente comando:

   ```
   python3 converter.py
   ```

8. Cuando la interfaz del intérprete de comandos le pida que introduzca el nombre del bucket de S3, introduzca el nombre del bucket de S3 y pulse Entrar.
9. Cuando la interfaz intérprete de comandos le pida que introduzca la ruta completa del archivo de salida de Comprehend, introduzca **`output`** y pulse Entrar.
10. Cuando la interfaz de intérprete de comandos le pida que introduzca la ruta completa del archivo a su carpeta de metadatos, introduzca **`metadata/`** y pulse Entrar .

> **Important**
>
> Para que los metadatos tengan el formato correcto, los valores de entrada de los pasos 8 a 10 deben ser exactos.

Para ejecutar el script de Python 3 (AWS CLI)

1. Para descargar el archivo `converter.py` para Python 3, ejecute el siguiente comando en una ventana de terminal:

   Linux

   ```
   curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/samples/converter.py.zip
   ```

   Donde:

   - *path/* es la ruta del archivo a la ubicación en la que desea guardar el archivo comprimido.

   macOS

   ```
   curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/samples/converter.py.zip
   ```

   Donde:

   - *path/* es la ruta del archivo a la ubicación en la que desea guardar el archivo comprimido.

   Windows

   ```
   curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/samples/converter.py.zip
   ```

   Donde:

   - *path/* es la ruta del archivo a la ubicación en la que desea guardar el archivo comprimido.

2. Para extraer el archivo Python 3, ejecute el siguiente comando en la ventana del terminal:

Linux

```
unzip path/converter.py.zip -d path/
```

Donde:

- *path/* es la ruta del archivo a su `converter.py.zip`.

macOS

```
unzip path/converter.py.zip -d path/
```

Donde:

- *path/* es la ruta del archivo a su `converter.py.zip`.

Windows

```
tar -xf path/converter.py.zip -C path/
```

Donde:

- *path/* es la ruta del archivo a su `converter.py.zip`.

3. Asegúrese de que Boto3 esté instalado en el dispositivo; para ello, ejecute el siguiente comando.

Linux

```
pip3 show boto3
```

macOS

```
pip3 show boto3
```

Windows

```
pip3 show boto3
```

> **Note**
>
> Si no tiene Boto3 instalado, ejecute `pip3 install boto3` para instalarlo.

4. Para ejecutar el script de Python 3 para convertir el archivo `output`, ejecute el siguiente comando.

   Linux

   ```
   python path/converter.py
   ```

   Donde:

   - *path/* es la ruta del archivo a su `converter.py.zip`.

   macOS

   ```
   python path/converter.py
   ```

   Donde:

   - *path/* es la ruta del archivo a su `converter.py.zip`.

   Windows

   ```
   python path/converter.py
   ```

   Donde:

   - *path/* es la ruta del archivo a su `converter.py.zip`.

5. Cuando se AWS CLI le pida que lo haga`Enter the name of your S3 bucket`, introduzca el nombre de su depósito de S3 y pulse enter.
6. Cuando se AWS CLI le pida que lo haga`Enter the full filepath to your Comprehend output file`, introduzca **`output`** y pulse enter.

7. Cuando se AWS CLI le pida que lo haga`Enter the full filepath to your metadata folder`, introduzca **`metadata/`** y pulse enter.

> **Important**
>
> Para que los metadatos tengan el formato correcto, los valores de entrada de los pasos 5 a 7 deben ser exactos.

Al final de este paso, los metadatos formateados se depositan en la carpeta `metadata` del bucket de S3.

Limpieza del bucket de Amazon S3

Dado que el índice de Amazon Kendra sincroniza todos los archivos almacenados en un bucket, le recomendamos que limpie su bucket de Amazon S3 para evitar resultados de búsqueda redundantes.

Para limpiar su bucket de Amazon S3 (Consola)

1. Abra la consola de Amazon S3 en https://console.aws.amazon.com/s3.
2. En Buckets, elija su bucket y, a continuación, seleccione la carpeta de salida del análisis de entidades de Amazon Comprehend, el archivo `.temp` de análisis de entidades de Amazon Comprehend y el archivo `output` de Amazon Comprehend extraído.
3. En la pestaña Descripción general, seleccione Eliminar.
4. En Eliminar objetos, elija ¿Eliminar objetos permanentemente? e ingrese **`permanently delete`** en el campo de entrada de texto.
5. Elija Eliminar objetos.

Para limpiar su bucket de Amazon S3 (AWS CLI)

1. Para eliminar todos los archivos y carpetas de su bucket de S3, excepto las carpetas `data` y `metadata`, utilice el comando remove en la AWS CLI:

Linux

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude
 "metadata/*"
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

macOS

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude
 "metadata/*"
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

Windows

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude
 "metadata/*"
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

2. Para asegurarse de que los objetos se han eliminado correctamente de su bucket de S3, compruebe su contenido utilizando el comando list:

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- DOC-EXAMPLE-BUCKET es el nombre del bucket de S3.

Al final de este paso, ha convertido el resultado del análisis de entidades de Amazon Comprehend en metadatos de Amazon Kendra. Ahora está listo para crear un índice de Amazon Kendra.

Paso 4: Creación de un índice de Amazon Kendra e ingesta de los metadatos

Para implementar su solución de búsqueda inteligente, debe crear un índice de Amazon Kendra e incorporar en él los datos y metadatos de S3.

Antes de añadir metadatos a su índice de Amazon Kendra, debe crear campos de índice personalizados correspondientes a los atributos de los documentos personalizados, que a su vez corresponden a los tipos de entidad de Amazon Comprehend. Amazon Kendra utiliza los campos de índice y los atributos de documentos personalizados que cree para buscar y filtrar sus documentos.

Para obtener más información, consulte Indexación y Creación de atributos de documentos personalizados.

Temas

- Creación de un índice de Amazon Kendra
- Actualización del rol de IAM para el acceso a Amazon S3

- Creación de campos de índice de búsqueda personalizados de Amazon Kendra
- Agregar el bucket de Amazon S3 como origen de datos para el índice
- Sincronización del índice de Amazon Kendra

Creación de un índice de Amazon Kendra

Para consultar los documentos fuente, debe crear un índice de Amazon Kendra.

Si utiliza este paso, debe crear y adjuntar un rol y una política de AWS IAM que permitan a Amazon Kendra acceder a CloudWatch sus registros antes de crear un índice. AWS CLI Para más información, consulte Requisitos previos.

Creación de un índice de Amazon Kendra (consola)

1. Abra la consola Amazon Kendra en https://console.aws.amazon.com/kendra/.

   > **Important**
   >
   > Asegúrese de que se encuentra en la misma región en la que creó su trabajo de análisis de entidades de Amazon Comprehend y su bucket de Amazon S3. Si se encuentra en otra región, elija la AWS región en la que creó su bucket de Amazon S3 en el selector de regiones de la barra de navegación superior.

2. Elija Crear índice.
3. Para los Detalles del índice en la página Especificar detalles del índice, haga lo siguiente:

   a. En Nombre de índice, ingrese el **`kendra-index`**.
   b. Mantenga el campo Descripción en blanco.
   c. En Rol de IAM, elija Crear un nuevo rol. Este rol proporciona acceso a su bucket de Amazon S3.
   d. En Nombre del rol, ingrese **`kendra-role`**. El rol de IAM tendrá el prefijo `AmazonKendra-`.
   e. Mantenga la configuración predeterminada para el Cifrado y las etiquetas y seleccione Siguiente.
4. En Configuración del control de acceso en la página Configurar el control de acceso de los usuarios, elija No y, a continuación, Siguiente.
5. Para las Ediciones de aprovisionamiento en la página de Detalles de aprovisionamiento, elija Edición de desarrollador y seleccione Crear.

Crea un índice de Amazon Kendra (AWS CLI)

1. Para crear y adjuntar un rol de IAM para Amazon Kendra que lo reconozca como una entidad de confianza, haga lo siguiente:

   a. Guarde la siguiente política de confianza como un archivo JSON llamado `kendra-trust-policy.json` en un editor de texto de su dispositivo local.

   ```
   {
       "Version": "2012-10-17",
       "Statement": {
           "Effect": "Allow",
           "Principal": {
               "Service": "kendra.amazonaws.com"
           },
           "Action": "sts:AssumeRole"
       }
   }
   ```

   b. Para crear un rol de IAM llamado `kendra-role` y adjuntarle el archivo `kendra-trust-policy.json` guardado, use el comando create-role:

   Linux

   ```
   aws iam create-role \
            --role-name kendra-role \
            --assume-role-policy-document file://path/kendra-trust-policy.json
   ```

   Donde:

   - *path/* es la ruta del archivo a `kendra-trust-policy.json` en su dispositivo local.

   macOS

   ```
   aws iam create-role \
            --role-name kendra-role \
            --assume-role-policy-document file://path/kendra-trust-policy.json
   ```

   Donde:

- *path/* es la ruta del archivo a `kendra-trust-policy.json` en su dispositivo local.

Windows

```
aws iam create-role ^
         --role-name kendra-role ^
         --assume-role-policy-document file://path/kendra-trust-policy.json
```

Donde:

- *path/* es la ruta del archivo a `kendra-trust-policy.json` en su dispositivo local.

c. Copie el nombre de recurso de Amazon (ARN) en el editor de texto y guárdelo de forma local como `kendra-role-arn`.

> **Note**
>
> El ARN tiene un formato similar a *arn:aws:iam::123456789012:role/kendra-role*. Necesita el ARN que guardó como `kendra-role-arn` para ejecutar los trabajos de Amazon Kendra.

2. Antes de crear un índice, debe proporcionar su `kendra-role` permiso para escribir en CloudWatch Logs. Para ello, siga los pasos que se describen a continuación:

a. Guarde la siguiente política de confianza como un archivo JSON llamado `kendra-cloudwatch-policy.json` en un editor de texto de su dispositivo local.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Action":"cloudwatch:PutMetricData",
         "Resource":"*",
         "Condition":{
            "StringEquals":{
               "cloudwatch:namespace":"Kendra"
            }
         }
      },
```

```
      {
         "Effect":"Allow",
         "Action":"logs:DescribeLogGroups",
         "Resource":"*"
      },
      {
         "Effect":"Allow",
         "Action":"logs:CreateLogGroup",
         "Resource":"arn:aws:logs:aws-region:aws-account-id:log-group:/aws/
kendra/*"
      },
      {
         "Effect":"Allow",
         "Action":[
            "logs:DescribeLogStreams",
            "logs:CreateLogStream",
            "logs:PutLogEvents"
         ],
         "Resource":"arn:aws:logs:aws-region:aws-account-id:log-group:/aws/
kendra/*:log-stream:*"
      }
   ]
}
```

Sustituya *aws-region* por su región *aws-account-id*y por AWS su ID de cuenta de 12 dígitos. AWS

b. Para crear una política de IAM para acceder a los CloudWatch registros, utilice el comando create-policy:

Linux

```
aws iam create-policy \
          --policy-name kendra-cloudwatch-policy \
          --policy-document file://path/kendra-cloudwatch-policy.json
```

Donde:

- *path/* es la ruta del archivo a `kendra-cloudwatch-policy.json` en su dispositivo local.

macOS

```
aws iam create-policy \
          --policy-name kendra-cloudwatch-policy \
          --policy-document file://path/kendra-cloudwatch-policy.json
```

Donde:

- *path/* es la ruta del archivo a `kendra-cloudwatch-policy.json` en su dispositivo local.

Windows

```
aws iam create-policy ^
          --policy-name kendra-cloudwatch-policy ^
          --policy-document file://path/kendra-cloudwatch-policy.json
```

Donde:

- *path/* es la ruta del archivo a `kendra-cloudwatch-policy.json` en su dispositivo local.

c. Copie el nombre de recurso de Amazon (ARN) en el editor de texto y guárdelo de forma local como `kendra-cloudwatch-arn`.

> **Note**
>
> El ARN tiene un formato similar a *arn:aws:iam: :123456789012:role/*. kendra-cloudwatch-policy Necesita el ARN que guardó como `kendra-cloudwatch-arn` para asociar el `kendra-cloudwatch-policy` a su rol de IAM.

d. Para `kendra-cloudwatch-policy` asociarlo a su función de IAM, utilice el comando: attach-role-policy

Linux

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
```

```
        --role-name kendra-role
```

Donde:

- *policy-arn* es su archivo guardado `kendra-cloudwatch-arn`.

macOS

```
aws iam attach-role-policy \
        --policy-arn policy-arn \
        --role-name kendra-role
```

Donde:

- *policy-arn* es su archivo guardado `kendra-cloudwatch-arn`.

Windows

```
aws iam attach-role-policy ^
        --policy-arn policy-arn ^
        --role-name kendra-role
```

Donde:

- *policy-arn* es su archivo guardado `kendra-cloudwatch-arn`.

3. Para crear un índice, utilice el comando create-index:

Linux

```
aws kendra create-index \
       --name kendra-index \
       --edition DEVELOPER_EDITION \
       --role-arn role-arn \
       --region aws-region
```

Donde:

- *role-arn* es su `kendra-role-arn` guardado,

- *aws-region* es tu región. AWS

macOS

```
aws kendra create-index \
        --name kendra-index \
        --edition DEVELOPER_EDITION \
        --role-arn role-arn \
        --region aws-region
```

Donde:

- *role-arn* es su `kendra-role-arn` guardado,
- *aws-region* es tu región. AWS

Windows

```
aws kendra create-index ^
        --name kendra-index ^
        --edition DEVELOPER_EDITION ^
        --role-arn role-arn ^
        --region aws-region
```

Donde:

- *role-arn* es su `kendra-role-arn` guardado,
- *aws-region* es tu región. AWS

4. Copie el índice `Id` y guárdelo en un editor de texto como `kendra-index-id`. El `Id` ayuda a realizar el seguimiento del estado de creación del índice.
5. Para realizar un seguimiento del progreso de su trabajo de creación de índices, utilice el comando describe-index:

Linux

```
aws kendra describe-index \
        --id kendra-index-id \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, `kendra-index-id`
- *aws-region* es tu región. AWS

macOS

```
aws kendra describe-index \
        --id kendra-index-id \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, `kendra-index-id`
- *aws-region* es tu región. AWS

Windows

```
aws kendra describe-index ^
        --id kendra-index-id ^
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, `kendra-index-id`
- *aws-region* es tu región. AWS

El proceso de creación del índice tarda una media de 15 minutos, pero puede tardar más. Cuando el estado del índice es activo, su índice está listo para ser utilizado. Mientras se crea el índice, puede iniciar el siguiente paso.

Si lo utiliza AWS CLI en este paso, debe crear y adjuntar una política de IAM a su rol de IAM de Amazon Kendra que otorgue a su índice permisos para acceder a su bucket de S3.

Actualización del rol de IAM para el acceso a Amazon S3

Mientras se crea el índice, usted actualiza su rol de IAM en Amazon Kendra para permitir que el índice que creó lea los datos de su bucket de Amazon S3. Para obtener más información, consulte Roles de IAM para Amazon Kendra.

Para actualizar su rol de IAM (Consola)

1. Abra la consola de IAM en https://console.aws.amazon.com/iam/.
2. En el panel de navegación izquierdo, elija Roles e introduzca **`kendra-role`** en el cuadro de búsqueda situado encima del nombre del rol.
3. En las opciones sugeridas, haga clic en `kendra-role`.
4. En Resumen, seleccione Asociar políticas.
5. En Adjuntar permisos, en el cuadro de búsqueda, introduce **S3** y selecciona la casilla situada junto a la ReadOnlyAccess política de Amazon S3 entre las opciones sugeridas.
6. Elija Asociar política. En la página de resumen, ahora verá dos políticas asociadas al rol de IAM.
7. Regrese a la consola de Amazon Kendra en https://console.aws.amazon.com/kendra/ y espere a que el estado del índice cambie de Creando a Activo antes de continuar con el siguiente paso.

Para actualizar su rol de IAM (AWS CLI)

1. Guarde el siguiente texto en un archivo JSON llamado `kendra-S3-access-policy.json` en un editor de texto en su dispositivo local.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Action":[
            "s3:GetObject"
         ],
         "Resource":[
            "arn:aws:s3:::DOC-EXAMPLE-BUCKET/*"
         ],
         "Effect":"Allow"
      },
      {
         "Action":[
            "s3:ListBucket"
```

```
            ],
            "Resource":[
               "arn:aws:s3:::DOC-EXAMPLE-BUCKET"
            ],
            "Effect":"Allow"
         },
         {
            "Effect":"Allow",
            "Action":[
               "kendra:BatchPutDocument",
               "kendra:BatchDeleteDocument",
               "kendra:ListDataSourceSyncJobs"
            ],
            "Resource":[
               "arn:aws:kendra:aws-region:aws-account-id:index/kendra-index-id"
            ]
         }
      ]
}
```

Sustituya DOC-EXAMPLE-BUCKET por el nombre de su bucket de S3, aws-region por su *región*, por su ID de cuenta de 12 AWS AWS dígitos y *aws-account-id*por su archivo guardado. *kendra-index-id*kendra-index-id

2. Para crear una política de IAM para acceder a su bucket de S3, utilice el comando create-policy:

   Linux

   ```
   aws iam create-policy \
            --policy-name kendra-S3-access-policy \
            --policy-document file://path/kendra-S3-access-policy.json
   ```

   Donde:

   - *path/* es la ruta del archivo a kendra-S3-access-policy.json en su dispositivo local.

   macOS

   ```
   aws iam create-policy \
            --policy-name kendra-S3-access-policy \
   ```

```
        --policy-document file://path/kendra-S3-access-policy.json
```

Donde:

- *path/* es la ruta del archivo a `kendra-S3-access-policy.json` en su dispositivo local.

Windows

```
aws iam create-policy ^
        --policy-name kendra-S3-access-policy ^
        --policy-document file://path/kendra-S3-access-policy.json
```

Donde:

- *path/* es la ruta del archivo a `kendra-S3-access-policy.json` en su dispositivo local.

3. Copie el nombre de recurso de Amazon (ARN) en el editor de texto y guárdelo de forma local como `kendra-S3-access-arn`.

> **Note**
>
> El ARN tiene un formato similar a *arn:aws:iam::123456789012:role/kendra-S3-access-policy*. Necesita el ARN que guardó como `kendra-S3-access-arn` para asociar el `kendra-S3-access-policy` a su rol de IAM.

4. Para asociarlo `kendra-S3-access-policy` a su rol de IAM de Amazon Kendra, utilice el comando: attach-role-policy

Linux

```
aws iam attach-role-policy \
        --policy-arn policy-arn \
        --role-name kendra-role
```

Donde:

- *policy-arn* es su archivo guardado `kendra-S3-access-arn`.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

Donde:

- *policy-arn* es su archivo guardado `kendra-S3-access-arn`.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name kendra-role
```

Donde:

- *policy-arn* es su archivo guardado `kendra-S3-access-arn`.

Creación de campos de índice de búsqueda personalizados de Amazon Kendra

Para preparar a Amazon Kendra para que reconozca sus metadatos como atributos de documentos personalizados, debe crear campos personalizados correspondientes a los tipos de entidad de Amazon Comprehend. Puede introducir los siguientes nueve tipos de entidades de Amazon Comprehend como campos personalizados:

- COMMERCIAL_ITEM
- FECHA
- EVENT
- UBICACIÓN
- ORGANIZATION
- OTHER
- PERSON

- QUANTITY
- TITLE

> **Important**
>
> El índice no reconocerá los tipos de entidades mal escritos.

Para crear campos personalizados para su índice de Amazon Kendra (Consola)

1. Abra la consola Amazon Kendra en https://console.aws.amazon.com/kendra/.
2. En la lista de índices, haga clic en `kendra-index`.
3. En el panel de navegación izquierdo, en Administración de datos, elija Definición de faceta.
4. En el menú Campos de índice, seleccione Agregar campo.
5. En el cuadro de diálogo Agregar campo de índice, haga lo siguiente:
   a. En el campo Nombre, escriba **`COMMERCIAL_ITEM`**.
   b. En Tipo de datos, elija Lista de cadenas.
   c. En Tipos de uso, seleccione Facetable, Aparece en búsquedas y Visualizable y, a continuación, seleccione Agregar.
   d. Repita los pasos a a c para cada tipo de entidad de Amazon Comprehend: COMMERCIAL_ITEM, DATE, EVENT, LOCATION, ORGANIZATION, OTHER, PERSON, QUANTITY, TITLE.

La consola muestra los mensajes de adición de campos realizados correctamente. Puede optar por cerrarlos antes de continuar con el siguiente paso.

Para crear campos personalizados para su índice de Amazon Kendra (AWS CLI)

1. Guarde el siguiente texto en un archivo JSON llamado `custom-attributes.json` en un editor de texto en su dispositivo local.

```
[
   {
       "Name": "COMMERCIAL_ITEM",
       "Type": "STRING_LIST_VALUE",
       "Search": {
```

```
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "DATE",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "EVENT",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "LOCATION",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "ORGANIZATION",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "OTHER",
        "Type": "STRING_LIST_VALUE",
```

```
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "PERSON",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "QUANTITY",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    },
    {
        "Name": "TITLE",
        "Type": "STRING_LIST_VALUE",
        "Search": {
            "Facetable": true,
            "Searchable": true,
            "Displayable": true
        }
    }
]
```

2. Para crear campos personalizados en el índice, use el comando update-index:

   Linux

```
aws kendra update-index \
        --id kendra-index-id \
        --document-metadata-configuration-updates file://path/custom-
attributes.json \
```

```
        --region aws-region
```

Donde:

- *kendra-index-id*está guardado, `kendra-index-id`
- *path/* es la ruta del archivo a `custom-attributes.json` en su dispositivo local,
- *aws-region* es tu región. AWS

macOS

```
aws kendra update-index \
        --id kendra-index-id \
        --document-metadata-configuration-updates file://path/custom-
attributes.json \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, `kendra-index-id`
- *path/* es la ruta del archivo a `custom-attributes.json` en su dispositivo local,
- *aws-region* es tu región. AWS

Windows

```
aws kendra update-index ^
        --id kendra-index-id ^
        --document-metadata-configuration-updates file://path/custom-
attributes.json ^
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, `kendra-index-id`
- *path/* es la ruta del archivo a `custom-attributes.json` en su dispositivo local,
- *aws-region* es tu región. AWS

3. Para comprobar que los atributos personalizados se han añadido a su índice, utilice el comando describe-index:

Linux

```
aws kendra describe-index \
        --id kendra-index-id \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

macOS

```
aws kendra describe-index \
        --id kendra-index-id \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Windows

```
aws kendra describe-index ^
        --id kendra-index-id ^
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Agregar el bucket de Amazon S3 como origen de datos para el índice

Antes de poder sincronizar el índice, debe conectar el origen de datos de S3 a él.

Para conectar un bucket de S3 a su índice de Amazon Kendra (Consola)

1. Abra la consola Amazon Kendra en https://console.aws.amazon.com/kendra/.
2. En la lista de índices, haga clic en `kendra-index`.
3. En el menú de navegación de la izquierda, en Administración de datos, seleccione Origen de datos.
4. En la sección Seleccione el tipo de conector de origen de datos, vaya a Amazon S3 y elija Agregar conector.
5. En la página Especificar detalles, haga lo siguiente:
   a. En Nombre y descripción, para Nombre del origen de datos, introduzca **`S3-data-source`**.
   b. Mantenga la sección Descripción en blanco.
   c. Mantenga la configuración predeterminada para Etiquetas.
   d. Elija Siguiente.
6. En la página Configurar los ajustes de sincronización, en la sección Ámbito de sincronización, haga lo siguiente:
   a. En Introducir la ubicación del origen de datos, elija Examinar S3.
   b. En Elegir recursos, seleccione su bucket de S3 y, a continuación, seleccione Elegir.
   c. En la Ubicación de la carpeta de prefijos de los archivos de metadatos, elija Examinar S3.
   d. En Elegir recursos, haga clic en el nombre de su bucket en la lista de buckets.
   e. Para Objetos, seleccione la caja de opción para `metadata` y seleccione Elegir. El campo de ubicación ahora debería decir `metadata/`.
   f. Mantenga los ajustes predeterminados para la ubicación del archivo de configuración de la lista de control de acceso, la clave de descifrado seleccionada y la configuración adicional.
7. Para el rol de IAM, en la página Configurar los ajustes de sincronización, elija `kendra-role`.
8. En la página Configurar los ajustes de sincronización, en Sincronizar programación de ejecución, para Frecuencia, selecciona Ejecutar bajo demanda y, a continuación, selecciona Siguiente.
9. En la página Revisar y crear, revise sus opciones para los detalles del origen de datos y seleccione Añadir origen de datos.

Para conectar un bucket de S3 a su índice de Amazon Kendra (AWS CLI)

1. Guarde el siguiente texto en un archivo JSON llamado `S3-data-connector.json` en un editor de texto en su dispositivo local.

```
{
   "S3Configuration":{
      "BucketName":"DOC-EXAMPLE-BUCKET",
      "DocumentsMetadataConfiguration":{
         "S3Prefix":"metadata"
      }
   }
}
```

   Sustituya DOC-EXAMPLE-BUCKET por el nombre del bucket de S3.

2. Para conectar su bucket de S3 a su índice, utilice el create-data-sourcecomando:

   Linux

```
aws kendra create-data-source \
        --index-id kendra-index-id \
        --name S3-data-source \
        --type S3 \
        --configuration file://path/S3-data-connector.json \
        --role-arn role-arn \
        --region aws-region
```

   Donde:

   - *kendra-index-id*es tu guardadokendra-index-id,
   - *path/* es la ruta del archivo a `S3-data-connector.json` en su dispositivo local,
   - *role-arn* es su `kendra-role-arn` guardado,
   - *aws-region* es tu región. AWS

   macOS

```
aws kendra create-data-source \
        --index-id kendra-index-id \
        --name S3-data-source \
```

```
        --type S3 \
        --configuration file://path/S3-data-connector.json \
        --role-arn role-arn \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, `kendra-index-id`
- *path/* es la ruta del archivo a `S3-data-connector.json` en su dispositivo local,
- *role-arn* es su `kendra-role-arn` guardado,
- *aws-region* es tu región. AWS

Windows

```
aws kendra create-data-source ^
        --index-id kendra-index-id ^
        --name S3-data-source ^
        --type S3 ^
        --configuration file://path/S3-data-connector.json ^
        --role-arn role-arn ^
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, `kendra-index-id`
- *path/* es la ruta del archivo a `S3-data-connector.json` en su dispositivo local,
- *role-arn* es su `kendra-role-arn` guardado,
- *aws-region* es tu región. AWS

3. Copie el conector `Id` y guárdelo en un editor de texto como `S3-connector-id`. El `Id` le ayuda a rastrear el estado del proceso de conexión de datos.
4. Para asegurarse de que la fuente de datos de S3 se conectó correctamente, utilice el describe-data-sourcecomando:

   Linux

```
aws kendra describe-data-source \
        --id S3-connector-id \
        --index-id kendra-index-id \
```

```
        --region aws-region
```

Donde:

- *S3-connector-id* es su S3-connector-id guardado,
- *kendra-index-id*está guardadokendra-index-id,
- *aws-region* es tu región. AWS

macOS

```
aws kendra describe-data-source \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

Donde:

- *S3-connector-id* es su S3-connector-id guardado,
- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Windows

```
aws kendra describe-data-source ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

Donde:

- *S3-connector-id* es su S3-connector-id guardado,
- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Al final de este paso, el origen de datos de Amazon S3 se conecta al índice.

Sincronización del índice de Amazon Kendra

Con el origen de datos de Amazon S3 añadido, ahora puede sincronizar su índice de Amazon Kendra con él.

Para sincronizar su índice de Amazon Kendra (Consola)

1. Abra la consola Amazon Kendra en https://console.aws.amazon.com/kendra/.
2. En la lista de índices, haga clic en `kendra-index`.
3. En el menú de navegación izquierdo, elija Origen de datos.
4. En Origen de datos, seleccione `S3-data-source`.
5. En la barra de navegación superior, elija Sincronizar ahora.

Para sincronizar su índice de Amazon Kendra (AWS CLI)

1. Para sincronizar el índice, utilice el comando start-data-source-sync-job:

   Linux

   ```
   aws kendra start-data-source-sync-job \
           --id S3-connector-id \
           --index-id kendra-index-id \
           --region aws-region
   ```

   Donde:

   - *S3-connector-id* es su `S3-connector-id` guardado,
   - *kendra-index-id*es tu guardado`kendra-index-id`,
   - *aws-region* es tu región. AWS

   macOS

   ```
   aws kendra start-data-source-sync-job \
           --id S3-connector-id \
           --index-id kendra-index-id \
           --region aws-region
   ```

   Donde:

- *S3-connector-id* es su S3-connector-id guardado,
- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Windows

```
aws kendra start-data-source-sync-job ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

Donde:

- *S3-connector-id* es su S3-connector-id guardado,
- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

2. Para comprobar el estado de la sincronización del índice, utilice el comando list-data-source-sync-jobs:

Linux

```
aws kendra list-data-source-sync-jobs \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

Donde:

- *S3-connector-id* es su S3-connector-id guardado,
- *kendra-index-id*está guardadokendra-index-id,
- *aws-region* es tu región. AWS

macOS

```
aws kendra list-data-source-sync-jobs \
        --id S3-connector-id \
```

```
        --index-id kendra-index-id \
        --region aws-region
```

Donde:

- *S3-connector-id* es su S3-connector-id guardado,
- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Windows

```
aws kendra list-data-source-sync-jobs ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

Donde:

- *S3-connector-id* es su S3-connector-id guardado,
- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Al final de este paso, ha creado un índice de Amazon Kendra que permite realizar búsquedas y filtrar para su conjunto de datos.

Paso 5: Consulta del índice de Amazon Kendra

Su índice de Amazon Kendra ya está listo para consultas en lenguaje natural. Cuando busca en su índice, Amazon Kendra utiliza todos los datos y metadatos que ha proporcionado para devolver las respuestas más precisas a su consulta de búsqueda.

Hay tres tipos de consultas a las que Amazon Kendra puede responder:

- Consultas sobre hechos (preguntas sobre “quién”, “qué”, “cuándo” o “dónde”)
- Consultas descriptivas (preguntas sobre el “cómo”)
- Búsquedas de palabras clave (preguntas cuya intención y alcance no están claros)

Temas

- Consulta del índice de Amazon Kendra
- Filtrar los resultados de búsqueda

Consulta del índice de Amazon Kendra

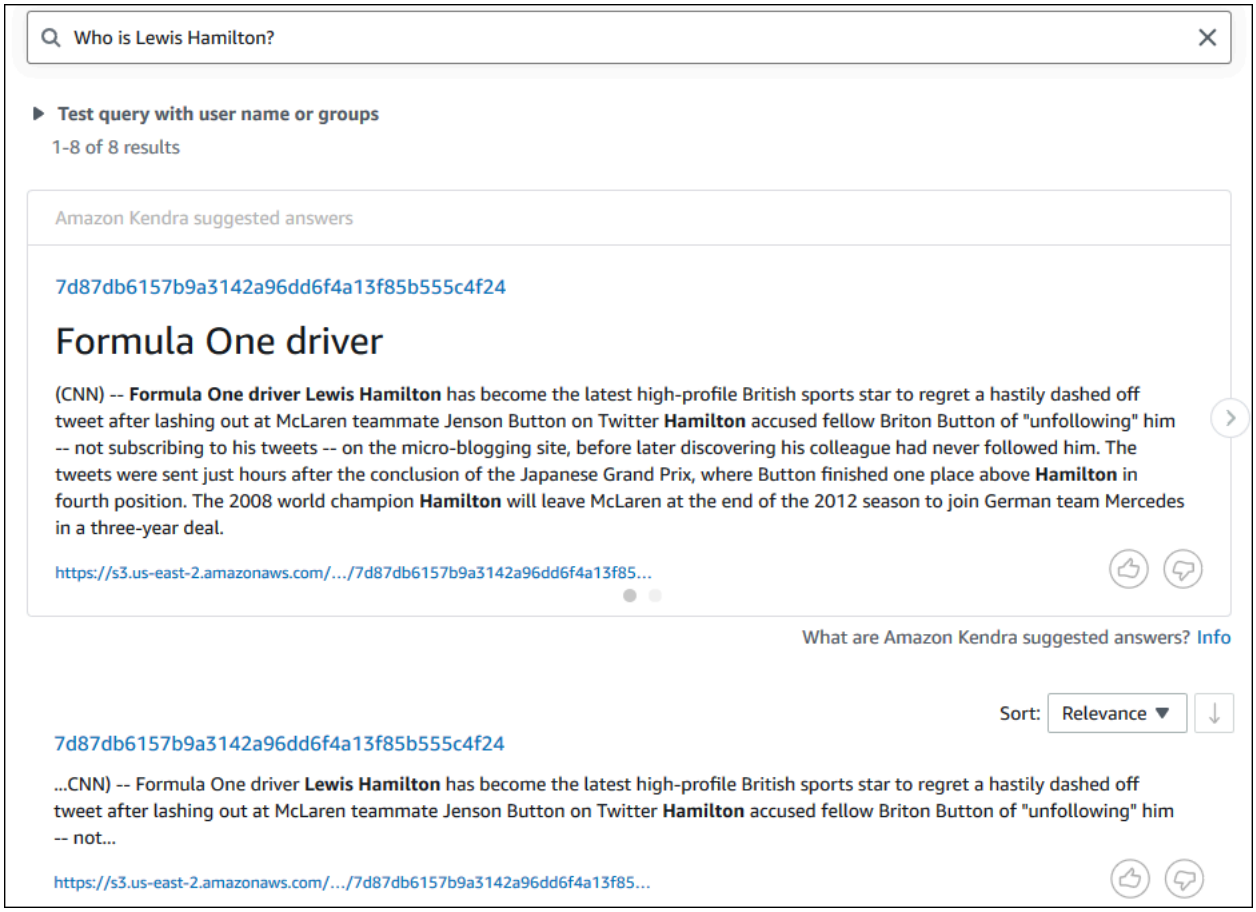
Puede consultar su índice de Amazon Kendra mediante preguntas que correspondan a los tres tipos de consultas que admite Amazon Kendra. Para más información, vea Consultas.

Las preguntas de ejemplo de esta sección se eligieron en función del conjunto de datos de muestra.

Para consultar su índice de Amazon Kendra (Consola)

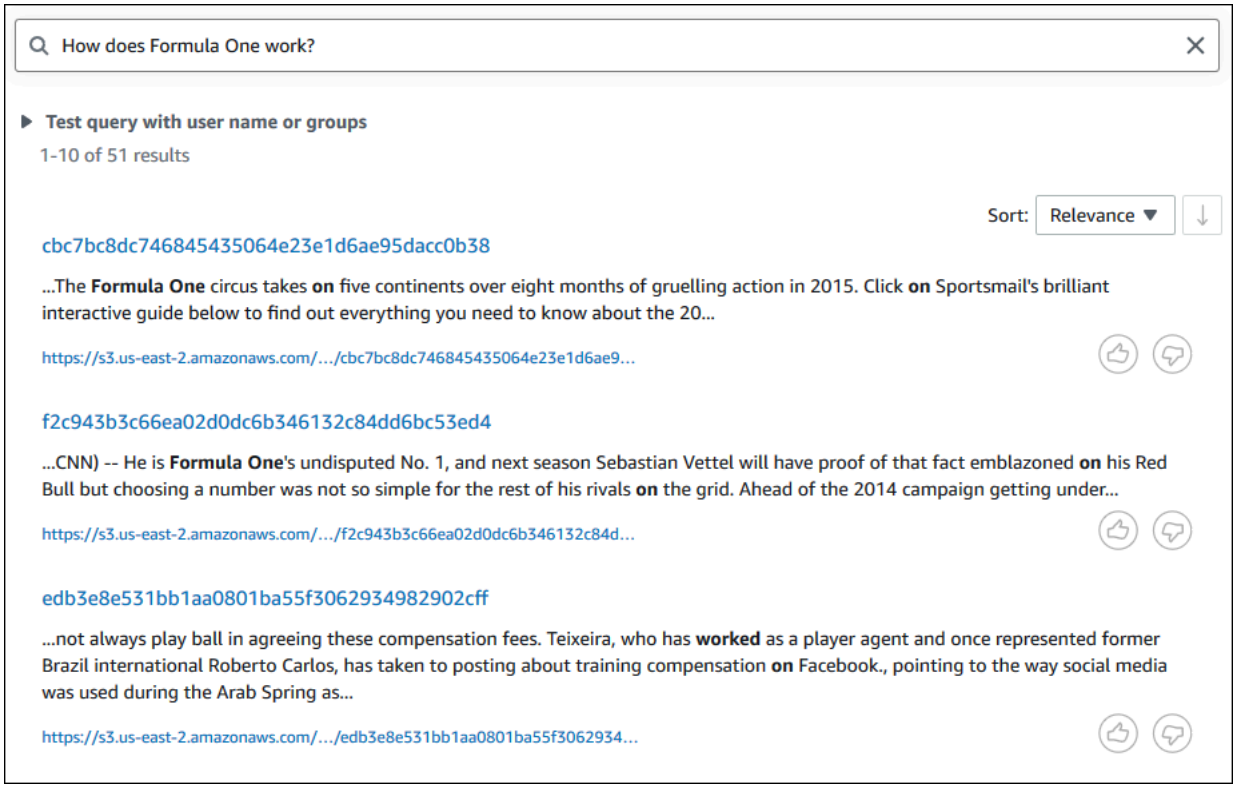
1. Abra la consola Amazon Kendra en https://console.aws.amazon.com/kendra/.
2. En la lista de índices, haga clic en `kendra-index`.
3. En el menú de navegación de la izquierda, elija la opción para buscar en su índice.
4. Para ejecutar un ejemplo de consulta de datos de muestra, escriba **`Who is Lewis Hamilton?`** en el cuadro de búsqueda y pulse Entrar.

   El primer resultado devuelto es la respuesta sugerida por Amazon Kendra, junto con el archivo de datos que contiene la respuesta. El resto de los resultados forman el conjunto de documentos recomendados.



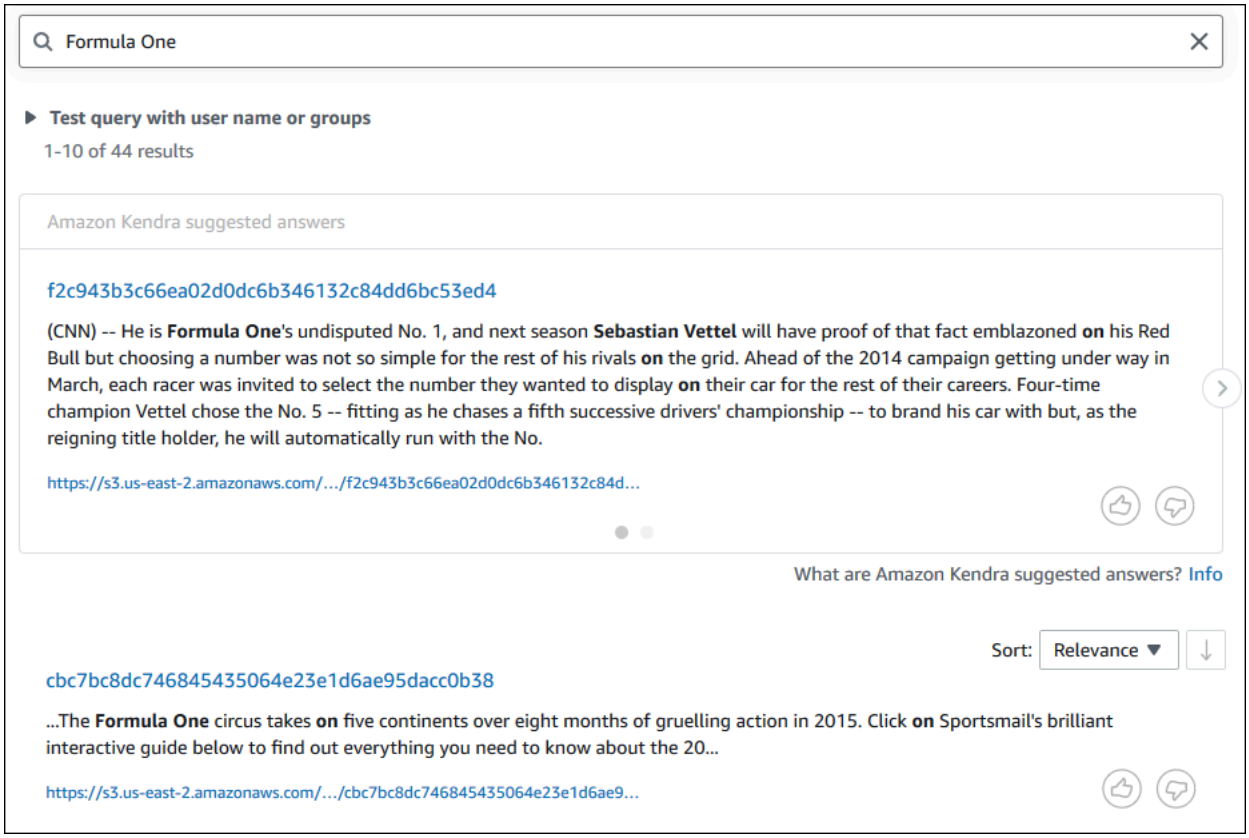
5. Para realizar una consulta descriptiva, introduzca **`How does Formula One work?`** en el cuadro de búsqueda y pulse Entrar.

   Verá otro resultado devuelto por la consola Amazon Kendra, esta vez con la frase relevante resaltada.



6. Para realizar una búsqueda de palabras clave, escriba **`Formula One`** en el cuadro de búsqueda y pulse Entrar.

   Verá otro resultado devuelto por la consola de Amazon Kendra, seguido de los resultados de todas las demás menciones de la frase en el conjunto de datos.



Para consultar su índice de Amazon Kendra (AWS CLI)

1. Para ejecutar una consulta sobre hechos de ejemplo, utilice el comando query:

   Linux

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Who is Lewis Hamilton?" \
           --region aws-region
   ```

   Donde:

   - *kendra-index-id*es tu salvado, kendra-index-id
   - *aws-region* es tu región. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Who is Lewis Hamilton?" \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Who is Lewis Hamilton?" ^
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

AWS CLI Muestra los resultados de la consulta.

2. Para ejecutar una consulta descriptiva de ejemplo, utilice el comando query:

Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "How does Formula One work?" \
        --region aws-region
```

Donde:

- *kendra-index-id*está guardadokendra-index-id,

- *aws-region* es tu región. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "How does Formula One work?" \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "How does Formula One work?" ^
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

AWS CLI Muestra los resultados de su consulta.

3. Para ejecutar un ejemplo de búsqueda por palabra clave, utilice el comando query:

Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Formula One" \
        --region aws-region
```

Donde:

- *kendra-index-id*está guardadokendra-index-id,
- *aws-region* es tu región. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Formula One" \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Formula One" ^
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

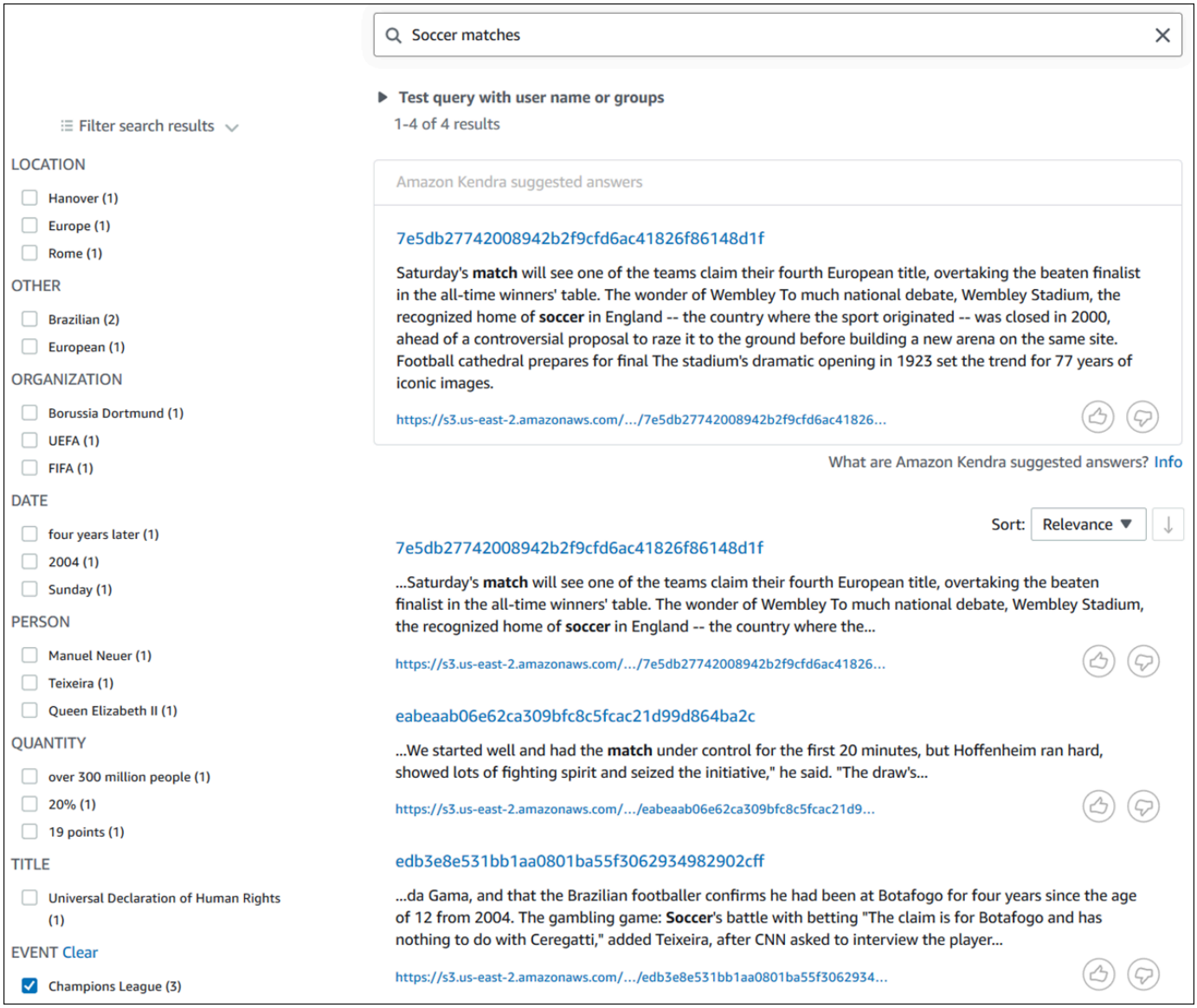
AWS CLI Muestra las respuestas devueltas a su consulta.

Filtrar los resultados de búsqueda

Puede filtrar y ordenar los resultados de la búsqueda mediante atributos de documento personalizados en la consola de Amazon Kendra. Para obtener más información sobre cómo Amazon Kendra procesa las consultas, consulte Filtrar consultas.

Para filtrar los resultados de la búsqueda (Consola)

1. Abra la consola Amazon Kendra en https://console.aws.amazon.com/kendra/.
2. En la lista de índices, haga clic en `kendra-index`.
3. En el menú de navegación de la izquierda, elija la opción para buscar en su índice.
4. En el cuadro de búsqueda, escriba **`Soccer matches`** como consulta y pulse Entrar.
5. En el menú de navegación de la izquierda, seleccione Filtrar resultados de búsqueda para ver una lista de facetas que puede utilizar para filtrar su búsqueda.
6. Seleccione la casilla “Liga de Campeones” en el subtítulo EVENTO, para ver los resultados de su búsqueda filtrados solo por los resultados que contengan “Liga de Campeones”.



Para filtrar los resultados de la búsqueda (AWS CLI)

1. Para ver las entidades de un tipo específico (por ejemplo, EVENT) que están disponibles para una búsqueda, utilice el comando query:

   Linux

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Soccer matches" \
           --facets '[{"DocumentAttributeKey":"EVENT"}]' \
           --region aws-region
   ```

   Donde:

   - *kendra-index-id*está guardadokendra-index-id,
   - *aws-region* es tu región. AWS

   macOS

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Soccer matches" \
           --facets '[{"DocumentAttributeKey":"EVENT"}]' \
           --region aws-region
   ```

   Donde:

   - *kendra-index-id*es tu salvado, kendra-index-id
   - *aws-region* es tu región. AWS

   Windows

   ```
   aws kendra query ^
           --index-id kendra-index-id ^
           --query-text "Soccer matches" ^
           --facets '[{"DocumentAttributeKey":"EVENT"}]' ^
           --region aws-region
   ```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

AWS CLI Muestra los resultados de la búsqueda. Para obtener una lista de las facetas del tipoEVENT, vaya a la sección «FacetResults» de la AWS CLI salida para ver una lista de las facetas filtrables con sus recuentos. Por ejemplo, una de las facetas es la “Liga de Campeones”.

> **Note**
>
> En lugar de EVENT, puede elegir cualquiera de los campos de índice que creó en the section called “Creación de un índice de Amazon Kendra” para el valor DocumentAttributeKey.

2. Para ejecutar la misma búsqueda pero filtrar solo por los resultados que contengan “Liga de Campeones”, utilice el comando query:

Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Soccer matches" \
        --attribute-filter '{"ContainsAny":{"Key":"EVENT","Value":
{"StringListValue":["Champions League"]}}}' \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu guardado, kendra-index-id
- *aws-region* es tu región. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Soccer matches" \
```

```
        --attribute-filter '{"ContainsAny":{"Key":"EVENT","Value":
{"StringListValue":["Champions League"]}}}' \
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Soccer matches" ^
        --attribute-filter '{"ContainsAny":{"Key":"EVENT","Value":
{"StringListValue":["Champions League"]}}}' ^
        --region aws-region
```

Donde:

- *kendra-index-id*es tu salvado, kendra-index-id
- *aws-region* es tu región. AWS

AWS CLI Muestra los resultados de búsqueda filtrados.

Paso 6: Limpieza

Limpieza de los archivos

Para dejar de incurrir en cargos en su AWS cuenta después de completar este tutorial, puede seguir los siguientes pasos:

1. Eliminar el bucket de Amazon S3

   Para obtener información acerca de cómo se elimina un bucket, consulte Eliminación de un bucket.

2. Elimine su índice de Amazon Kendra

Para obtener información sobre la eliminación de un índice de Amazon Kendra, consulte Eliminación de un índice.

3. Eliminar**`converter.py`**

   - Para la consola: ve a AWS CloudShelly asegúrate de que la región esté configurada como la tuya AWS . Una vez que se haya cargado el intérprete de comandos bash, escribe el siguiente comando en el entorno y pulsa Intro.

     ```
     rm converter.py
     ```

   - Para AWS CLI: ejecuta el siguiente comando en una ventana de terminal.

     Linux

     ```
     rm file/converter.py
     ```

     Donde:

     - *`file/`* es la ruta del archivo a `converter.py` en su dispositivo local.

     macOS

     ```
     rm file/converter.py
     ```

     Donde:

     - *`file/`* es la ruta del archivo a `converter.py` en su dispositivo local.

     Windows

     ```
     rm file/converter.py
     ```

     Donde:

     - *`file/`* es la ruta del archivo a `converter.py` en su dispositivo local.

Más información

Para obtener más información sobre la integración de Amazon Kendra en su flujo de trabajo, puede consultar las siguientes entradas de blog:

- Etiquetado de metadatos de contenido para una búsqueda mejorada

- Cree una solución de búsqueda inteligente con enriquecimiento de contenido automatizado

Para obtener más información sobre Amazon Comprehend, puede consultar la Guía del desarrollador de Amazon Comprehend.

Registro y monitoreo de Amazon Kendra

Temas

- Monitorización de su índice (consola)
- Registro de llamadas a la API de Amazon Kendra con registros AWS CloudTrail
- Registro de las llamadas a la API de Amazon Kendra Intelligent Ranking con registros AWS CloudTrail.
- Monitorización de Amazon Kendra con Amazon CloudWatch
- Monitorización de Amazon Kendra con Registros de Amazon CloudWatch

Monitorización de su índice (consola)

Utilice la consola de Amazon Kendra para supervisar el estado de los índices y los orígenes de datos. Puede utilizar esta información para realizar un seguimiento del tamaño y los requisitos de almacenamiento del índice y para supervisar el progreso y el éxito de la sincronización entre el índice y los orígenes de datos.

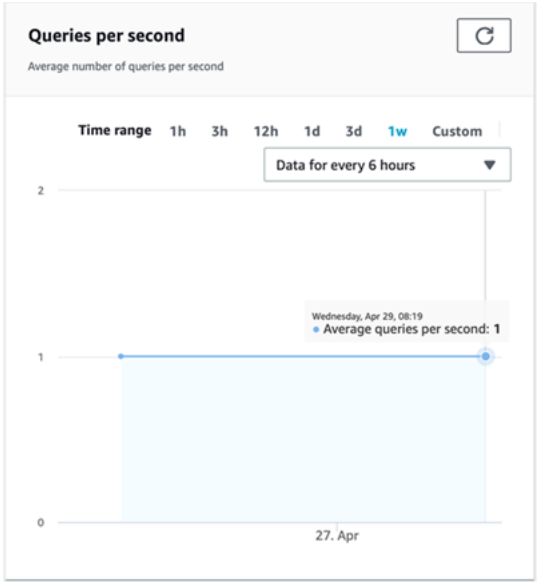
Para ver métricas de índices de (consola)

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon Kendra en https://console.aws.amazon.com/kendra/home.
2. En la lista de índices, seleccione el índice que desea visualizar.
3. Desplácese por la pantalla para ver las métricas del índice.

Puede ver las siguientes métricas sobre su índice.

- Recuento de documentos: el número total de documentos indexados. Incluye todos los documentos de todos los orígenes de datos. Utilice esta métrica para determinar si necesita comprar más o menos unidades de almacenamiento para su índice.

- Consultas por segundo: el número de consultas de índice que se solicitan cada segundo. Utilice esta métrica para determinar si necesita comprar más o menos unidades de consulta para su índice.



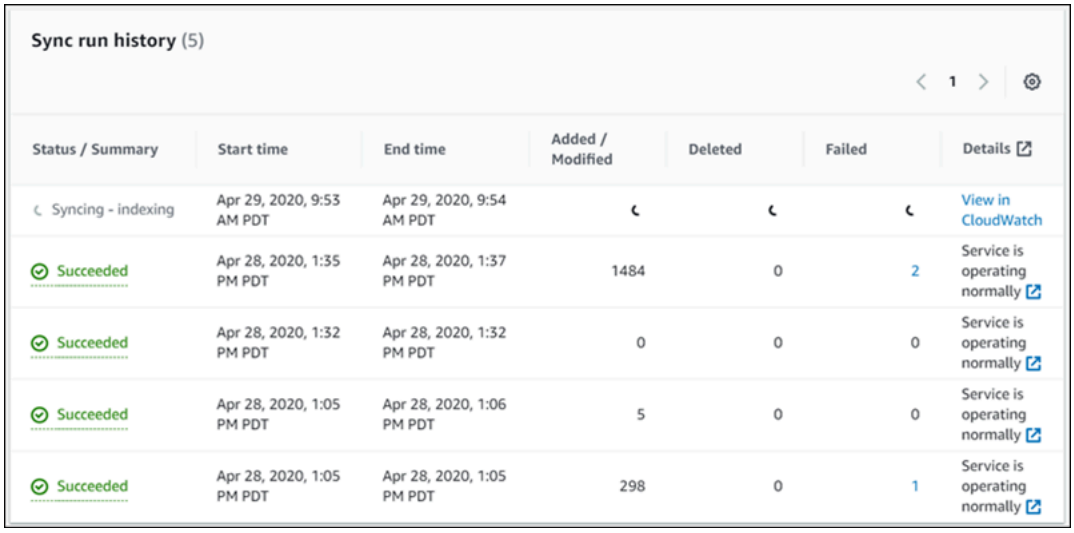
Para supervisar el progreso y el éxito de la sincronización entre el índice y un origen de datos, utilice la consola de Amazon Kendra. Utilice esta información para determinar el estado de su origen de datos.

Para ver las métricas de sincronización de (consola)

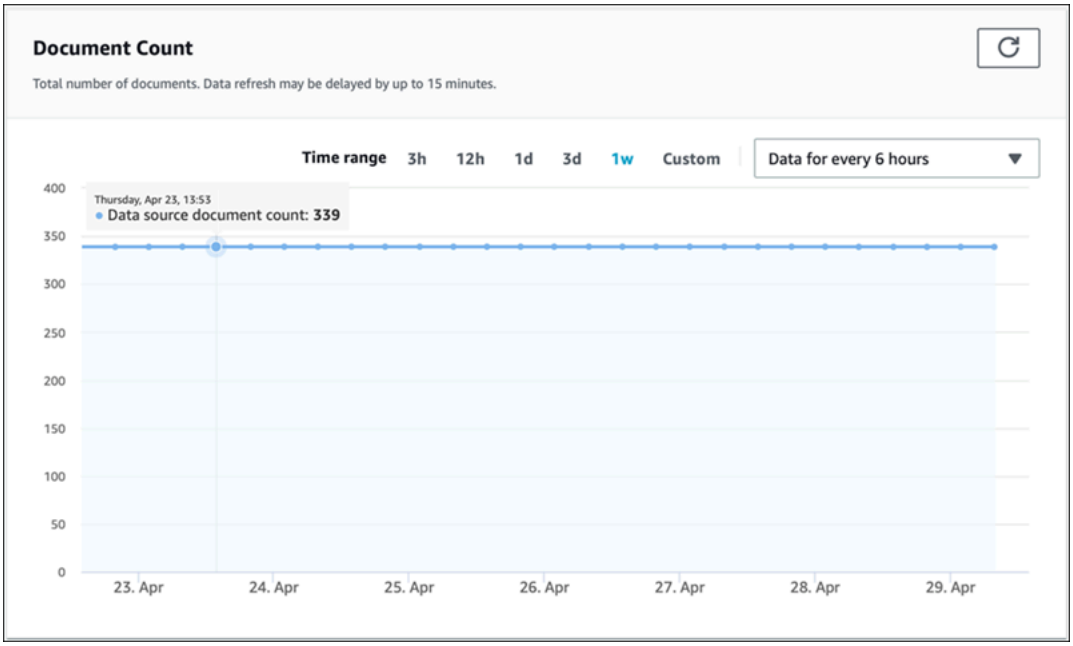
1. Inicie sesión en la AWS Management Console y abra la consola de Amazon Kendra en https://console.aws.amazon.com/kendra/home.
2. En la lista de índices, elija el índice para el que desea ver las métricas de sincronización.
3. En el menú izquierdo, elija Origen de datos.
4. En la lista de orígenes de datos, elija el origen de datos que desee visualizar.
5. Desplácese por la pantalla para ver las métricas de ejecución de la sincronización.

Puede ver la siguiente información.

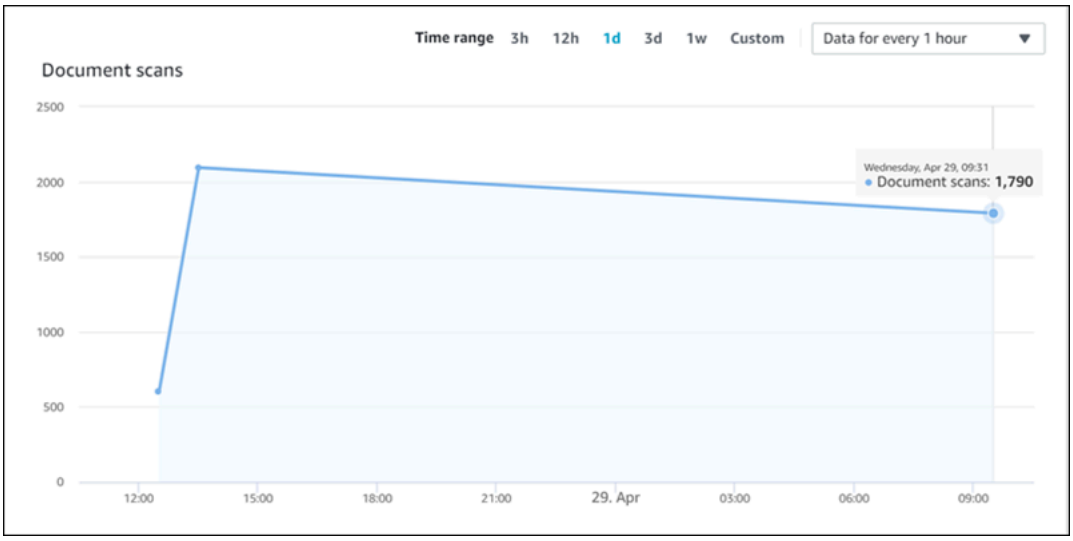
- Historial de ejecuciones de sincronización: estadísticas sobre la ejecución de la sincronización, incluida la hora de inicio y finalización, el número de documentos agregados, eliminados y fallidos. Si se produce un error en la ejecución de la sincronización, hay un enlace a CloudWatch Logs con más información. Seleccione el icono de configuración en la esquina superior izquierda para cambiar las columnas que se muestran en el historial. Utilice esta información para determinar el estado general del origen de datos.



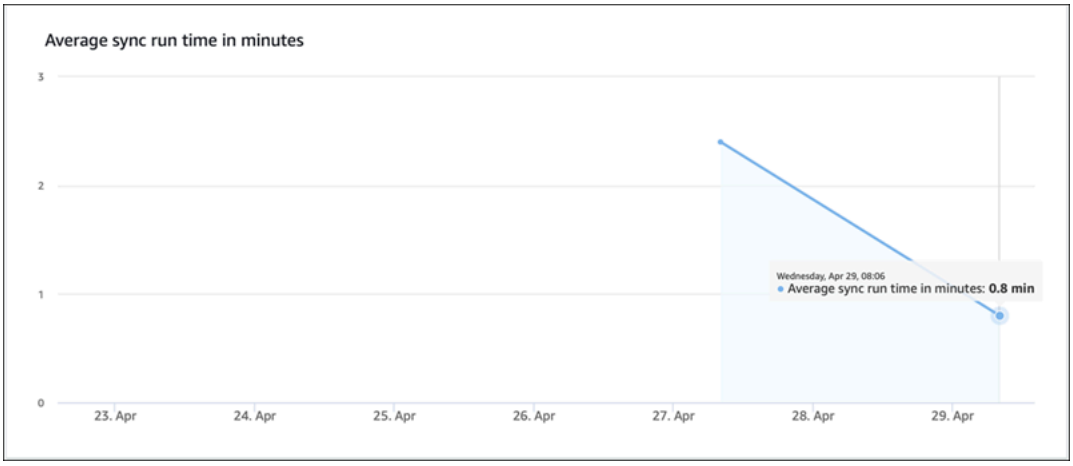
- Recuento de documentos: el número total de documentos indexados desde este origen de datos. Es el total de todos los documentos agregados al origen de datos menos el total de todos los documentos eliminados del origen de datos. Utilice esta información para determinar cuántos documentos de este origen de datos se incluyen en el índice.



- Escaneos de documentos: el número total de documentos escaneados durante la ejecución de la sincronización. Incluye todos los documentos del origen de datos, incluidos los que se agregaron, actualizaron, eliminaron o no cambiaron. Utilice esta información para determinar si Amazon Kendra escanea todos los documentos del origen de datos. La cantidad de documentos escaneados afecta al importe que se cobra por el servicio.



- Tiempo medio de ejecución de la sincronización en minutos: el tiempo medio que tarda una ejecución de sincronización en completarse. El tiempo que se tarda en sincronizar un origen de datos afecta al importe que se cobra por el servicio.



Registro de llamadas a la API de Amazon Kendra con registros AWS CloudTrail

Amazon Kendra se integra con AWS CloudTrail, un servicio que proporciona un registro de las acciones que realiza un usuario, un rol o un servicio de AWS en Amazon Kendra. CloudTrail obtiene todas las llamadas a la API de Amazon Kendra como eventos, incluidas las llamadas procedentes de la consola de Amazon Kendra y de las llamadas de código a las API de Amazon Kendra. Si crea un registro de seguimiento, puede habilitar la entrega continua de eventos de CloudTrail a un bucket de Amazon S3, incluidos los eventos de Amazon Kendra. Si no configura un registro de seguimiento, puede ver los eventos más recientes de la consola de CloudTrail en el Historial de eventos. Mediante la información recopilada por CloudTrail, puede determinar la solicitud que se realizó a Amazon Kendra, la dirección IP de origen desde la que se realizó, quién la realizó y cuándo, etc.

Para obtener más información acerca de CloudTrail, incluso cómo configurarlo y activarlo, consulte la Guía del usuario de AWS CloudTrail.

Información de Amazon Kendra en CloudTrail

CloudTrail se habilita en su cuenta AWS cuando crea la cuenta. Cuando se produce una actividad en Amazon Kendra, esta se registra en un evento de CloudTrail junto con los eventos de los demás servicios de AWS en Historial de eventos de CloudTrail. Puede ver, buscar y descargar los últimos eventos de la cuenta de AWS. Para obtener más información, consulte Ver eventos con el historial de eventos de CloudTrail.

Para mantener un registro continuo de los eventos de la cuenta de AWS, incluidos los eventos de Amazon Kendra, cree un registro de seguimiento. Un registro de seguimiento es una configuración

que permite a CloudTrail entregar eventos como archivos de registro a un bucket de S3 especificado. De manera predeterminada, cuando se crea un registro de seguimiento en la consola, el registro de seguimiento se aplica a todas las regiones de AWS. El registro de seguimiento registra los eventos de todas las regiones de la partición de AWS y envía los archivos de registro al bucket de S3 especificado. También es posible configurar otros servicios de AWS para analizar en profundidad y actuar en función de los datos de eventos recopilados en los registros de CloudTrail. Para obtener más información, consulte:

- Introducción a la creación de registros de seguimiento
- Servicios e integraciones compatibles con CloudTrail
- Configuración de notificaciones de Amazon SNS para CloudTrail
- Recibir archivos de registro de CloudTrail de varias regiones y Recibir archivos de registro de CloudTrail de varias cuentas

CloudTrail registra todas las acciones de Amazon Kendra, que se documentan en la referencia de la API. Por ejemplo, las llamadas a las operaciones `CreateIndex`, `CreateDataSource` y `Query` generan entradas en los archivos de registro de CloudTrail.

Cada entrada de registro o evento contiene información sobre quién generó la solicitud. Para obtener más información, consulte el Elemento userIdentity de CloudTrail.

Ejemplo: Entradas del archivo de registro de Amazon Kendra

Un registro de seguimiento es una configuración que permite la entrega de eventos como archivos de registro en el bucket de S3 especificado. Los archivos de registro de CloudTrail pueden contener una o varias entradas de registro. Un evento representa una solicitud específica realizada desde un origen y contiene información sobre la acción solicitada, la fecha y la hora de la acción, los parámetros de la solicitud, etc. Los archivos de registro de CloudTrail no rastrean el orden en la pila de las llamadas públicas a la API, por lo que estas no aparecen en ningún orden específico.

La llamada a la operación `Query` crea la siguiente entrada.

```
       {
           "eventVersion": "1.05",
           "userIdentity": {
               "type": "AssumedRole | FederatedUser | IAMUser | Root | SAMLUser |
WebIdentityUser",
               "principalId": "principal ID",
               "arn": "ARN",
```

```
                "accountId": "account ID",
                "accessKeyId": "access key ID",
                "sessionContext": {
                    "sessionIssuer": {
                        "type": "Role",
                        "principalId": "principal Id",
                        "arn": "ARN",
                        "accountId": "account ID",
                        "userName": "user name"
                    },
                    "webIdFederationData": {

                    },
                    "attributes": {
                        "mfaAuthenticated": false,
                        "creationDate": "timestamp"
                    }
                }
            },
            "eventTime": "timestamp",
            "eventSource": "kendra.amazonaws.com",
            "eventName": "Query",
            "awsRegion": "region",
            "sourceIPAddress": "source IP address",
            "userAgent": "user agent",
            "requestParameters": {
                "indexId": "index ID"
            },
            "responseElements": null,
            "requestID": "request ID",
            "eventID": "event ID",
            "eventType": "AwsApiCall",
            "recipientAccountId": "account ID"
        },
```

Registro de las llamadas a la API de Amazon Kendra Intelligent Ranking con registros AWS CloudTrail.

Amazon Kendra Intelligent Ranking se integra con AWS CloudTrail, un servicio que proporciona un registro de las acciones que realiza un usuario, un rol o un servicio de AWS en Amazon Kendra Intelligent Ranking. CloudTrail captura todas las llamadas a la API de Amazon Kendra Intelligent

Ranking como eventos, incluidas las llamadas de código a las API de Amazon Kendra Intelligent Ranking. Si crea un registro de seguimiento, puede habilitar la entrega continua de eventos de CloudTrail a un bucket de Amazon S3, incluidos los eventos de Amazon Kendra Intelligent Ranking. Si no configura un registro de seguimiento, puede ver los eventos más recientes de la consola de CloudTrail en el Historial de eventos. Mediante la información recopilada por CloudTrail, puede determinar la solicitud que se realizó a Amazon Kendra Intelligent Ranking, la dirección IP de origen desde la que se realizó, quién la realizó y cuándo, etc.

Para obtener más información acerca de CloudTrail, incluso cómo configurarlo y activarlo, consulte la Guía del usuario de AWS CloudTrail.

Información de Amazon Kendra Intelligent Ranking en CloudTrail

CloudTrail se habilita en su cuenta AWS cuando crea la cuenta. Cuando se produce una actividad en Amazon Kendra Intelligent Ranking, esta se registra en un evento de CloudTrail junto con los eventos de los demás servicios de AWS en Historial de eventos de CloudTrail. Puede ver, buscar y descargar los últimos eventos de la cuenta de AWS. Para obtener más información, consulte Ver eventos con el historial de eventos de CloudTrail.

Para mantener un registro continuo de los eventos de la cuenta de AWS, incluidos los eventos de Amazon Kendra Intelligent Ranking, cree un registro de seguimiento. Un registro de seguimiento es una configuración que permite a CloudTrail entregar eventos como archivos de registro a un bucket de S3 especificado. De manera predeterminada, cuando se crea un registro de seguimiento en la consola, el registro de seguimiento se aplica a todas las regiones de AWS. El registro de seguimiento registra los eventos de todas las regiones de la partición de AWS y envía los archivos de registro al bucket de S3 especificado. También es posible configurar otros servicios de AWS para analizar en profundidad y actuar en función de los datos de eventos recopilados en los registros de CloudTrail. Para obtener más información, consulte:

- Introducción a la creación de registros de seguimiento
- Servicios e integraciones compatibles con CloudTrail
- Configuración de notificaciones de Amazon SNS para CloudTrail
- Recibir archivos de registro de CloudTrail de varias regiones y Recibir archivos de registro de CloudTrail de varias cuentas

CloudTrail registra todas las acciones de Intelligent Ranking de Amazon Kendra, que se documentan en la referencia de la API. Por ejemplo, las llamadas a la `CreateRescoreExecutionPlan` generan entradas en los archivos de registro de CloudTrail.

Cada entrada de registro o evento contiene información sobre quién generó la solicitud. Para obtener más información, consulte el Elemento userIdentity de CloudTrail.

Ejemplo: Entradas del archivo de registro de Amazon Kendra Intelligent Ranking

Un registro de seguimiento es una configuración que permite la entrega de eventos como archivos de registro en el bucket de S3 especificado. Los archivos de registro de CloudTrail pueden contener una o varias entradas de registro. Un evento representa una solicitud específica realizada desde un origen y contiene información sobre la acción solicitada, la fecha y la hora de la acción, los parámetros de la solicitud, etc. Los archivos de registro de CloudTrail no rastrean el orden en la pila de las llamadas públicas a la API, por lo que estas no aparecen en ningún orden específico.

La llamada a la operación `CreateRescoreExecutionPlan` crea la siguiente entrada.

```
{
            "eventVersion": "1.08",
            "userIdentity": {
                "type": "AssumedRole",
                "principalId": "principal ID",
                "arn": "ARN",
                "accountId": "account ID",
                "accessKeyId": "access key ID",
                "sessionContext": {
                    "sessionIssuer": {
                        "type": "Role",
                        "principalId": "principal ID",
                        "arn": "ARN",
                        "accountId": "account ID",
                        "userName": "user name"
                    },
                    "webIdFederationData": {},
                    "attributes": {
                        "creationDate": "yyyy-mm-ddThh:mm:ssZ",
                        "mfaAuthenticated": "false"
                    }
                }
```

```
        },
        "eventTime": "yyyy-mm-ddThh:mm:ssZ",
        "eventSource": "kendra-ranking.amazonaws.com",
        "eventName": "CreateRescoreExecutionPlan",
        "awsRegion": "region",
        "sourceIPAddress": "source IP address",
        "userAgent": "user agent",
        "requestParameters": {
            "name": "name",
            "description": "description",
            "clientToken": "client token"
        },
        "responseElements": {
            "id": "rescore execution plan ID",
            "arn": "rescore execution plan ARN"
        },
        "requestID": "request ID",
        "eventID": "event ID",
        "readOnly": false,
        "eventType": "AwsApiCall",
        "managementEvent": true,
        "recipientAccountId": "account ID",
        "eventCategory": "Management",
        "tlsDetails": {
            "tlsVersion": "TLS version",
            "cipherSuite": "cipher suite",
            "clientProvidedHostHeader": "kendra-ranking.[region].api.aws"
        }
    }
```

Monitorización de Amazon Kendra con Amazon CloudWatch

Para realizar un seguimiento del estado de sus índices, utilice Amazon CloudWatch. Con CloudWatch, puede obtener métricas para la sincronización de documentos para su índice. También puede configurar alarmas de CloudWatch para recibir una notificación cuando una o varias métricas superen el umbral que defina. Por ejemplo, puede controlar el número de documentos enviados para indexarlos o el número de documentos que no se han podido indexar.

Debe tener los permisos de CloudWatch adecuados para supervisar Amazon Kendra con CloudWatch. Para obtener más información, consulte Autenticación y control de acceso para Amazon CloudWatch en la Guía del usuario de Amazon CloudWatch.

Visualización de métricas de Amazon Kendra

Vea las métricas de Amazon Kendra mediante la consola CloudWatch.

Pasos para ver las métricas (consola de CloudWatch)

1. Inicie sesión en la AWS Management Console y abra la consola de CloudWatch en https://console.aws.amazon.com/cloudwatch/.
2. Elija Métricas, elija Todas las métricas y después AWS/Lex.
3. Elija la dimensión, un nombre de métrica y, a continuación, Add to graph (Añadir al gráfico).
4. Elija un valor para el intervalo de fechas. El recuento de las métricas del intervalo de fechas seleccionado se muestra en el gráfico.

Creación de una alarma

Las alarmas de CloudWatch controlan una única métrica durante el periodo de tiempo especificado y realizan una o más acciones: enviar notificaciones a un tema de Amazon Simple Notification Service (Amazon SNS) o a una política de escalado automático. Las acciones se basan en el valor de la métrica con respecto a un umbral determinado durante los períodos de tiempo que se especifiquen. CloudWatch también puede enviar mensajes de Amazon SNS cuando la alarma cambia de estado.

Las alarmas de CloudWatch solo invocan acciones cuando el estado cambia y se mantiene durante el período especificado.

Para configurar una alarma

1. Inicie sesión en la AWS Management Console y abra la consola de CloudWatch en https://console.aws.amazon.com/cloudwatch/.
2. Elija Alarms (Alarmas) y, a continuación, seleccione Create Alarm (Crear alarma).
3. Seleccionar una métrica. Elija una métrica de Kendra para su índice y origen de datos. Configure también la hora como número establecido de horas, días, semanas o de forma personalizada.
4. Elija su estadística. Por ejemplo, Promedio. Elija también el período de activación de la alarma como un número fijo de minutos, horas, por día o personalizado.
5. Elija su umbral para activar la alarma, ya sea que utilice un valor estático o una banda, y la condición que debe cumplir con el umbral.

6. Elija el estado de alarma para el activador, ya sea que la métrica deba superar el umbral establecido o cualquier otro estado. Seleccione a quién o qué correo electrónico desea enviar la notificación de alarma.
7. Si está satisfecho con la alarma, elija Crear alarma.

> **Note**
>
> Debe proporcionar un nombre para la alarma de CloudWatch.

Métricas de CloudWatch para trabajos de sincronización de índices

En la siguiente tabla se describen las métricas de Amazon Kendra para trabajos de sincronización de Origen de datos.

Si usa la API o la CLI, debe especificar "AWS/Kendra" `Namespace` además de la que prefiera cuando utilice la API `MetricName` GetMetricStatistics.

Métrica	Descripción
`DocumentsCrawled`	El número de documentos que el trabajo de sincronización escaneó o descubrió durante la ejecución. Dimensiones: • IndexId • DataSourceId Unidad: recuento
`DocumentsSubmittedForIndexing`	El número de documentos que el trabajo de sincronización ha enviado al índice. Dimensiones: • IndexId • DataSourceId

Métrica	Descripción
	Unidad: recuento
DocumentsSubmittedForIndexingFailed	El número de documentos que no se han podido indexar. Consulte el contenido del registro de CloudWatch del trabajo de sincronización para obtener más información. Dimensiones: • IndexId • DataSourceId Unidad: recuento
DocumentsSubmittedForDeletion	El número de documentos que el trabajo de sincronización ha solicitado eliminar del índice. Dimensiones: • IndexId • DataSourceId Unidad: recuento
DocumentsSubmittedForDeletionFailed	Número de documentos que no se han podido eliminar. Consulte el contenido del registro de CloudWatch del trabajo de sincronización para obtener más información. Dimensiones: • IndexId • DataSourceId Unidad: recuento

Métricas para Origen de datos de Amazon Kendra

En la siguiente tabla se describen las métricas de Amazon Kendra para trabajos de sincronización de Origen de datos. Las métricas marcadas con un asterisco (*) se utilizan únicamente para Origen de datos de Amazon S3.

Si usa la API o la CLI, debe especificar “AWS/Kendra” `Namespace` además de la que prefiera cuando utilice la API `MetricName` GetMetricStatistics.

Métrica	Descripción
`DocumentsSkippedNoChange *`	El número de documentos examinados y comprobados que no han cambiado, por lo que no se enviaron para su indexación. Dimensiones: • IndexId • DataSourceId Unidad: recuento
`DocumentsSkippedInvalidMetadata *`	El número de documentos omitidos porque se produjo un problema con el archivo de metadatos asociado. Consulte el contenido del registro de CloudWatch de la ejecución de sincronización para obtener más información. Dimensiones: • IndexId • DataSourceId Unidad: recuento
`DocumentsCrawled`	El número de archivos de documentos examinados.

Métrica	Descripción
	Dimensiones: • IndexId • DataSourceId Unidad: recuento
DocumentsSubmittedForDeletion	El número de documentos examinados que se eliminaron del origen de datos y se enviaron para su eliminación. Dimensiones: • IndexId • DataSourceId Unidad: recuento
DocumentsSubmittedForDeletionFailed	El número de documentos de un origen de datos que no se han podido eliminar. Dimensiones: • IndexId • DataSourceId Unidad: recuento

Métrica	Descripción
`DocumentsSubmittedForIndexing`	El número de documentos examinados y enviados para su indexación. Dimensiones: • IndexId • DataSourceId Unidad: recuento
`DocumentsSubmittedForIndexingFailed`	El número de documentos presentados para su indexación que no se pudieron indexar. Dimensiones: • IndexId • DataSourceId Unidad: recuento

Métricas de los documentos indexados

En la siguiente tabla se describen las métricas de Amazon Kendra para documentos indexados. Para los documentos que se indexan mediante la operación BatchPutDocument, solo se admite la dimensión `IndexId`.

Si usa la API o la CLI, debe especificar “AWS/Kendra” `Namespace` además de la que prefiera cuando utilice la API `MetricName` GetMetricStatistics.

Métrica	Descripción
`DocumentsIndexed`	El número de documentos indexados. Dimensiones: • IndexId

Métrica	Descripción
	• DataSourceId Unidad: recuento
DocumentsFailedToIndex	El número de documentos que no se han podido indexar. Compruebe el contenido del registro de CloudWatch para más detalles. Dimensiones: • IndexId • DataSourceId Unidad: recuento
IndexQueryCount	El número de consultas de índice por minuto. Dimensiones: • IndexId Unidad: recuento

Monitorización de Amazon Kendra con Registros de Amazon CloudWatch

Amazon Kendra utiliza Amazon CloudWatch Logs para proporcionarle información sobre el funcionamiento de sus orígenes de datos. Amazon Kendra registra los detalles del proceso de los documentos a medida que se indexan. Registra los errores del origen de datos que se producen mientras se indexan los documentos. Usted utiliza CloudWatch Logs para supervisar, almacenar y acceder a los archivos de registro.

CloudWatch Logs almacena los eventos de registro en una secuencia de registros que forma parte de un grupo de registros. Amazon Kendra utiliza estas funciones de la siguiente manera:

- Grupos de registros: Amazon Kendra almacena todos los flujos de registros en un único grupo de registros para cada índice. Amazon Kendra crea el grupo de registros cuando se crea el índice. El identificador del grupo de registros siempre comienza por “aws/kendra/”.
- Flujo de registro: Amazon Kendra crea un nuevo flujo de registro de origen de datos en el grupo de registros para cada trabajo de sincronización de índices que ejecute. También crea un nuevo flujo de registro de documentos cuando un flujo alcanza aproximadamente 500 entradas.
- Entradas de registro: Amazon Kendra crea una entrada de registro en el flujo de registro a medida que indexa los documentos. Cada entrada proporciona información sobre el procesamiento del documento o sobre cualquier error que se produzca.

Para obtener más información sobre el uso de CloudWatch Logs, consulte Qué es Amazon Cloud Watch Logs en la Guía del usuario de Amazon Cloud Watch Logs.

Amazon Kendra crea dos tipos de flujos de registro:

- Flujos de registro de Origen de datos
- Flujo de registro de documentos

Flujos de registro de Origen de datos

Los flujos de registro de Origen de datos publican entradas sobre sus trabajos de sincronización de índices. Cada trabajo de sincronización crea un nuevo flujo de registro que se utiliza para publicar las entradas. El nombre del flujo de registro es:

```
data source id/YYYY-MM-DD-HH/data source sync job ID
```

Se crea un nuevo flujo de registro para cada trabajo de sincronización que se ejecute.

Hay tres tipos de mensajes de registro publicados en un flujo de registro de un origen de datos:

- Un mensaje de registro de un documento que no se pudo enviar para su indexación. A continuación, se muestra un ejemplo de este mensaje para un documento de un origen de datos de S3:

```
{
    "DocumentId": "document ID",
    "S3Path": "s3://bucket/prefix/object",
    "Message": "Failed to ingest document via BatchPutDocument.",
```

```
    "ErrorCode": "InvalidRequest",
    "ErrorMessage": "No document metadata configuration found for document attribute
 key  city."
}
```

- Mensaje de registro de un documento que no se ha podido enviar para su eliminación. A continuación se muestra un ejemplo de este mensaje:

```
{
    "DocumentId": "document ID",
    "Message": "Failed to delete document via BatchDeleteDocument.",
    "ErrorCode": "InvalidRequest",
    "ErrorMessage": "Document can't be deleted because it doesn't exist."
}
```

- Un mensaje de registro cuando se encuentra un archivo de metadatos no válido para un documento en un bucket de Amazon S3. A continuación se muestra un ejemplo de este mensaje.

```
{
    "Message": "Found invalid metadata
 file bucket/prefix/filename.extension.metadata.json."
}
```

- Para los conectores de bases de datos y SharePoint, Amazon Kendra solo escribe mensajes en el flujo de registro si un documento no se puede indexar. A continuación, se muestra un ejemplo del mensaje de error que Amazon Kendra registra.

```
{
    "DocumentID": "document ID",
    "IndexID": "index ID",
    "SourceURI": "",
    "CrawlStatus": "FAILED",
    "ErrorCode": "403",
    "ErrorMessage": "Access Denied",
    "DataSourceErrorCode": "403"
}
```

Flujo de registro de documentos

Amazon Kendra registra información sobre el procesamiento de documentos mientras se indexan. Registro de un conjunto de mensajes para documentos almacenados en un origen de datos de

Amazon S3. Registra los errores solo en los documentos almacenados en un origen de datos de Microsoft SharePoint o de una base de datos.

Si los documentos se agregaron al índice mediante la operación BatchPutDocument, el flujo de registro se denomina de la siguiente manera:

```
YYYY-MM-DD-HH/UUID
```

Si los documentos se agregaron al índice mediante un origen de datos, el flujo de registro se denomina de la siguiente manera:

```
dataSourceId/YYYY-MM-DD-HH/UUID
```

Cada flujo de registro contiene hasta 500 mensajes.

Si se produce un error al indexar un documento, se envía este mensaje al flujo de registro:

```
{
    "DocumentId": "document ID",
    "IndexName": "index name",
    "IndexId": "index ID"
    "SourceURI": "source URI"
    "IndexingStatus": "DocumentFailedToIndex",
    "ErrorCode": "400 | 500",
    "ErrorMessage": "message"
}
```

Seguridad en Amazon Kendra

La seguridad en la nube AWS es la máxima prioridad. Como AWS cliente, usted se beneficia de una arquitectura de centro de datos y red diseñada para cumplir con los requisitos de las organizaciones más sensibles a la seguridad.

La seguridad es una responsabilidad compartida entre usted AWS y usted. El modelo de responsabilidad compartida la describe como seguridad de la nube y seguridad en la nube:

- Seguridad de la nube:AWS es responsable de proteger la infraestructura que ejecuta AWS los servicios en la AWS nube. AWS también le proporciona servicios que puede utilizar de forma segura. Los auditores externos prueban y verifican periódicamente la eficacia de nuestra seguridad como parte de los AWS programas de de . Para obtener más información sobre los programas de conformidad que se aplican a Amazon Kendra, consulte AWS Servicios dentro del alcance por programa de conformidad Servicios incluidos en el ámbito de aplicación por programa AWS .
- Seguridad en la nube: su responsabilidad viene determinada por el AWS servicio que utilice. Usted también es responsable de otros factores, incluida la confidencialidad de los datos, los requisitos de la empresa y la legislación y los reglamentos aplicables.

Esta documentación le permite comprender cómo aplicar el modelo de responsabilidad compartida cuando se utiliza Amazon Kendra. En los siguientes temas, se mostrará cómo configurar Amazon Kendra para satisfacer sus objetivos de seguridad y conformidad. También aprenderá a utilizar otros AWS servicios que le ayudan a supervisar y proteger sus recursos de Amazon Kendra.

Temas

- Protección de los datos en Amazon Kendra
- Amazon Kendra Amazon Kendra Intelligent Ranking y puntos de conexión de VPC de interfaz ()AWS PrivateLink
- Administración de identidades y accesos para Amazon Kendra
- Prácticas recomendadas de seguridad
- Registro y monitoreo en Amazon Kendra
- Validación de la conformidad de Amazon Kendra
- Resiliencia en Amazon Kendra
- Seguridad de infraestructuras en Amazon Kendra

- Análisis de configuración y vulnerabilidad en AWS Identity and Access Management

Protección de los datos en Amazon Kendra

El modelo de se aplica a protección de datos en Amazon Kendra. Como se describe en este modelo, AWS es responsable de proteger la infraestructura global en la que se ejecutan todos los Nube de AWS. Usted es responsable de mantener el control sobre el contenido alojado en esta infraestructura. Usted también es responsable de las tareas de administración y configuración de seguridad para los Servicios de AWS que utiliza. Para obtener más información sobre la privacidad de los datos, consulte las Preguntas frecuentes sobre la privacidad de datos. Para obtener información sobre la protección de datos en Europa, consulte la publicación de blog sobre el Modelo de responsabilidad compartida de AWS y GDPR en el Blog de seguridad de AWS .

Con fines de protección de datos, le recomendamos que proteja Cuenta de AWS las credenciales y configure los usuarios individuales con AWS IAM Identity Center o AWS Identity and Access Management (IAM). De esta manera, solo se otorgan a cada usuario los permisos necesarios para cumplir sus obligaciones laborales. También recomendamos proteger sus datos de la siguiente manera:

- Utilice la autenticación multifactor (MFA) en cada cuenta.
- Utilice SSL/TLS para comunicarse con los recursos. AWS Se recomienda el uso de TLS 1.2 y recomendamos TLS 1.3.
- Configure la API y el registro de actividad de los usuarios con. AWS CloudTrail
- Utilice soluciones de AWS cifrado, junto con todos los controles de seguridad predeterminados Servicios de AWS.
- Utilice servicios de seguridad administrados avanzados, como Amazon Macie, que lo ayuden a detectar y proteger los datos confidenciales almacenados en Amazon S3.
- Si necesita módulos criptográficos validados por FIPS 140-2 para acceder a AWS través de una interfaz de línea de comandos o una API, utilice un punto final FIPS. Para obtener más información sobre los puntos de conexión de FIPS disponibles, consulte Estándar de procesamiento de la información federal (FIPS) 140-2.

Se recomienda encarecidamente no introducir nunca información confidencial o sensible, como, por ejemplo, direcciones de correo electrónico de clientes, en etiquetas o campos de formato libre, tales como el campo Nombre. Esto incluye cuando trabaja con Amazon Kendra u otro dispositivo Servicios de AWS mediante la consola, la API o AWS los AWS CLI SDK. Cualquier dato que

ingrese en etiquetas o campos de formato libre utilizados para nombres se puede emplear para los registros de facturación o diagnóstico. Si proporciona una URL a un servidor externo, recomendamos encarecidamente que no incluya la información de las credenciales en la URL para validar la solicitud para ese servidor.

Cifrado en reposo

Amazon Kendra cifra los datos en reposo con la clave de cifrado que elija. Puede elegir una de las siguientes opciones:

- Una clave de AWS KMS de su propiedad AWS . Si no especifica una clave de cifrado, los datos se cifran con esta clave de forma predeterminada.
- Una clave KMS AWS administrada en tu cuenta. Amazon Kendra crea, administra y utiliza esta clave en su nombre. El nombre de la clave es `aws/kendra`.
- Una clave administrada por el cliente. Puede proporcionar el ARN de una clave de cifrado que haya creado en su cuenta. Cuando utilice una clave de KMS gestionada por el cliente, debe asignar a la clave una política de claves que permita a Amazon Kendra utilizarla. Seleccione una clave de KMS gestionada por el cliente de cifrado simétrico, Amazon Kendra no admite claves de KMS asimétricas. Para obtener más información, consulte Administración de claves.

Cifrado en tránsito

Amazon Kendra utiliza el protocolo HTTPS para comunicarse con la aplicación cliente. Utiliza HTTPS y AWS firmas para comunicarse con otros servicios en nombre de tu aplicación. Si utiliza una VPC, puede utilizarla AWS PrivateLink para establecer una conexión privada entre su VPC y Amazon Kendra.

Administración de claves

Amazon Kendra cifra el contenido del índice mediante uno de los tres tipos de claves. Puede elegir una de las siguientes opciones:

- Un AWS KMS de su propiedad. AWS Esta es la opción predeterminada.
- Una clave AWS de KMS administrada. Amazon Kendra crea esta clave en su cuenta y la administra y utiliza en su nombre.
- Una clave de KMS administrada por el cliente. Puede crear la clave al crear un índice o un origen de datos de Amazon Kendra o puede crear la clave mediante la consola de AWS KMS .

Seleccione una clave de KMS gestionada por el cliente de cifrado simétrico. Amazon Kendra no es compatible con claves de KMS asimétricas. Para obtener más información, consulte Uso de claves simétricas y asimétricas en la Guía para desarrolladores de AWS Key Management Service.

Amazon Kendra Amazon Kendra Intelligent Ranking y puntos de conexión de VPC de interfaz ()AWS PrivateLink

Puede establecer una conexión privada entre la VPC y Amazon Kendra mediante la creación de un punto de conexión de VPC de interfaz. Los puntos de enlace de la interfaz funcionan con una tecnología que le permite acceder de forma privada a las API de Amazon Kendra sin una puerta de enlace a Internet, un dispositivo NAT, una conexión VPN o una conexión AWS Direct Connect. AWS PrivateLink Las instancias de la VPC no necesitan direcciones IP públicas para comunicarse con las API de Amazon Kendra. El tráfico entre su VPC y Amazon Kendra no sale de la red de Amazon.

Cada punto de conexión de la interfaz está representado por una o más interfaces de red elásticas en las subredes.

Consideraciones sobre los puntos de enlace de VPC de Amazon Kendra y Amazon Kendra Intelligent Ranking

Antes de configurar un punto final de interfaz de VPC para Amazon Kendra o Amazon Kendra Intelligent Ranking, asegúrese de revisar los requisitos previos en la Guía del usuario de Amazon VPC.

Amazon Kendra y Amazon Kendra Intelligent Ranking permiten realizar llamadas a todas sus acciones de API desde su VPC.

Creación de un punto final de VPC de interfaz para Amazon Kendra y Amazon Kendra Intelligent Ranking

Puede crear un punto de conexión de VPC para el servicio Amazon Kendra o Amazon Kendra Intelligent Ranking mediante la consola de Amazon VPC o el (). AWS Command Line Interface AWS CLI

Cree un punto de conexión de VPC para Amazon Kendra mediante el siguiente nombre de servicio:

- com.amazonaws.*region*.kendra

Cree un punto de enlace de VPC para Amazon Kendra Intelligent Ranking con el siguiente nombre de servicio:

- aws.api.*region*.kendra-ranking

Tras crear un punto de enlace de VPC, puede utilizar el siguiente AWS CLI comando de ejemplo que utiliza el `endpoint-url` parámetro para especificar un punto de enlace de interfaz para la API de Amazon Kendra:

```
aws kendra list-indices --endpoint-url https://VPC endpoint
```

El *punto final de la VPC* es el nombre DNS que se genera cuando se crea el punto final de la interfaz. Este nombre incluye el ID del punto de conexión de la VPC y el nombre del servicio de Amazon Kendra, que incluye la región. Por ejemplo, `vpce-1234-abcdef.kendra.us-west-2.vpce.amazonaws.com`.

Si activa el DNS privado para el punto final, puede realizar solicitudes de API a Amazon Kendra utilizando su nombre de DNS predeterminado para la región. Por ejemplo, `kendra.us-east-1.amazonaws.com`.

Para más información, consulte Creación de un punto de conexión de interfaz en la Guía del usuario de Amazon VPC.

Creación de una política de puntos de conexión de VPC para Amazon Kendra y Amazon Kendra Intelligent Ranking

Puede adjuntar una política de punto final a su punto de enlace de VPC que controle el acceso a Amazon Kendra o Amazon Kendra Intelligent Ranking.

La política de Amazon Kendra o Amazon Kendra Intelligent Ranking especifica la siguiente información:

- El usuario principal/autorizado que puede realizar acciones.
- Las acciones que se pueden realizar.
- Los recursos en los que se pueden llevar a cabo las acciones.

Ejemplo: política de punto de conexión de VPC para acciones de Amazon Kendra

A continuación, se muestra un ejemplo de una política de punto de conexión para Amazon Kendra. Cuando se adjunta a un punto final, esta política otorga acceso a todas las acciones de Amazon Kendra disponibles a todos los principales o usuarios autorizados de todos los recursos.

```
{
   "Statement":[
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "kendra:*"
         ],
         "Resource":"*"
      }
   ]
}
```

Ejemplo: política de puntos de conexión de VPC para las acciones de Amazon Kendra Intelligent Ranking

El siguiente es un ejemplo de una política de puntos finales para Amazon Kendra Intelligent Ranking. Cuando se adjunta a un punto final, esta política otorga acceso a todas las acciones de Amazon Kendra Intelligent Ranking disponibles a todos los principales o usuarios autorizados de todos los recursos.

```
{
   "Statement":[
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "kendra-ranking:*"
         ],
         "Resource":"*"
      }
   ]
}
```

Para obtener más información, consulte Control del acceso a los puntos de enlace de la VPC mediante políticas de puntos de enlace en la Guía del usuario de Amazon VPC.

Administración de identidades y accesos para Amazon Kendra

AWS Identity and Access Management (IAM) es una herramienta Servicio de AWS que ayuda al administrador a controlar de forma segura el acceso a los recursos. AWS Los administradores de IAM controlan quién se puede autenticar (iniciar sesión) y autorizar (tener permisos) para utilizar los recursos de Amazon Kendra. La IAM es una Servicio de AWS herramienta que puede utilizar sin coste adicional.

Temas

- Público
- Autenticación con identidades
- Administración de acceso mediante políticas
- Cómo funciona Amazon Kendra con IAM
- Ejemplos de políticas basadas en identidades de Amazon Kendra
- AWS políticas gestionadas para Amazon Kendra
- Solución de problemas de identidad y acceso de Amazon Kendra

Público

La forma de usar AWS Identity and Access Management (IAM) varía según el trabajo que realice en Amazon Kendra.

Usuario de servicio: si utiliza el servicio Amazon Kendra para realizar el trabajo, el administrador proporciona las credenciales y los permisos que necesita. A medida que utilice más características de Amazon Kendra para realizar el trabajo, es posible que necesite permisos adicionales. Entender cómo se administra el acceso puede ayudarlo a solicitar los permisos correctos al administrador. Si no puede acceder a una característica de Amazon Kendra, consulte Solución de problemas de identidad y acceso de Amazon Kendra.

Administrador de servicio: si está a cargo de los recursos de Amazon Kendra de su empresa, probablemente tenga acceso completo a Amazon Kendra. Su trabajo consiste en determinar a qué características y recursos de Amazon Kendra deben acceder sus usuarios del servicio. Luego, debe enviar solicitudes a su administrador de IAM para cambiar los permisos de los usuarios de su servicio. Revise la información de esta página para conocer los conceptos básicos de IAM. Para obtener más información acerca de cómo la empresa puede utilizar IAM con Amazon Kendra, consulte Cómo funciona Amazon Kendra con IAM.

Administrador de IAM: si es un administrador de IAM, es posible que desee obtener información sobre cómo escribir políticas para administrar el acceso a Amazon Kendra. Para consultar ejemplos de políticas de Amazon Kendra basadas en identidades que puede utilizar en IAM, consulte Ejemplos de políticas basadas en identidades de Amazon Kendra.

Autenticación con identidades

La autenticación es la forma de iniciar sesión para AWS usar sus credenciales de identidad. Debe estar autenticado (con quien haya iniciado sesión AWS) como usuario de IAM o asumiendo una función de IAM. Usuario raíz de la cuenta de AWS

Puede iniciar sesión AWS como una identidad federada mediante las credenciales proporcionadas a través de una fuente de identidad. AWS IAM Identity Center Los usuarios (Centro de identidades de IAM), la autenticación de inicio de sesión único de su empresa y sus credenciales de Google o Facebook son ejemplos de identidades federadas. Al iniciar sesión como una identidad federada, su administrador habrá configurado previamente la federación de identidades mediante roles de IAM. Cuando accedes AWS mediante la federación, estás asumiendo un rol de forma indirecta.

Según el tipo de usuario que sea, puede iniciar sesión en el portal AWS Management Console o en el de AWS acceso. Para obtener más información sobre cómo iniciar sesión AWS, consulte Cómo iniciar sesión Cuenta de AWS en su Guía del AWS Sign-In usuario.

Si accede AWS mediante programación, AWS proporciona un kit de desarrollo de software (SDK) y una interfaz de línea de comandos (CLI) para firmar criptográficamente sus solicitudes con sus credenciales. Si no utilizas AWS herramientas, debes firmar las solicitudes tú mismo. Para obtener más información sobre cómo usar el método recomendado para firmar las solicitudes usted mismo, consulte Firmar las solicitudes de la AWS API en la Guía del usuario de IAM.

Independientemente del método de autenticación que use, es posible que deba proporcionar información de seguridad adicional. Por ejemplo, le AWS recomienda que utilice la autenticación multifactor (MFA) para aumentar la seguridad de su cuenta. Para obtener más información, consulte Autenticación multifactor en la Guía del usuario de AWS IAM Identity Center y Uso de la autenticación multifactor (MFA) en AWSen la Guía del usuario de IAM.

Cuenta de AWS usuario root

Al crear una Cuenta de AWS, comienza con una identidad de inicio de sesión que tiene acceso completo a todos Servicios de AWS los recursos de la cuenta. Esta identidad se denomina usuario Cuenta de AWS raíz y se accede a ella iniciando sesión con la dirección de correo electrónico y la contraseña que utilizaste para crear la cuenta. Recomendamos encarecidamente que no utilice el

usuario raíz para sus tareas diarias. Proteja las credenciales del usuario raíz y utilícelas solo para las tareas que solo el usuario raíz pueda realizar. Para ver la lista completa de las tareas que requieren que inicie sesión como usuario raíz, consulte Tareas que requieren credenciales de usuario raíz en la Guía del usuario de IAM.

Usuarios y grupos de IAM

Un usuario de IAM es una identidad propia Cuenta de AWS que tiene permisos específicos para una sola persona o aplicación. Siempre que sea posible, recomendamos emplear credenciales temporales, en lugar de crear usuarios de IAM que tengan credenciales de larga duración como contraseñas y claves de acceso. No obstante, si tiene casos de uso específicos que requieran credenciales de larga duración con usuarios de IAM, recomendamos rotar las claves de acceso. Para más información, consulte Rotar las claves de acceso periódicamente para casos de uso que requieran credenciales de larga duración en la Guía del usuario de IAM.

Un grupo de IAM es una identidad que especifica un conjunto de usuarios de IAM. No puede iniciar sesión como grupo. Puede usar los grupos para especificar permisos para varios usuarios a la vez. Los grupos facilitan la administración de los permisos de grandes conjuntos de usuarios. Por ejemplo, podría tener un grupo cuyo nombre fuese IAMAdmins y conceder permisos a dicho grupo para administrar los recursos de IAM.

Los usuarios son diferentes de los roles. Un usuario se asocia exclusivamente a una persona o aplicación, pero la intención es que cualquier usuario pueda asumir un rol que necesite. Los usuarios tienen credenciales permanentes a largo plazo y los roles proporcionan credenciales temporales. Para más información, consulte Cuándo crear un usuario de IAM (en lugar de un rol) en la Guía del usuario de IAM.

Roles de IAM

Un rol de IAM es una identidad dentro de usted Cuenta de AWS que tiene permisos específicos. Es similar a un usuario de IAM, pero no está asociado a una determinada persona. Puede asumir temporalmente una función de IAM en el AWS Management Console cambiando de función. Puede asumir un rol llamando a una operación de AWS API AWS CLI o utilizando una URL personalizada. Para más información sobre los métodos para el uso de roles, consulte Uso de roles de IAM en la Guía del usuario de IAM.

Los roles de IAM con credenciales temporales son útiles en las siguientes situaciones:

- Acceso de usuario federado: para asignar permisos a una identidad federada, puede crear un rol y definir sus permisos. Cuando se autentica una identidad federada, se asocia la identidad

al rol y se le conceden los permisos define el rol. Para obtener información acerca de roles para federación, consulte Creación de un rol para un proveedor de identidades de terceros en la Guía del usuario de IAM. Si utiliza IAM Identity Center, debe configurar un conjunto de permisos. IAM Identity Center correlaciona el conjunto de permisos con un rol en IAM para controlar a qué pueden acceder las identidades después de autenticarse. Para obtener información acerca de los conjuntos de permisos, consulte Conjuntos de permisos en la Guía del usuario de AWS IAM Identity Center .

- Permisos de usuario de IAM temporales: un usuario de IAM puede asumir un rol de IAM para recibir temporalmente permisos distintos que le permitan realizar una tarea concreta.
- Acceso entre cuentas: puede utilizar un rol de IAM para permitir que alguien (una entidad principal de confianza) de otra cuenta acceda a los recursos de la cuenta. Los roles son la forma principal de conceder acceso entre cuentas. Sin embargo, con algunas Servicios de AWS, puedes adjuntar una política directamente a un recurso (en lugar de usar un rol como proxy). Para conocer la diferencia entre las funciones y las políticas basadas en recursos para el acceso entre cuentas, consulte el tema sobre el acceso a recursos entre cuentas en IAM en la Guía del usuario de IAM.
- Acceso entre servicios: algunos utilizan funciones en otros. Servicios de AWS Servicios de AWS Por ejemplo, cuando realiza una llamada en un servicio, es común que ese servicio ejecute aplicaciones en Amazon EC2 o almacene objetos en Amazon S3. Es posible que un servicio haga esto usando los permisos de la entidad principal, usando un rol de servicio o usando un rol vinculado al servicio.
  - Sesiones de acceso directo (FAS): cuando utilizas un usuario o un rol de IAM para realizar acciones en ellas AWS, se te considera director. Cuando utiliza algunos servicios, es posible que realice una acción que desencadene otra acción en un servicio diferente. El FAS utiliza los permisos del principal que llama Servicio de AWS y los solicita Servicio de AWS para realizar solicitudes a los servicios descendentes. Las solicitudes de FAS solo se realizan cuando un servicio recibe una solicitud que requiere interacciones con otros Servicios de AWS recursos para completarse. En este caso, debe tener permisos para realizar ambas acciones. Para obtener información sobre las políticas a la hora de realizar solicitudes de FAS, consulte Reenviar sesiones de acceso.
  - Rol de servicio: un rol de servicio es un rol de IAM que adopta un servicio para realizar acciones en su nombre. Un administrador de IAM puede crear, modificar y eliminar un rol de servicio desde IAM. Para obtener más información, consulte Creación de un rol para delegar permisos a un Servicio de AWS en la Guía del usuario de IAM.
  - Función vinculada al servicio: una función vinculada a un servicio es un tipo de función de servicio que está vinculada a un. Servicio de AWS El servicio puede asumir el rol para realizar

una acción en su nombre. Los roles vinculados al servicio aparecen en usted Cuenta de AWS y son propiedad del servicio. Un administrador de IAM puede ver, pero no editar, los permisos de los roles vinculados a servicios.

- Aplicaciones que se ejecutan en Amazon EC2: puede usar un rol de IAM para administrar las credenciales temporales de las aplicaciones que se ejecutan en una instancia EC2 y realizan AWS CLI solicitudes a la API. AWS Es preferible hacerlo de este modo a almacenar claves de acceso en la instancia de EC2. Para asignar una AWS función a una instancia EC2 y ponerla a disposición de todas sus aplicaciones, debe crear un perfil de instancia adjunto a la instancia. Un perfil de instancia contiene el rol y permite a los programas que se ejecutan en la instancia de EC2 obtener credenciales temporales. Para más información, consulte Uso de un rol de IAM para conceder permisos a aplicaciones que se ejecutan en instancias Amazon EC2 en la Guía del usuario de IAM.

Para obtener información sobre el uso de los roles de IAM, consulte Cuándo crear un rol de IAM (en lugar de un usuario) en la Guía del usuario de IAM.

Administración de acceso mediante políticas

El acceso se controla AWS creando políticas y adjuntándolas a AWS identidades o recursos. Una política es un objeto AWS que, cuando se asocia a una identidad o un recurso, define sus permisos. AWS evalúa estas políticas cuando un director (usuario, usuario raíz o sesión de rol) realiza una solicitud. Los permisos en las políticas determinan si la solicitud se permite o se deniega. La mayoría de las políticas se almacenan AWS como documentos JSON. Para obtener más información sobre la estructura y el contenido de los documentos de política JSON, consulte Información general de políticas JSON en la Guía del usuario de IAM.

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

De forma predeterminada, los usuarios y los roles no tienen permisos. Un administrador de IAM puede crear políticas de IAM para conceder permisos a los usuarios para realizar acciones en los recursos que necesitan. A continuación, el administrador puede añadir las políticas de IAM a roles y los usuarios pueden asumirlos.

Las políticas de IAM definen permisos para una acción independientemente del método que se utilice para realizar la operación. Por ejemplo, suponga que dispone de una política que permite la acción `iam:GetRole`. Un usuario con esa política puede obtener información sobre el rol de la API AWS Management Console AWS CLI, la o la AWS API.

Políticas basadas en identidades

Las políticas basadas en identidad son documentos de políticas de permisos JSON que puede asociar a una identidad, como un usuario de IAM, un grupo de usuarios o un rol. Estas políticas controlan qué acciones pueden realizar los usuarios y los roles, en qué recursos y en qué condiciones. Para obtener más información sobre cómo crear una política basada en identidad, consulte Creación de políticas de IAM en la Guía del usuario de IAM.

Las políticas basadas en identidades pueden clasificarse además como políticas insertadas o políticas administradas. Las políticas insertadas se integran directamente en un único usuario, grupo o rol. Las políticas administradas son políticas independientes que puede adjuntar a varios usuarios, grupos y roles de su Cuenta de AWS empresa. Las políticas administradas incluyen políticas AWS administradas y políticas administradas por el cliente. Para más información sobre cómo elegir una política administrada o una política insertada, consulte Elegir entre políticas administradas y políticas insertadas en la Guía del usuario de IAM.

Políticas basadas en recursos

Las políticas basadas en recursos son documentos de política JSON que se asocian a un recurso. Ejemplos de políticas basadas en recursos son las políticas de confianza de roles de IAM y las políticas de bucket de Amazon S3. En los servicios que admiten políticas basadas en recursos, los administradores de servicios pueden utilizarlos para controlar el acceso a un recurso específico. Para el recurso al que se asocia la política, la política define qué acciones puede realizar una entidad principal especificada en ese recurso y en qué condiciones. Debe especificar una entidad principal en una política en función de recursos. Los principales pueden incluir cuentas, usuarios, roles, usuarios federados o. Servicios de AWS

Las políticas basadas en recursos son políticas insertadas que se encuentran en ese servicio. No puedes usar políticas AWS gestionadas de IAM en una política basada en recursos.

Listas de control de acceso (ACL)

Las listas de control de acceso (ACL) controlan qué entidades principales (miembros de cuentas, usuarios o roles) tienen permisos para acceder a un recurso. Las ACL son similares a las políticas basadas en recursos, aunque no utilizan el formato de documento de políticas JSON.

Amazon S3 y Amazon VPC son ejemplos de servicios que admiten las ACL. AWS WAF Para obtener más información sobre las ACL, consulte Información general de Lista de control de acceso (ACL) en la Guía para desarrolladores de Amazon Simple Storage Service.

Otros tipos de políticas

AWS admite tipos de políticas adicionales y menos comunes. Estos tipos de políticas pueden establecer el máximo de permisos que los tipos de políticas más frecuentes le conceden.

- Límites de permisos: un límite de permisos es una característica avanzada que le permite establecer los permisos máximos que una política basada en identidad puede conceder a una entidad de IAM (usuario o rol de IAM). Puede establecer un límite de permisos para una entidad. Los permisos resultantes son la intersección de las políticas basadas en la identidad de la entidad y los límites de permisos. Las políticas basadas en recursos que especifiquen el usuario o rol en el campo `Principal` no estarán restringidas por el límite de permisos. Una denegación explícita en cualquiera de estas políticas anulará el permiso. Para obtener más información sobre los límites de los permisos, consulte Límites de permisos para las entidades de IAM en la Guía del usuario de IAM.
- Políticas de control de servicios (SCP): las SCP son políticas de JSON que especifican los permisos máximos para una organización o unidad organizativa (OU). AWS Organizations AWS Organizations es un servicio para agrupar y gestionar de forma centralizada varios de los Cuentas de AWS que son propiedad de su empresa. Si habilita todas las características en una organización, entonces podrá aplicar políticas de control de servicio (SCP) a una o a todas sus cuentas. El SCP limita los permisos de las entidades en las cuentas de los miembros, incluidas las de cada una. Usuario raíz de la cuenta de AWS Para obtener más información acerca de Organizations y las SCP, consulte Funcionamiento de las SCP en la Guía del usuario de AWS Organizations .
- Políticas de sesión: las políticas de sesión son políticas avanzadas que se pasan como parámetro cuando se crea una sesión temporal mediante programación para un rol o un usuario federado. Los permisos de la sesión resultantes son la intersección de las políticas basadas en identidades del rol y las políticas de la sesión. Los permisos también pueden proceder de una política en función de recursos. Una denegación explícita en cualquiera de estas políticas anulará el permiso. Para más información, consulte Políticas de sesión en la Guía del usuario de IAM.

Varios tipos de políticas

Cuando se aplican varios tipos de políticas a una solicitud, los permisos resultantes son más complicados de entender. Para saber cómo AWS determinar si se debe permitir una solicitud cuando se trata de varios tipos de políticas, consulte la lógica de evaluación de políticas en la Guía del usuario de IAM.

Cómo funciona Amazon Kendra con IAM

Antes de utilizar IAM para administrar el acceso a Amazon Kendra, debe conocer qué características de IAM están disponibles con Amazon Kendra. Para obtener una visión general de cómo Amazon Kendra y otros AWS servicios funcionan con IAM, consulte AWS Servicios que funcionan con IAM en la Guía del usuario de IAM.

Temas

- Políticas de Amazon Kendra basadas en identidades
- Políticas de Amazon Kendra basadas en recursos
- Listas de control de acceso (ACL)
- Autorización basada en etiquetas de Amazon Kendra
- Roles de IAM en Amazon Kendra

Políticas de Amazon Kendra basadas en identidades

Con las políticas basadas en identidades de IAM, puede especificar las acciones y los recursos permitidos o denegados, así como las condiciones en las que se permiten o deniegan las acciones. Amazon Kendra admite acciones, claves de condiciones y recursos específicos. Para obtener información sobre todos los elementos que utiliza en una política JSON, consulte Referencia de los elementos de las políticas JSON de IAM en la Guía del usuario de IAM.

Acciones

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Action` de una política JSON describe las acciones que puede utilizar para conceder o denegar el acceso en una política. Las acciones políticas suelen tener el mismo nombre que la operación de AWS API asociada. Hay algunas excepciones, como acciones de solo permiso que no tienen una operación de API coincidente. También hay algunas operaciones que requieren varias acciones en una política. Estas acciones adicionales se denominan acciones dependientes.

Incluya acciones en una política para conceder permisos y así llevar a cabo la operación asociada.

Las acciones de políticas de Amazon Kendra utilizan el siguiente prefijo antes de la acción: `kendra:`. Por ejemplo, para conceder permiso a alguien para que publique índices de Amazon Kendra con la operación de ListIndicesAPI, debe incluir la `kendra:ListIndices` acción en

su política. Las instrucciones de la política deben incluir un elemento `Action` o un elemento `NotAction`. Amazon Kendra define su propio conjunto de acciones que describen las tareas que se pueden realizar con este servicio.

Para especificar varias acciones en una única instrucción, sepárelas con comas del siguiente modo:

```
"Action": [
      "kendra:action1",
      "kendra:action2"
```

Puede utilizar caracteres comodín para especificar varias acciones (*). Por ejemplo, para especificar todas las acciones que comiencen con la palabra `Describe`, incluya la siguiente acción:

```
"Action": "kendra:Describe*"
```

Para ver una lista de las acciones de Amazon Kendra, consulte Acciones definidas por Amazon Kendra en la Guía del usuario de IAM.

Recursos

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Resource` de la política JSON especifica el objeto u objetos a los que se aplica la acción. Las instrucciones deben contener un elemento `Resource` o `NotResource`. Como práctica recomendada, especifique un recurso utilizando el Nombre de recurso de Amazon (ARN). Puede hacerlo para acciones que admitan un tipo de recurso específico, conocido como permisos de nivel de recurso.

Para las acciones que no admiten permisos de nivel de recurso, como las operaciones de descripción, utilice un carácter comodín (*) para indicar que la instrucción se aplica a todos los recursos.

```
"Resource": "*"
```

El recurso de índice de Amazon Kendra tiene el siguiente ARN:

```
arn:${Partition}:kendra:${Region}:${Account}:index/${IndexId}
```

Para obtener más información sobre el formato de los ARN, consulte Nombres de recursos de Amazon (ARN) y espacios de nombres de AWS servicio.

Por ejemplo, para especificar un índice en su instrucción, utilice el GUID del índice en el ARN siguiente:

```
"Resource": "arn:aws:kendra:${Region}:${Account}:index/${GUID}"
```

Para especificar todos los índices que pertenecen a una cuenta específica, utilice el carácter comodín (*):

```
"Resource": "arn:aws:${Region}:${Account}:index/*"
```

Algunas acciones de Amazon Kendra, como las que se utilizan para crear recursos, no se pueden llevar a cabo en un recurso específico. En dichos casos, debe utilizar el carácter comodín (*).

```
"Resource": "*"
```

Para ver una lista de los tipos de recursos de Amazon Kendra y los ARN, consulte Recursos definidos por Amazon Kendra en la Guía del usuario de IAM. Para obtener información acerca de las acciones con las que puede especificar el ARN de cada recurso, consulte Acciones definidas por Amazon Kendra.

Claves de condición

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Condition` (o bloque de `Condition`) permite especificar condiciones en las que entra en vigor una instrucción. El elemento `Condition` es opcional. Puede crear expresiones condicionales que utilicen operadores de condición, tales como igual o menor que, para que la condición de la política coincida con los valores de la solicitud.

Si especifica varios elementos de `Condition` en una instrucción o varias claves en un único elemento de `Condition`, AWS las evalúa mediante una operación `AND` lógica. Si especifica varios valores para una única clave de condición, AWS evalúa la condición mediante una `OR` operación lógica. Se deben cumplir todas las condiciones antes de que se concedan los permisos de la instrucción.

También puede utilizar variables de marcador de posición al especificar condiciones. Por ejemplo, puede conceder un permiso de usuario de IAM para acceder a un recurso solo si está etiquetado con su nombre de usuario de IAM. Para más información, consulte Elementos de la política de IAM: variables y etiquetas en la Guía del usuario de IAM.

AWS admite claves de condición globales y claves de condición específicas del servicio. Para ver todas las claves de condición AWS globales, consulte las claves de contexto de condición AWS globales en la Guía del usuario de IAM.

Amazon Kendra no proporciona ninguna clave de condición específica del servicio, pero sí admite el uso de algunas claves de condición globales. Para ver todas las claves de condición AWS globales, consulte las claves de contexto de condición AWS globales en la Guía del usuario de IAM.

Ejemplos

Para ver ejemplos de políticas basadas en identidad de Amazon Kendra, consulte Ejemplos de políticas basadas en identidades de Amazon Kendra.

Políticas de Amazon Kendra basadas en recursos

Amazon Kendra no admite las políticas basadas en recursos.

Listas de control de acceso (ACL)

Amazon Kendra no admite listas de control de acceso (ACL) para acceder a servicios y recursos de AWS .

Autorización basada en etiquetas de Amazon Kendra

Puede asociar etiquetas a determinados tipos de recursos de Amazon Kendra para autorizar el acceso a dichos recursos. Para controlar el acceso utilizando etiquetas, debe proporcionar información de las etiquetas en el elemento de condición de una política utilizando las claves de condición `aws:RequestTag/`*`key-name`* o `aws:TagKeys`.

En la tabla siguiente se enumeran las acciones, los tipos de recursos correspondientes y las claves de condición para el control de acceso basado en etiquetas. Cada acción se autoriza en función de las etiquetas asociadas al tipo de recurso correspondiente.

Acción	Tipo de recurso	Claves de condición
CreateDataFuente		`aws:RequestTag` , `aws:TagKeys`
CreateFaq		`aws:RequestTag` , `aws:TagKeys`
CreateIndex		`aws:RequestTag` , `aws:TagKeys`
API_ ListTags ForResource	origen de datos, preguntas frecuentes, índice	
TagResource	origen de datos, preguntas frecuentes, índice	`aws:RequestTag` , `aws:TagKeys`
UntagResource	origen de datos, preguntas frecuentes, índice	`aws:TagKeys`

Para obtener información acerca del etiquetado de recursos de Amazon Kendra, consulte Etiquetas. Para obtener un ejemplo de política basada en identidad que limita el acceso a un recurso basado en etiquetas de recurso, consulte Ejemplo de política basada en etiquetas. Para obtener más información sobre el uso de etiquetas para limitar el acceso a los recursos, consulte Control del acceso mediante etiquetas en la Guía del usuario de IAM.

Roles de IAM en Amazon Kendra

Un rol de IAM es una entidad de tu AWS cuenta que tiene permisos específicos.

Uso de credenciales temporales con Amazon Kendra

Puede utilizar credenciales temporales para iniciar sesión con federación, asumir un rol de IAM o asumir un rol de acceso entre cuentas. Las credenciales de seguridad temporales se obtienen llamando a operaciones de la AWS STS API, como AssumeRoleo GetFederationToken.

Amazon Kendra admite el uso de credenciales temporales.

Roles de servicio

Esta característica permite que un servicio asuma un rol de servicio en su nombre. Este rol permite que el servicio obtenga acceso a los recursos de otros servicios para completar una acción en su nombre. Los roles de servicio aparecen en su cuenta de IAM y son propiedad de la cuenta. Esto significa que un administrador de IAM puede cambiar los permisos de este rol. Sin embargo, hacerlo podría deteriorar la funcionalidad del servicio.

Amazon Kendra admite roles de servicio.

Elegir un rol de IAM en Amazon Kendra

Al crear un índice, llamar a la operación `BatchPutDocument`, crear un origen de datos o crear una sección de preguntas frecuentes, debe proporcionar un rol de acceso Nombre de recurso de Amazon (ARN) que Amazon Kendra utilice para acceder a los recursos necesarios en su nombre. Si ya ha creado un rol, la consola de Amazon Kendra proporciona una lista de roles para elegir. Es importante elegir un rol que permita el acceso a los recursos que necesita. Para obtener más información, consulte IAM roles de acceso para Amazon Kendra.

Ejemplos de políticas basadas en identidades de Amazon Kendra

De forma predeterminada, los usuarios y roles no tienen permiso para crear ni modificar los recursos de Amazon Kendra. Tampoco pueden realizar tareas con la AWS API AWS Management Console AWS CLI, o. Un administrador de IAM debe crear políticas de IAM que concedan permisos a los usuarios y a los roles para realizar operaciones de la API concretas en los recursos especificados que necesiten. El administrador debe asociar esas políticas a los usuarios o grupos que necesiten esos permisos.

Para obtener información acerca de cómo crear una política basada en identidad de IAM con estos documentos de políticas JSON de ejemplo, consulte Creación de políticas en la pestaña JSON en la Guía del usuario de IAM.

Temas

- Prácticas recomendadas relativas a políticas
- Políticas administradas (predefinidas) por AWS para Amazon Kendra
- Cómo permitir a los usuarios consultar sus propios permisos
- Acceso a un índice de Amazon Kendra
- Ejemplo de política basada en etiquetas

Prácticas recomendadas relativas a políticas

Las políticas basadas en identidades determinan si alguien puede crear, eliminar o acceder a los recursos de Amazon Kendra de la cuenta. Estas acciones pueden generar costos adicionales para su Cuenta de AWS. Siga estas directrices y recomendaciones al crear o editar políticas basadas en identidades:

- Comience con las políticas AWS administradas y avance hacia los permisos con privilegios mínimos: para empezar a conceder permisos a sus usuarios y cargas de trabajo, utilice las políticas AWS administradas que otorgan permisos en muchos casos de uso comunes. Están disponibles en su. Cuenta de AWS Le recomendamos que reduzca aún más los permisos definiendo políticas administradas por el AWS cliente que sean específicas para sus casos de uso. Con el fin de obtener más información, consulte las políticas administradas por AWS o las políticas administradas por AWS para funciones de trabajo en la Guía de usuario de IAM.
- Aplique permisos de privilegio mínimo: cuando establezca permisos con políticas de IAM, conceda solo los permisos necesarios para realizar una tarea. Para ello, debe definir las acciones que se pueden llevar a cabo en determinados recursos en condiciones específicas, también conocidos como permisos de privilegios mínimos. Con el fin de obtener más información sobre el uso de IAM para aplicar permisos, consulte Políticas y permisos en IAM en la Guía del usuario de IAM.
- Utilice condiciones en las políticas de IAM para restringir aún más el acceso: puede agregar una condición a sus políticas para limitar el acceso a las acciones y los recursos. Por ejemplo, puede escribir una condición de políticas para especificar que todas las solicitudes deben enviarse utilizando SSL. También puedes usar condiciones para conceder el acceso a las acciones del servicio si se utilizan a través de una acción específica Servicio de AWS, por ejemplo AWS CloudFormation. Para obtener más información, consulte Elementos de la política de JSON de IAM: Condición en la Guía del usuario de IAM.
- Utilice el analizador de acceso de IAM para validar las políticas de IAM con el fin de garantizar la seguridad y funcionalidad de los permisos: el analizador de acceso de IAM valida políticas nuevas y existentes para que respeten el lenguaje (JSON) de las políticas de IAM y las prácticas recomendadas de IAM. El analizador de acceso de IAM proporciona más de 100 verificaciones de políticas y recomendaciones procesables para ayudar a crear políticas seguras y funcionales. Para más información, consulte Política de validación de Analizador de acceso de IAM en la Guía de usuario de IAM.
- Requerir autenticación multifactor (MFA): si tiene un escenario que requiere usuarios de IAM o un usuario raíz en Cuenta de AWS su cuenta, active la MFA para mayor seguridad. Para solicitar la MFA cuando se invocan las operaciones de la API, agregue las condiciones de la MFA a sus

políticas. Para más información, consulte Configuración del acceso a una API protegido por MFA en la Guía de usuario de IAM.

Para obtener más información sobre las prácticas recomendadas de IAM, consulte las Prácticas recomendadas de seguridad en IAM en la Guía del usuario de IAM.

Políticas administradas (predefinidas) por AWS para Amazon Kendra

AWS aborda muchos casos de uso comunes al proporcionar políticas de IAM independientes que son creadas y administradas por. AWS Estas políticas se denominan políticas AWS gestionadas. AWS Las políticas administradas le permiten asignar permisos a los usuarios, grupos y roles con más facilidad que si tuviera que escribir las políticas usted mismo. Para obtener más información, consulte Adición de permisos a un usuario en la Guía del usuario de IAM.

Las siguientes políticas AWS gestionadas, que puede adjuntar a los grupos y roles de su cuenta, son específicas de Amazon Kendra:

- AmazonKendraReadOnly— Otorga acceso de solo lectura a los recursos de Amazon Kendra.
- AmazonKendraFullAccess— Otorga acceso completo para crear, leer, actualizar, eliminar, etiquetar y ejecutar todos los recursos de Amazon Kendra.

En el caso de la consola, su rol también debe tener permisos `iam:CreateRole`, `iam:CreatePolicy`, `iam:AttachRolePolicy` y `s3:ListBucket`.

> **Note**
>
> Para consultar estos permisos, inicie sesión en la consola de IAM y busque las políticas específicas.

También puede crear sus propias políticas personalizadas con el fin de conceder permisos para realizar acciones de la API de Amazon Kendra. Puede asociar estas políticas personalizadas a los roles o grupos de IAM que requieran esos permisos. Para obtener ejemplos de políticas de IAM para Amazon Kendra, consulte Ejemplos de políticas basadas en identidades de Amazon Kendra.

Cómo permitir a los usuarios consultar sus propios permisos

En este ejemplo, se muestra cómo podría crear una política que permita a los usuarios de IAM ver las políticas administradas e insertadas que se asocian a la identidad de sus usuarios. Esta política

incluye permisos para completar esta acción en la consola o mediante programación mediante la AWS CLI API o. AWS

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "ViewOwnUserInfo",
            "Effect": "Allow",
            "Action": [
                "iam:GetUserPolicy",
                "iam:ListGroupsForUser",
                "iam:ListAttachedUserPolicies",
                "iam:ListUserPolicies",
                "iam:GetUser"
            ],
            "Resource": ["arn:aws:iam::*:user/${aws:username}"]
        },
        {
            "Sid": "NavigateInConsole",
            "Effect": "Allow",
            "Action": [
                "iam:GetGroupPolicy",
                "iam:GetPolicyVersion",
                "iam:GetPolicy",
                "iam:ListAttachedGroupPolicies",
                "iam:ListGroupPolicies",
                "iam:ListPolicyVersions",
                "iam:ListPolicies",
                "iam:ListUsers"
            ],
            "Resource": "*"
        }
    ]
}
```

Acceso a un índice de Amazon Kendra

En este ejemplo, quieres conceder a un usuario de tu AWS cuenta acceso para consultar un índice.

```
{
    "Version": "2012-10-17",
    "Statement": [
```

```
        {
            "Sid": "QueryIndex",
            "Effect": "Allow",
            "Action": [
                "kendra:Query"
            ],
            "Resource": "arn:aws:kendra:${Region}:${Account}:index/${Index ID}"
        }
    ]
}
```

Ejemplo de política basada en etiquetas

Las políticas basadas en etiquetas son documentos de política JSON que especifican las acciones que una entidad principal puede realizar en recursos etiquetados.

Ejemplo: usar una etiqueta para acceder a un recurso

Este ejemplo de política otorga a un usuario o rol de su AWS cuenta permiso para usar la `Query` operación con cualquier recurso etiquetado con la clave **`department`** y el valor**`finance`**.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kendra:Query"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/department": "finance"
                }
            }
        }
   ]
}
```

Ejemplo: usar una etiqueta para activar operaciones de Amazon Kendra

Este ejemplo de política otorga a un usuario o rol de su AWS cuenta permiso para usar cualquier operación de Amazon Kendra, excepto la `TagResource` operación con cualquier recurso etiquetado con la clave **`department`** y el valor. **`finance`**

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra:*",
            "Resource": "*"
        },
        {
            "Effect": "Deny",
            "Action": [
                "kendra:TagResource"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/department": "finance"
                }
            }
        }
    ]
}
```

Ejemplo: usar una etiqueta para restringir el acceso a una operación

Este ejemplo de política restringe el acceso de un usuario o rol de su AWS cuenta para usar la `CreateIndex` operación, a menos que el usuario proporcione la **`department`** etiqueta y tenga los valores **`finance`** permitidos y. **`IT`**

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra:CreateIndex",
            "Resource": "*"
        },
```

```
        {
            "Effect": "Deny",
            "Action": "kendra:CreateIndex",
            "Resource": "*",
            "Condition": {
                "Null": {
                    "aws:RequestTag/department": "true"
                }
            }
        },
        {
            "Effect": "Deny",
            "Action": "kendra:CreateIndex",
            "Resource": "*",
            "Condition": {
                "ForAnyValue:StringNotEquals": {
                    "aws:RequestTag/department": [
                        "finance",
                        "IT"
                    ]
                }
            }
        }
    ]
}
```

AWS políticas gestionadas para Amazon Kendra

Para añadir permisos a usuarios, grupos y roles, es más fácil usar políticas AWS administradas que escribirlas usted mismo. Se necesita tiempo y experiencia para crear políticas administradas por el cliente de IAM que proporcionen a su equipo solo los permisos necesarios. Para empezar rápidamente, puedes usar nuestras políticas AWS gestionadas. Estas políticas cubren casos de uso comunes y están disponibles en tu AWS cuenta. Para obtener más información sobre las políticas AWS administradas, consulte las políticas AWS administradas en la Guía del usuario de IAM.

AWS los servicios mantienen y AWS actualizan las políticas gestionadas. No puede cambiar los permisos en las políticas AWS gestionadas. En ocasiones, los servicios agregan permisos adicionales a una política administrada por AWS para admitir características nuevas. Este tipo de actualización afecta a todas las identidades (usuarios, grupos y roles) donde se asocia la política. Es más probable que los servicios actualicen una política administrada por AWS cuando se lanza

una nueva característica o cuando se ponen a disposición nuevas operaciones. Los servicios no eliminan los permisos de una política AWS administrada, por lo que las actualizaciones de la política no afectarán a los permisos existentes.

Además, AWS admite políticas administradas para funciones laborales que abarcan varios servicios. Por ejemplo, la política de ReadOnlyacceso AWS gestionado proporciona acceso de solo lectura a todos los AWS servicios y recursos. Cuando un servicio lanza una nueva función, AWS agrega permisos de solo lectura para nuevas operaciones y recursos. Para obtener una lista y descripciones de las políticas de funciones de trabajo, consulte Políticas administradas de AWS para funciones de trabajo en la Guía del usuario de IAM.

AWS política gestionada: AmazonKendraReadOnly

Concede acceso de solo lectura a los recursos de Amazon Kendra. Esta política incluye los siguientes permisos.

- `kendra`: permite a los usuarios realizar acciones que devuelven una lista de elementos o detalles sobre un elemento. Esto incluye las operaciones de la API que comienzan con `Describe`, `List`, `Query`, `BatchGetDocumentStatus`, `GetQuerySuggestions` o `GetSnapshots`.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "kendra:Describe*",
                "kendra:List*",
                "kendra:Query",
                "kendra:BatchGetDocumentStatus",
                "kendra:GetQuerySuggestions",
                "kendra:GetSnapshots"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

AWS política gestionada: AmazonKendraFullAccess

Concede acceso para crear, leer, actualizar, eliminar, etiquetar y ejecutar todos los recursos de Amazon Kendra. Esta política incluye los siguientes permisos.

- `kendra`: permite a las entidades principales acceso de lectura y escritura a todas las acciones de Amazon Kendra.
- `s3`: permite a las entidades principales obtener ubicaciones de buckets de Amazon S3 y enumerar buckets.
- `iam`: permite a las entidades principales transmitir y enumerar roles.
- `kms`—Permite a los directores describir y enumerar AWS KMS las claves y los alias.
- `secretsmanager`: permite a las entidades principales crear, describir y enumerar secretos.
- `ec2`: permite a las entidades principales describir grupos de seguridad, VCP (nube privada virtual) y subredes.
- `cloudwatch`: permite a las entidades principales ver las métricas de Cloud Watch.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "iam:PassRole",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "iam:PassedToService": "kendra.amazonaws.com"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "iam:ListRoles"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
```

```
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeVpcs",
                "ec2:DescribeSubnets"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kms:ListKeys",
                "kms:ListAliases",
                "kms:DescribeKey"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:ListAllMyBuckets",
                "s3:GetBucketLocation"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:ListSecrets"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "cloudwatch:GetMetricData"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:CreateSecret",
                "secretsmanager:DescribeSecret"
            ],
            "Resource": "arn:aws:secretsmanager:*:*:secret:AmazonKendra-*"
```

```
        },
        {
            "Effect": "Allow",
            "Action": "kendra:*",
            "Resource": "*"
        }
    ]
}
```

Amazon Kendra actualiza las políticas gestionadas AWS

Consulte los detalles sobre las actualizaciones de las políticas AWS gestionadas de Amazon Kendra desde que este servicio comenzó a realizar el seguimiento de estos cambios. Para obtener alertas automáticas sobre cambios en esta página, suscríbase a la fuente RSS en la página de historial de documentos de Amazon Kendra.

Cambio	Descripción	Fecha
AmazonKendraReadOnly—Añadir permiso para dar soporte a las API GetSnapshots BatchGetDocumentStatus	Amazon Kendra agregó las nuevas API `GetSnapshots` y `BatchGetDocumentStatus` . `GetSnapshots` proporciona datos que muestran cómo interactúan sus usuarios con su aplicación de búsqueda. `BatchGetDocumentStatus` supervisa el progreso de la indexación de los documentos.	3 de enero de 2022
AmazonKendraReadOnly—Añadir permiso para respaldar la operación GetQuerySuggestions	Amazon Kendra agregó una nueva API `GetQuerySuggestions` que permite acceder a sugerencias de consultas de búsqueda populares, lo que ayuda	27 de mayo de 2021

Cambio	Descripción	Fecha
	a guiar la búsqueda de los usuarios. Cuando los usuarios escriben su consulta de búsqueda, la consulta sugerida ayuda a completar automáticamente la búsqueda.	
Amazon Kendra comenzó a realizar el seguimiento de los cambios	Amazon Kendra comenzó a realizar un seguimiento de los cambios en sus políticas AWS gestionadas.	27 de mayo de 2021

Solución de problemas de identidad y acceso de Amazon Kendra

Utilice la siguiente información para diagnosticar y solucionar los problemas habituales que pueden surgir cuando se trabaja con Amazon Kendra e IAM.

Temas

- No tengo autorización para realizar una acción en Amazon Kendra
- No estoy autorizado a realizar lo siguiente: PassRole
- Soy administrador y deseo permitir que otras personas accedan a Amazon Kendra
- Quiero permitir que personas ajenas a mi AWS cuenta accedan a mis recursos de Amazon Kendra

No tengo autorización para realizar una acción en Amazon Kendra

Si AWS Management Console le indica que no está autorizado a realizar una acción, debe ponerse en contacto con su administrador para obtener ayuda. El administrador es la persona que le proporcionó las credenciales de inicio de sesión.

En el siguiente ejemplo, el error se produce cuando el usuario `mateojackson` intenta utilizar la consola para ver detalles sobre un índice, pero no tiene permisos `kendra:DescribeIndex`.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 kendra:DescribeIndex on resource: index ARN
```

En este caso, Mateo pide a su administrador que actualice sus políticas de forma que pueda obtener acceso al recurso `index` mediante la acción `kendra:DescribeIndex`.

No estoy autorizado a realizar lo siguiente: PassRole

Si recibe un error que indica que no tiene autorización para llevar a cabo la acción `iam:PassRole`, las políticas se deben actualizar para permitirle pasar un rol a Amazon Kendra.

Algunas Servicios de AWS permiten transferir una función existente a ese servicio en lugar de crear una nueva función de servicio o una función vinculada a un servicio. Para ello, debe tener permisos para transferir el rol al servicio.

En el siguiente ejemplo, el error se produce cuando un usuario de IAM denominado `marymajor` intenta utilizar la consola para realizar una acción en Amazon Kendra. Sin embargo, la acción requiere que el servicio cuente con permisos que otorguen un rol de servicio. Mary no tiene permisos para transferir el rol al servicio.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

En este caso, las políticas de Mary se deben actualizar para permitirle realizar la acción `iam:PassRole`.

Si necesita ayuda, póngase en contacto con su administrador. AWS El administrador es la persona que le proporcionó las credenciales de inicio de sesión.

Soy administrador y deseo permitir que otras personas accedan a Amazon Kendra

Para permitir que otros accedan a Amazon Kendra, debe crear una entidad de IAM (usuario o rol) para la persona o aplicación que necesita acceso. Esta persona utilizará las credenciales de la entidad para acceder a AWS. A continuación, debe asociar una política a la entidad que les conceda los permisos correctos en Amazon Kendra.

Para comenzar de inmediato, consulte Creación del primer grupo y usuario delegado de IAM en la Guía del usuario de IAM.

Quiero permitir que personas ajenas a mi AWS cuenta accedan a mis recursos de Amazon Kendra

Puede crear un rol que los usuarios de otras cuentas o las personas externas a la organización puedan utilizar para acceder a sus recursos. Puede especificar una persona de confianza para que

asuma el rol. En el caso de los servicios que admitan las políticas basadas en recursos o las listas de control de acceso (ACL), puede utilizar dichas políticas para conceder a las personas acceso a sus recursos.

Para más información, consulte lo siguiente:

- Para saber si Amazon Kendra admite estas características, consulte Cómo funciona Amazon Kendra con IAM.
- Para obtener información sobre cómo proporcionar acceso a sus recursos a través de Cuentas de AWS los suyos, consulte Proporcionar acceso a un usuario de IAM en otro de su propiedad Cuenta de AWS en la Guía del usuario de IAM.
- Para obtener información sobre cómo proporcionar acceso a tus recursos a terceros Cuentas de AWS, consulta Cómo proporcionar acceso a recursos que Cuentas de AWS son propiedad de terceros en la Guía del usuario de IAM.
- Para obtener información sobre cómo proporcionar acceso mediante una federación de identidades, consulte Proporcionar acceso a usuarios autenticados externamente (identidad federada) en la Guía del usuario de IAM.
- Para conocer la diferencia entre usar roles y políticas basadas en recursos para el acceso entre cuentas, consulte el tema Acceso a recursos entre cuentas en IAM en la Guía del usuario de IAM.

Prácticas recomendadas de seguridad

Amazon Kendra proporciona una serie de características de seguridad que debe tener en cuenta a la hora de desarrollar e implementar sus propias políticas de seguridad. Las siguientes prácticas recomendadas son directrices generales y no suponen una solución de seguridad completa. Puesto que es posible que estas prácticas recomendadas no sean adecuadas o suficientes para el entorno, considérelas como consideraciones útiles en lugar de como normas.

Aplicación del principio de privilegios mínimos

Amazon Kendra proporciona una política de acceso pormenorizada para las aplicaciones que utilizan IAM roles. Recomendamos que a los roles solo se les otorguen los privilegios mínimos necesarios para el trabajo, como cubrir su aplicación y el acceso al destino del registro. También recomendamos auditar los trabajos para detectar permisos de forma regular y ante cualquier cambio en su aplicación.

Permisos de control de acceso basado en roles (RBAC)

Los administradores deben controlar estrictamente los permisos de control de acceso basado en roles (RBAC) para aplicaciones de Amazon Kendra.

Registro y monitoreo en Amazon Kendra

El monitoreo es una parte importante del mantenimiento de la fiabilidad, la disponibilidad y el rendimiento de las aplicaciones de Amazon Kendra. Para supervisar las llamadas a la API de Amazon Kendra, puede utilizar. AWS CloudTrail Para supervisar el estado de sus trabajos, utilice Amazon CloudWatch Logs.

- Amazon CloudWatch Alarms: al usar CloudWatch las alarmas, usted observa una única métrica durante un período de tiempo que especifique. Si la métrica supera una política. CloudWatch las alarmas no invocan acciones cuando una métrica se encuentra en un estado determinado. En su lugar, el estado debe haber cambiado y debe mantenerse durante el número de periodos especificado. Para obtener más información, consulte Monitorización de Amazon Kendra con Amazon CloudWatch.
- AWS CloudTrail Registros: CloudTrail proporciona un registro de las acciones realizadas por un usuario, un rol o un AWS servicio en Amazon Kendra o Amazon Kendra Intelligent Ranking. Con la información recopilada por CloudTrail, puede determinar la solicitud que se realizó a Amazon Kendra, la dirección IP desde la que se realizó la solicitud, quién la hizo, cuándo se realizó y detalles adicionales. Para obtener más información, consulte Registro de llamadas a la API de Amazon Kendra con registros AWS CloudTrail y Registro de las llamadas a la API de Amazon Kendra Intelligent Ranking con registros AWS CloudTrail..

Validación de la conformidad de Amazon Kendra

Los auditores externos evalúan la seguridad y la conformidad de Amazon Kendra en distintos programas de conformidad de Amazon Kendra. Amazon Kendra cumple con lo siguiente:

- Ley de Portabilidad y Responsabilidad de Seguros Médicos de EE. UU (Health Insurance Portability and Accountability Act, HIPAA).
- Controles del Sistema y Organizaciones (System and Organization Controls, SOC) 2
- Programa de Asesores Registrados de Seguridad de la Información (Information Security Registered Assessors Program, IRAP)

- Programa Federal de Administración de Riesgos y Autorizaciones (Federal Risk and Authorization Management Program, FedRAMP) Moderado en las regiones EE. UU. Este y EE. UU. Oeste
- El Programa Federal de Gestión de Riesgos y Autorizaciones (FedRAMP) tiene un alto nivel en la región GovCloud AWS (EE. UU. Oeste)

Para ver una lista de AWS los servicios incluidos en el ámbito de los programas de conformidad específicos, consulte los AWS servicios incluidos en el ámbito de aplicación por programa de conformidad y los programa de conformidad. Para obtener información general, consulte Programas de AWS cumplimiento > Programas AWS .

Puede descargar informes de auditoría de terceros utilizando AWS Artifact. Para obtener más información, consulte Descarga de informes en AWS Artifact .

Su responsabilidad de conformidad al utilizar Amazon Kendra se determina en función de la sensibilidad de los datos, los objetivos de cumplimiento de su empresa y la legislación y los reglamentos correspondientes. AWS proporciona los siguientes recursos para ayudar con la conformidad:

- Guías de inicio rápido de seguridad y cumplimiento de implementación analizan las consideraciones arquitectónicas y proporcionan los pasos para implementar entornos básicos centrados en la seguridad y el cumplimiento. AWS
- Documento técnico sobre cómo diseñar una arquitectura basada en la seguridad y el cumplimiento de la HIPAA: este documento técnico describe cómo pueden utilizar las empresas para crear aplicaciones que cumplan con la HIPAA. AWS
- AWS Recursos de cumplimiento Recursos de cumplimiento: esta colección y guías puede aplicarse a su sector y ubicación.
- Evaluación de los recursos con reglas en la guía para AWS Config desarrolladores: el AWS Config servicio evalúa en qué medida las configuraciones de los recursos cumplen con las prácticas internas, las directrices del sector y las normas.
- AWS Security Hub—Este AWS servicio proporciona una visión integral del estado de su seguridad AWS que le ayuda a comprobar el cumplimiento de los estándares y las mejores prácticas del sector de la seguridad.

Resiliencia en Amazon Kendra

La infraestructura AWS global se basa en AWS regiones y zonas de disponibilidad. AWS Las regiones proporcionan varias zonas de disponibilidad aisladas y separadas físicamente, que están conectadas mediante redes de baja latencia, alto rendimiento y alta redundancia. Con las zonas de disponibilidad, puede diseñar y utilizar aplicaciones y bases de datos que realizan una conmutación por error automática entre las zonas sin interrupciones. Las zonas de disponibilidad tienen una mayor disponibilidad, tolerancia a errores y escalabilidad que las infraestructuras tradicionales de uno o varios centros de datos.

Para obtener más información sobre AWS las regiones y las zonas de disponibilidad, consulte Infraestructura global.AWS

Con una infraestructura AWS global, Amazon Kendra Enterprise Edition es tolerante a errores, escalable y de alta disponibilidad. Actualmente no se admite la reversión a versiones anteriores de un índice, pero puede actualizar o volver a crear partes del índice si elimina y vuelve a agregar los orígenes de datos existentes al índice.

Seguridad de infraestructuras en Amazon Kendra

Como servicio gestionado, Amazon Kendra está protegido por la seguridad de la red AWS global. Para obtener información sobre los servicios AWS de seguridad y cómo se AWS protege la infraestructura, consulte Seguridad AWS en la nube. Para diseñar su AWS entorno utilizando las mejores prácticas de seguridad de la infraestructura, consulte Protección de infraestructuras en un marco de buena AWS arquitectura basado en el pilar de la seguridad.

Utiliza las llamadas a la API AWS publicadas para acceder a Amazon Kendra a través de la red. Los clientes deben admitir lo siguiente:

- Seguridad de la capa de transporte (TLS). Exigimos TLS 1.2 y recomendamos TLS 1.3.
- Conjuntos de cifrado con confidencialidad directa total (PFS) como DHE (Ephemeral Diffie-Hellman) o ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). La mayoría de los sistemas modernos como Java 7 y posteriores son compatibles con estos modos.

Además, las solicitudes deben estar firmadas mediante un ID de clave de acceso y una clave de acceso secreta que esté asociada a una entidad de seguridad de IAM principal. También puede utilizar AWS Security Token Service (AWS STS) para generar credenciales de seguridad temporales para firmar solicitudes.

Análisis de configuración y vulnerabilidad en AWS Identity and Access Management

AWS se encarga de las tareas de seguridad básicas, como la aplicación de parches al sistema operativo (SO) huésped y a las bases de datos, la configuración del firewall y la recuperación ante desastres. Estos procedimientos han sido revisados y certificados por los terceros pertinentes. Para obtener más detalles, consulte los siguientes recursos de :

- Modelo de responsabilidad compartida
- AWS: Información general acerca de los procesos de seguridad (documento técnico)

Los siguientes recursos también abordan la configuración y el análisis de vulnerabilidades en AWS Identity and Access Management (IAM):

- Validación de conformidad para AWS Identity and Access Management
- Mejores prácticas de seguridad y casos de uso en AWS Identity and Access Management.

Cuotas para Amazon Kendra

Regiones de admitidas

Para obtener una lista de AWS las regiones en las Amazon Kendra que está disponible, consulte Amazon Kendra las regiones y puntos de enlace en la Referencia general de Amazon Web Services.

Cuotas

Las cuotas de servicio, también denominadas límites, son la cantidad máxima de recursos de servicio para su AWS cuenta. Para obtener más información, consulte Amazon Kendra service quotas en la Referencia general de AWS .

Cuotas indexadas

Descripción	Predeterminado	Edición	Ajustable
Número máximo de índices por cuenta	10	Desarrollador, empresa	Sí
Cantidad de texto extraída para un índice en una sola unidad (desarrollador). No puede añadir unidades adicionales para extraer texto en la Developer Edition.	3 GB	Desarrollador	No
Cantidad de texto extraída para un índice en una sola unidad (Enterprise). Puede añadir	30 GB	Enterprise	Sí

Descripción	Predeterminado	Edición	Ajustable
hasta 100 unidades adicionales para extraer texto para la edición Enterpris e, o simplemente ponerse en contacto con Soporte.			

Cuotas del conector de fuente de datos

Descripción	Predeterminado	Edición	Ajustable
Número máximo de conectores de fuentes de datos por índice (desarrollador)	5	Desarrollador	No
Número máximo de conectores de fuentes de datos por índice (Enterprise)	50	Enterprise	Sí
Tamaño máximo de un único documento o archivo sin procesar cuando se utiliza un conector de fuente de datos	50 MB	Desarrollador, empresa	Sí
Número máximo de prefijos S3 en el archivo de configura ción de la lista de control de acceso incluido en el conector	100	Desarrollador, empresa	No

Descripción	Predeterminado	Edición	Ajustable
de la fuente Amazon S3 de datos			
Tamaño máximo del archivo de configuración de la lista de control de acceso incluido en el conector de la fuente de Amazon S3 datos	50 MB	Desarrollador, empresa	Sí

Preguntas frecuentes sobre cuotas

Descripción	Predeterminado	Edición	Ajustable
Número máximo de preguntas frecuentes por índice	30	Desarrollador, empresa	Sí
Tamaño máximo de 1 FAQ	5 MB	Desarrollador, empresa	Sí
Número máximo de resultados devueltos para FAQ	4	Desarrollador, empresa	Sí
Número máximo de caracteres permitido para una pregunta de preguntas frecuentes	300	Desarrollador, empresa	No
Número máximo de caracteres en una respuesta a las preguntas frecuentes	2000	Desarrollador, empresa	No

Cuotas del tesauro

Descripción	Predeterminado	Edición	Ajustable
Número máximo de tesauros por índice	1	Desarrollador, empresa	No
Tamaño máximo de un archivo de tesauro	5 MB	Desarrollador, empresa	Sí
Número máximo de reglas de sinónimos por tesauro	10 000	Desarrollador, empresa	Sí
Número máximo de sinónimos por término en todos los tesauros de un índice	10	Desarrollador, empresa	No

Amazon Kendra cuotas de experiencia

Descripción	Predeterminado	Edición	Ajustable
Número máximo de Amazon Kendra experiencias por índice	50	Desarrollador, empresa	Sí

Cuotas de consultas y resultados de búsqueda

Descripción	Predeterminado	Edición	Ajustable
Cantidad de consultas por segundo para un índice en una	0,05	Desarrollador	No

Descripción	Predeterminado	Edición	Ajustable
sola unidad (desarrollador). No puede añadir unidades adicionales para consultas para la Developer Edition.			
Cantidad de consultas por segundo para un índice en una sola unidad (Enterprise). Puede añadir hasta 100 unidades adicionales para consultas sobre la edición Enterprise, o simplemente póngase en contacto con Soporte.	0.1	Enterprise	Sí
Número máximo de caracteres por texto de consulta	1 000	Desarrollador, empresa	Sí
Número máximo de resultados de búsqueda por consulta. El valor predeterminado es 100. Para obtener más de 100 resultados, solo tiene que ponerse en contacto con Support.	100	Desarrollador, empresa	Sí

Descripción	Predeterminado	Edición	Ajustable
Número máximo de resultados de búsqueda por página	100	Desarrollador, empresa	Sí
Número máximo de palabras simbólicas por texto de consulta antes del truncamiento. El valor predeterminado es 30. Para permitir más de 30 palabras, simplemente póngase en contacto con Soporte.	30	Desarrollador, empresa	Sí
Tamaño máximo de la lista de grupos de usuarios por atributo de consulta	1 000	Desarrollador, empresa	Sí
Tamaño máximo de la lista de cadenas por atributo de consulta	10	Desarrollador, empresa	Sí

Cuotas de sugerencias de consultas

Descripción	Predeterminado	Edición	Ajustable
Número máximo de sugerencias de consulta devueltas por llamada de GetQuerysugerencias	10	Desarrollador, empresa	Sí

Descripción	Predeterminado	Edición	Ajustable
Número máximo de campos o atributos para las sugerencias de consulta por GetQuery llamada de sugerencias	10	Desarrollador, empresa	Sí
Número máximo de campos o atributos adicionales para las sugerencias de consulta por GetQuery llamada de sugerencias	5	Desarrollador, empresa	Sí
Número máximo de listas de bloqueo por índice	1	Desarrollador, empresa	No
Tamaño máximo de un archivo de texto de lista de bloqueo	2 MB	Desarrollador, empresa	Sí
Número máximo de elementos (palabras o frases) en una lista de bloqueados	20 000	Desarrollador, empresa	Sí
Número máximo de sugerencias de consultas con corrección ortográfica que se pueden devolver en una llamada a la API `Query`.	1	Desarrollador, empresa	Sí

Cuotas de documentos

Descripción	Predeterminado	Edición	Ajustable
Cantidad de texto extraída para un índice en una sola unidad (desarrollador). No puede añadir unidades adicionales para extraer texto en la Developer Edition.	3 GB	Desarrollador	No
Cantidad de texto extraída para un índice en una sola unidad (Enterprise). Puede añadir hasta 100 unidades adicionales para extraer texto para la edición Enterprise, o simplemente ponerse en contacto con Soporte.	30 GB	Enterprise	Sí
Tamaño máximo de un único documento o archivo sin procesar cuando se utiliza un conector de fuente de datos	50 MB	Desarrollador, empresa	Sí
Tamaño máximo de un único documento o archivo sin procesar	5 MB	Desarrollador, empresa	Sí

Descripción	Predeterminado	Edición	Ajustable
cuando se utiliza la `BatchPutDocument` API			
Cantidad máxima de texto extraída de un único documento	5 MB	Desarrollador, empresa	No
Número máximo de campos/atributos personalizados por índice	500	Desarrollador, empresa	No

Cuotas de resultados de búsqueda destacados

Descripción	Predeterminado	Edición	Ajustable
Número máximo de documentos destacados por conjunto de resultados destacados	4	Enterprise	Sí
Número máximo de textos de consulta por conjunto de resultados destacados	49	Enterprise	No
Número máximo de caracteres por texto de consulta en un conjunto de resultados destacados	1 000	Enterprise	Sí

Descripción	Predeterminado	Edición	Ajustable
Número máximo de conjuntos de resultados destacados por índice	50	Enterprise	Sí

Recupera las cuotas de resultados de búsqueda

Descripción	Predeterminado	Edición	Ajustable
Número máximo de solicitudes `Rescore` por segundo para un plan de ejecución rescore o una sola unidad de capacidad. Puede añadir hasta 1000 unidades adicionales.	0.01	Enterprise	No
Número máximo de planes de ejecución rescore por cuenta.	50	Enterprise	Sí
Número máximo de tokens en `Title` para un documento en una solicitud `Rescore`.	100	Enterprise	No
Número máximo de tokens en `Body` para un documento en una solicitud `Rescore`.	200	Enterprise	No

Descripción	Predeterminado	Edición	Ajustable
Número máximo de documentos en una solicitud `Rescore`.	25	Enterprise	No
Número máximo de documentos por grupo en una solicitud `Rescore`.	3	Enterprise	No

Para obtener más información sobre las cuotas de Amazon Kendra servicio y solicitar un aumento de cuota, consulte Service Quotas.

Solución de problemas

Esta sección puede ayudarle a resolver problemas comunes que puede encontrar al trabajar con ellos Amazon Kendra.

Temas

- Solución de problemas con los orígenes de datos
- Solución de problemas con los resultados de búsqueda de documentos
- Solución de problemas generales

Solución de problemas con los orígenes de datos

Esta sección puede ayudarle a resolver problemas comunes al configurar y utilizar Amazon Kendra los conectores de fuentes de datos.

No se han indexado mis documentos

Al sincronizar el Amazon Kendra índice con una fuente de datos, es posible que se produzcan problemas que impidan la indexación de los documentos. La indexación es un proceso que consta de dos pasos. En primer lugar, se comprueba el origen de datos para ver si hay documentos nuevos y actualizados que se deben indexar y se buscan documentos que se deben eliminar del índice. En segundo lugar, en el nivel del documento, se accede a cada documento y se indexa.

Se puede producir un error en cualquiera de estos pasos. Los errores de origen de datos se indican en la consola, en la sección Historial de ejecuciones de sincronización de la página de detalles del origen de datos. El estado del trabajo de sincronización puede ser correcto, incompleto o erróneo. También puede ver el número de documentos indexados y eliminados durante el trabajo. Si el estado es erróneo, se muestra un mensaje en la columna Detalles.

Los errores a nivel de documento se notifican en. Amazon CloudWatch Logs Puede ver los errores en la CloudWatch consola.

Para generar un informe de estado de sincronización de documentos, consulte Deseo generar un informe de estado de sincronización para mis documentos.

Ha fallado mi trabajo de sincronización

Un trabajo de sincronización suele fallar cuando hay un error de configuración en el índice o en el origen de datos. En la consola, encontrará el mensaje de error en la sección Historial de ejecuciones de sincronización de la página de detalles del origen de datos, en la columna Detalles. Los errores de documento se indican en Amazon CloudWatch Logs. El mensaje de error proporciona información sobre lo que ha fallado. El problema suele ser que el índice o la fuente de datos no tienen los IAM permisos adecuados. El mensaje de error describe los permisos que faltan. A continuación se muestran algunos de los mensajes de error que puede recibir:

`Failed to create log group for job. Please make sure that the IAM role provided has sufficient permissions.`

Si su función de índice no tiene permiso de uso CloudWatch, la fuente de datos no podrá crear un CloudWatch registro. Si recibe este error, debe añadir CloudWatch permisos al rol de índice.

`Failed to access Amazon S3 file prefix (`*`bucket name`*`) while trying to crawl your metadata files. Please make sure the IAM role (`*`ARN`*`) provided has sufficient permissions.`

Si utiliza una fuente de Amazon S3 datos, Amazon Kendra debe tener permiso para acceder al depósito que contiene los documentos. Debe añadir permiso para leer el depósito Amazon Kendra a la IAM función de fuente de datos.

`The provided IAM role (`*`ARN`*`) could not be assumed. Please make sure Amazon Kendra is a trusted entity that is allowed to assume the role.`

Amazon Kendra necesita permiso para asumir las IAM funciones de índice y fuente de datos. Debe añadir una política de confianza a los roles con permiso para la acción `sts:AssumeRole`.

Para conocer las IAM políticas que Amazon Kendra deben indexar una fuente de datos, consulte IAM las funciones.

Para generar un informe de estado de sincronización de documentos, consulte Deseo generar un informe de estado de sincronización para mis documentos.

Mi trabajo de sincronización está incompleto

Por lo general, los trabajos quedan incompletos cuando han finalizado el proceso de origen de datos, pero se produce algún error durante el proceso a nivel de documento. Cuando un trabajo está

incompleto, es posible que algunos de los documentos no se hayan indexado correctamente. En el caso de un origen de datos de Amazon S3 , las causas por las que un trabajo está incompleto son:

- Los metadatos de uno o más documentos no eran válidos.
- Cuando se envían documentos para su indexación pero no se ha enviado al menos un documento.
- Cuando se envían documentos para su eliminación del índice pero no se ha enviado al menos un documento.

Para solucionar problemas relacionados con un trabajo de sincronización incompleto, consulte primero los CloudWatch registros.

1. En la columna de detalles, selecciona Ver detalles en CloudWatch.
2. Revise los mensajes de error para ver qué causó el error en el documento.

Para generar un informe de estado de sincronización de documentos, consulte Deseo generar un informe de estado de sincronización para mis documentos.

Mi trabajo de sincronización se ha realizado correctamente, pero no hay documentos indexados

En algunas ocasiones, la ejecución de un trabajo de sincronización de índices se marca como correcto, pero no hay ningún documento nuevo o actualizado indexado en el momento esperado. Algunas de las causas posibles son:

- Comprueba la CloudWatch `DocumentsSubmittedForIndexingFailed` métrica para ver si algún documento no se ha sincronizado. Comprueba tus CloudWatch registros para obtener más información.
- En el caso de una fuente de Amazon S3 datos, es posible que hayas introducido un nombre Amazon Kendra de depósito o un prefijo incorrectos. Asegúrese de que el depósito que Amazon Kendra está utilizando es el que contiene los documentos que se van a indexar.
- Al volver a indexar un documento que no se pudo indexar en un trabajo anterior, Amazon Kendra no lo indexará a menos que haya cambiado el documento o el archivo de metadatos asociado.

Para generar un informe de estado de sincronización de documentos, consulte Deseo generar un informe de estado de sincronización para mis documentos.

Tengo problemas con el formato de los archivos al sincronizar mi origen de datos

Si tiene problemas con el formato de los archivos al añadir archivos al origen de datos o al sincronizar el origen de datos, asegúrese de que los tipos de documento sean compatibles con Amazon Kendra . Para obtener una lista de los tipos de documentos compatibles, Amazon Kendra consulte Tipos o formatos de documentos.

Si utiliza la API `BatchPutDocument` con archivos de texto sin formato, especifique `PLAIN_TEXT` como el tipo de contenido.

Quiero generar un informe del historial de sincronización de mis documentos

Al sincronizar el conector de la fuente de Amazon Kendra datos, Amazon Kendra puede generar informes de estado de sincronización para cada documento de la fuente de datos y copiarlos en un Amazon S3 depósito. Durante este proceso, los datos se cifran mediante claves de AWS KMS y solo usted puede verlos. El estado del documento del que se ha informado puede ser uno de los siguientes: erróneo, completado o satisfactorio con errores.

Antes de poder generar informes de estado de sincronización, debe hacer lo siguiente:

- Añada el siguiente principio Amazon Kendra de servicio a su política de Amazon S3 acceso

```
{
    “Version”: “2012-10-17",
    “Statement”: [
        {
            “Sid”: “KendraS3Access”,
            “Effect”: “Allow”,
            “Principal”: {
                “Service”: “kendra.amazonaws.com”
            },
            “Action”: “s3:PutObject”,
            “Resource”: “arn:aws:s3:::your-manifest-bucket-name/*”
        }
    ]
}
```

- Cree un Amazon S3 depósito con permisos de acceso a Amazon Kendra

Si utiliza la consola, para generar un informe de estado de sincronización, elija activar la opción de generación del historial de sincronización en la página de detalles del origen de datos. A continuación, introduzca la ubicación del Amazon S3 depósito y elija entre las opciones de configuración disponibles. Los informes se generarán a partir de la siguiente sincronización, una vez que haya activado la opción de generar informes.

Si eliminas el Amazon S3 depósito, perderás tus datos de registro y tendrás que configurar uno nuevo para almacenar los nuevos informes de sincronización.

Actualmente, solo se admite el estado de generación de informes de sincronización para el conector de Amazon S3.

¿Cuánto tiempo lleva sincronizar un origen de datos?

Si no hay actualizaciones en los documentos, el tiempo de sincronización de un Amazon Kendra índice aumenta en proporción lineal al número de documentos. Por ejemplo, 1000 documentos sin ninguna actualización tardarían unos cinco minutos en sincronizarse y 2000 documentos sin ninguna actualización tardarían unos 10 minutos. Si hay actualizaciones en los documentos, el tiempo de sincronización aumentará en función del número de documentos actualizados.

¿Cuánto cuesta sincronizar un origen de datos?

Al sincronizar el índice, tarda dos minutos en calentarse y activarse Amazon EC2 para establecer las conexiones necesarias. No se le cobrará nada durante este proceso. El medidor de uso solo comienza después de que se inicie el trabajo de sincronización. Para obtener más información sobre Amazon Kendra los precios, consulta Amazon Kendra los precios.

Recibo un error Amazon EC2 de autorización

Si se produce un error de operación Amazon EC2 no autorizada durante la sincronización de una fuente de datos de nube privada virtual (VPC), es probable que su IAM función de VPC carezca de los permisos necesarios. Compruebe que la IAM función que utiliza para la fuente de datos tenga los permisos adjuntos. Para obtener más información, consulte IAM Función de nube privada virtual.

No puedo usar los enlaces del índice de búsqueda para abrir mis Amazon S3 objetos

Su Amazon Kendra índice solo puede acceder a los archivos a los que una fuente de Amazon S3 datos le haya otorgado permisos de acceso. Por ejemplo, Amazon Kendra no puede modificar

los Amazon S3 permisos que determinan si un objeto debe ser público o cifrado. Amazon Kendra tampoco tiene los permisos predeterminados para crear o devolver un enlace firmado para Amazon S3 objetos. Si desea activar la vinculación firmada para Amazon S3 los objetos de un Amazon Kendra índice, tiene dos opciones:

- Puede firmar los resultados de la consulta del índice con el objeto URI de origen antes de devolver el resultado a la página de búsqueda. Para ver un step-by-step resumen de este proceso, consulte Compartir objetos mediante direcciones URL prefirmadas.
- Puedes anular el uri de la fuente de metadatos del Amazon S3 objeto y hacer que tu servicio esté disponible a través de una red de entrega de CloudFront contenido (CDN) conectada a un bucket. Amazon S3 O bien, puedes usar un punto final API Gateway proxy que devuelva una URL prefirmada y la redirija a ella.

Aparece un mensaje de error AccessDenied al usar un archivo de certificado SSL

Si aparece un error de acceso denegado al utilizar un certificado SSL con su fuente de datos, asegúrese de que su IAM función tenga el permiso para acceder al archivo del certificado SSL en la ubicación especificada. Si el certificado está cifrado con una AWS KMS clave, su IAM función también debe tener permiso para descifrar mediante la AWS KMS clave. Para obtener más información, consulte Autenticación y control de acceso para AWS KMS.

Aparece un error de autorización al utilizar una fuente de SharePoint datos

Si se produce un error de autorización al sincronizar el índice con una fuente de SharePoint datos, confirme que se le ha asignado una función de administrador del sitio. SharePoint

Mi índice no rastrea los documentos de mi origen de datos de Confluence

Si tu Amazon Kendra índice no rastrea documentos de tu fuente de datos de Confluence durante el proceso de sincronización, confirma que formas parte de los grupos de administradores de Confluence.

Solución de problemas con los resultados de búsqueda de documentos

Esta sección puede ayudarte a solucionar problemas en los resultados de búsqueda. Amazon Kendra

Los resultados de búsqueda no son relevantes para mi consulta de búsqueda

Si los resultados de búsqueda parecen irrelevantes, puede deberse a los siguientes motivos:

- Se incluyen resultados con confianza `LOW` en los resultados. Puedes filtrar los resultados con `LOW` confianza utilizando el `ScoreAttributes` campo QueryResultItems para excluir cualquier resultado con un valor de`LOW`. Amazon Kendra asigna a cada resultado un valor de intervalo de confianza igual o igual a `VERY_HIGH`HIGH, `MEDIUM` y`LOW`. Estos valores indican el nivel de confianza en que un resultado es relevante para una consulta. Además, independientemente de los niveles de confianza, Amazon Kendra devuelve tres tipos de resultados en el siguiente orden: `ANSWER` (extracto de la respuesta sugerida), `QUESTION_ANSWER` (preguntas frecuentes) y `DOCUMENT` (extracto del documento). Por lo tanto, es posible que un resultado con confianza `LOW` del tipo `QUESTION_ANSWER` se posicione por encima de otro resultado con confianza `VERY_HIGH` del tipo `DOCUMENT`. Sin embargo, no siempre es necesariamente cierto que la confianza `LOW` del tipo `QUESTION_ANSWER` sea un resultado mejor que la confianza `VERY_HIGH` del tipo `DOCUMENT`.
- Algunos campos o atributos de metadatos tienen un valor muy alto, lo que afecta a la clasificación de los resultados. Amazon Kendra busca en el índice mediante varios parámetros, como el título del documento, el texto, la fecha y los campos o atributos de texto personalizados. Puede probar con diferentes valores de priorización para obtener los mejores resultados en todas las consultas. También puede utilizar el ajuste de relevancia dinámico de consulta para usar diferentes valores de priorización para cada consulta.
- Los usuarios utilizan términos especializados cuando buscan información y no hay sinónimos personalizados configurados en el índice para gestionar estos términos especializados. Para obtener más información sobre cómo y cuándo usar sinónimos, consulte Añadir sinónimos personalizados a un índice.

¿Por qué solo veo 100 resultados?

Amazon Kendra devuelve el recuento total de los documentos relevantes. De forma predeterminada, se muestran los 100 primeros por consulta. Los resultados están paginados. Puede utilizar `PageNumber` para acceder a diferentes páginas.

Puede configurarlo Amazon Kendra para que devuelva hasta 1000 documentos o resultados de búsqueda por consulta, con un máximo de 100 resultados por página. Si quiere obtener más de 100 resultados, puede solicitarlo poniéndose en contacto con Soporte de cuotas. Aumentar el número de resultados de búsqueda podría afectar a la latencia.

¿Por qué faltan los documentos que espero ver?

Amazon Kendra admite listas de control de acceso (ACL) basadas en usuarios y grupos. Amazon Kendra ingiere las políticas de ACL a través de conectores. Si un índice no configura una ACL, solo se mostrarán los documentos que coincidan con el filtro de atributos para usuario y grupo. Si se proporciona un filtro de atributos de usuario o grupo, no se mostrarán los documentos sin una ACL.

Si utiliza un control de acceso basado en tokens, se mostrarán los documentos sin una política de ACL y los documentos que coincidan con los usuarios y los grupos.

¿Por qué veo documentos que tienen una política de ACL?

Si un índice no configura una política de control de acceso, el filtro puede proporcionar los usuarios y los grupos. Si no se aplica ningún filtro de usuarios y grupos, se devolverán todos los documentos relacionados. Se ignorará cualquier política de ACL.

Solución de problemas generales

Amazon Kendra utiliza CloudWatch métricas y registros para proporcionar información sobre la sincronización de las fuentes de datos. Puede usar las métricas y los registros para determinar qué ha fallado en una ejecución de sincronización y cómo solucionarlo.

Para solucionar problemas generales, comience con sus CloudWatch métricas.

- Compruebe la métrica `DocumentsCrawled` para ver cuántos documentos ha comprobado su origen de datos. En el caso de un Amazon S3 segmento, si el número es inferior al esperado, compruebe que la fuente de datos apunta al segmento correcto.

- Compruebe la métrica `DocumentsSkippedNoChange` para ver cuántos documentos se omitieron porque no han cambiado desde la última sincronización. Si el número no coincide con el esperado, compruebe que su repositorio se haya actualizado correctamente.
- Compruebe la métrica `DocumentsSkippedInvalidMetadata` para ver cuántos documentos tenían metadatos no válidos. Revise sus CloudWatch registros para ver los errores específicos que se produjeron.
- Compruebe la métrica `DocumentsSubmittedForIndexingFailed` para ver cuántos documentos se han enviado desde el origen de datos al índice pero no se han podido indexar. Por ejemplo, si utiliza un atributo de metadatos en un origen de datos de Amazon S3 que no se ha definido como un campo de índice personalizado, el documento no se indexará. Revisa tus CloudWatch registros para ver los errores específicos que se produjeron.
- Compruebe la métrica `DocumentsSubmittedForDeletionFailed` para ver cuántos documentos el origen de datos ha intentado eliminar del índice y no se han podido eliminar. Revisa tus CloudWatch registros para ver los errores específicos que se produjeron.

Puedes consultar los CloudWatch registros de una ejecución de sincronización concreta para obtener detalles de los errores que se produjeron durante la ejecución. Para obtener más información sobre CloudWatch los registros con Amazon Kendra, consulte CloudWatch Logs.

Intelligent Ranking Amazon Kendra

Intelligent Ranking Amazon Kendra utiliza capacidades de búsqueda semántica Amazon Kendra para volver a clasificar de forma inteligente los resultados de un servicio de búsqueda.

Temas

- Amazon Kendra Clasificación inteligente para autogestión OpenSearch
- Clasificación semántica de los resultados de un servicio de búsqueda

Amazon Kendra Clasificación inteligente para autogestión OpenSearch

Puedes aprovechar las capacidades Amazon Kendra de búsqueda semántica que ofrece para mejorar los resultados de búsqueda desde OpenSearchel servicio de búsqueda autogestionado de código abierto basado en la licencia Apache 2.0. El complemento Amazon Kendra Intelligent Ranking reclasifica OpenSearch semánticamente los resultados utilizando. Amazon Kendra Para ello, comprende el significado y el contexto de una consulta de búsqueda utilizando campos específicos, como el cuerpo o el título del documento, de los resultados de OpenSearch búsqueda predeterminados.

Tomemos, por ejemplo, esta consulta: “dirección principal de la nota clave”. Dado que «dirección» tiene varios significados, Amazon Kendra puede deducir el significado de la consulta para obtener información relevante alineada con el significado deseado. En este contexto, se trata del discurso de apertura de una conferencia. Un servicio de búsqueda más simple podría no tener en cuenta la intención y arrojar resultados para una dirección postal de Main Street, por ejemplo.

El complemento Intelligent Ranking OpenSearch está disponible para la versión 2.4.0 y versiones posteriores OpenSearch (autogestionables). Puedes instalar el complemento mediante un script Bash de inicio rápido para crear una nueva imagen de Docker OpenSearch con el complemento Intelligent Ranking incluido. Vea Configuración del complemento de búsqueda inteligente: este es un ejemplo de una configuración para ponerse en marcha rápidamente.

Cómo funciona el complemento de búsqueda inteligente

El proceso general del complemento Intelligent Ranking para OpenSearch (autogestionado) es el siguiente:

1. Un OpenSearch usuario realiza una consulta y OpenSearch proporciona una respuesta a la consulta o una lista de documentos que son relevantes para la consulta.
2. El complemento Intelligent Ranking toma la respuesta a la consulta y extrae la información de los documentos.
3. El complemento Intelligent Ranking realiza una llamada a la API Rescore de Amazon Kendra Intelligent Ranking.
4. La API `Rescore` toma la información extraída de los documentos y reclasifica semánticamente los resultados de la búsqueda.
5. La API `Rescore` devuelve los resultados de búsqueda reclasificados al complemento. El complemento reorganiza los resultados de búsqueda en la respuesta de OpenSearch búsqueda para reflejar la nueva clasificación semántica.

El complemento de clasificación inteligente vuelve a clasificar los resultados utilizando los campos “cuerpo” y “título”. Estos campos del plugin se pueden asignar a los campos del OpenSearch índice que mejor se ajusten a la definición de cuerpo y título de un documento. Por ejemplo, si su índice contiene capítulos de un libro con campos como “chapter_heading” y “chapter_contents”, puede asignar el primero a “title” y el segundo a “body” para obtener los mejores resultados.

Configuración del complemento de búsqueda inteligente

A continuación, se describe cómo configurarlo rápidamente OpenSearch (autogestionado) con el complemento Intelligent Ranking.

Configuración OpenSearch (autogestionada) con el complemento Intelligent Ranking (configuración rápida)

Si ya utilizas la imagen de Docker`opensearch:2.4.0`, puedes usar este Dockerfile para crear una nueva imagen de la OpenSearch versión 2.4.0 con el complemento Intelligent Ranking. Incluya un contenedor para la nueva imagen en su archivo docker-compose.yml u opensearch.yml. También debe incluir el ID del plan de ejecución de rescate generado al crear un plan de ejecución de rescate, junto con la información sobre la región y el punto de conexión; consulte el paso 2 para crear un plan de ejecución de rescate.

Si anteriormente descargó una versión de la imagen de Docker `opensearch` anterior a la 2.4.0, debe usar la imagen de Docker `opensearch:2.4.0` o una versión posterior y crear una nueva imagen con el complemento Intelligent Ranking incluido.

1. Descargue e instale Docker Desktop para su sistema operativo. Docker Desktop incluye Docker Compose y Docker Engine. Se recomienda comprobar si el ordenador cumple los requisitos del sistema mencionados en los detalles de instalación de Docker.

   También puede aumentar los requisitos de uso de memoria en la configuración de su escritorio Docker. Usted es responsable de los requisitos de uso de Docker fuera de los límites de uso disponibles de forma gratuita para los servicios de Docker. Consulte las suscripciones de Docker.

   Compruebe que el estado de Docker Desktop es “en ejecución”.

2. Aprovisione Amazon Kendra Intelligent Ranking y sus requisitos de capacidad. Una vez que aprovisione Intelligent Ranking Amazon Kendra , se le cobrará por hora en función de las unidades de capacidad establecidas. Consulte la información sobre los niveles y precios gratuitos.

   Utiliza la CreateRescoreExecutionPlanAPI para aprovisionar el`Rescore API`. Si no necesita más unidades de capacidad que la unidad predeterminada, no añada más unidades y proporcione solo un nombre para su plan de ejecución de rescore. También puede actualizar sus requisitos de capacidad mediante la UpdateRescoreExecutionPlanAPI. Para más información, consulte Clasificación semántica de los resultados de un servicio de búsqueda.

   Si lo desea, puede ir al paso 3 para crear un plan de ejecución de rescore predeterminado al ejecutar el script Bash de inicio rápido.

   Anote en el paso 4 el identificador del plan de ejecución de rescore incluido en la respuesta.

   CLI

   ```
   aws kendra-ranking create-rescore-execution-plan \
     --name MyRescoreExecutionPlan \
     --capacity-units '{"RescoreCapacityUnits":<integer number of additional
    capacity units>}'

   Response:

   {
       "Id": "<rescore execution plan ID>",
       "Arn": "arn:aws:kendra-ranking:<region>:<account-id>:rescore-execution-plan/
   <rescore-execution-plan-id>"
   }
   ```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra_ranking = boto3.client("kendra-ranking")

print("Create a rescore execution plan.")

# Provide a name for the rescore execution plan
name = "MyRescoreExecutionPlan"
# Set your required additional capacity units
# Don't set capacity units if you don't require more than 1 unit given by
 default
capacity_units = 1

try:
    rescore_execution_plan_response =
 kendra_ranking.create_rescore_execution_plan(
        Name = name,
        CapacityUnits = {"RescoreCapacityUnits":capacity_units}
    )

    pprint.pprint(rescore_execution_plan_response)

    rescore_execution_plan_id = rescore_execution_plan_response["Id"]

    print("Wait for Amazon Kendra to create the rescore execution plan.")

    while True:
        # Get the details of the rescore execution plan, such as the status
        rescore_execution_plan_description =
 kendra_ranking.describe_rescore_execution_plan(
            Id = rescore_execution_plan_id
        )
        # When status is not CREATING quit.
        status = rescore_execution_plan_description["Status"]
        print(" Creating rescore execution plan. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break
```

```
except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

3. Descarga el script de inicio rápido de Bash GitHub para tu versión de OpenSearch seleccionando la rama de versión en el menú desplegable de la rama principal.

   Este script utiliza imágenes de Docker OpenSearch y OpenSearch paneles de control con la versión que haya seleccionado en el GitHub repositorio del script. Descarga un archivo zip para el complemento Intelligent Ranking y genera un archivo `Dockerfile` para crear una nueva imagen de Docker OpenSearch que incluya el complemento. También crea un archivo docker-compose.yml que incluye contenedores para OpenSearch el complemento Intelligent Ranking y los paneles de control. OpenSearch El script agrega el ID del plan de ejecución de rescore, la información de la región y el punto final (usa la región) al archivo docker-compose.yml. A continuación, el script se ejecuta `docker-compose up` para iniciar los contenedores con Intelligent Ranking incluido y los paneles de control. OpenSearch OpenSearch Para detener los contenedores sin quitarlos, ejecute `docker-compose stop`. Para retirar los contenedores, ejecute `docker-compose down`.

4. Abra su terminal y en el directorio del script Bash, ejecute el siguiente comando.

   ```
   bash search_processing_kendra_quickstart.sh -p <execution-plan-id> -r <region>
   ```

   Al ejecutar este comando, proporciona el ID del plan de ejecución de rescore que anotó en el paso 2 al aprovisionar Amazon Kendra Intelligent Ranking, junto con la información de su región. Si lo desea, puede aprovisionar Intelligent Ranking Amazon Kendra mediante la opción `--create-execution-plan`. Esto crea un plan de ejecución para volver a puntuar con un nombre y una capacidad predeterminados.

   Para no perder el índice cuando se elimine el contenedor efímero predeterminado, puede hacer que el índice se mantenga durante las ejecuciones proporcionando el nombre del volumen de datos mediante la opción `--volume-name`. Si ha creado un índice anteriormente, puede especificar el volumen en el archivo docker-compose.yml u opensearch.yml. Para dejar sus volúmenes intactos, no ejecute`docker-compose down -v`.

   El script Bash de inicio rápido configura tus AWS credenciales en el OpenSearch almacén de claves para conectarte a Intelligent Ranking. Amazon Kendra Para proporcionar sus

AWS credenciales al script, utilice la `--profile` opción para especificar el perfil. AWS Si no se especifica la `--profile` opción, el script Bash de inicio rápido intentará leer AWS las credenciales (clave de acceso/secreta, token de sesión opcional) de las variables de entorno y, a continuación, del perfil predeterminado. AWS Si no se especifica la `--profile` opción y no se encuentra ninguna credencial, el script no pasará las credenciales al almacén de claves. OpenSearch Si no se especifica ninguna credencial en el OpenSearch almacén de claves, el complemento sigue comprobando las credenciales en la cadena de proveedores de credenciales predeterminada, incluidas las credenciales de Amazon ECS contenedor o las credenciales de perfil de instancia entregadas a través del Amazon EC2 servicio de metadatos.

Asegúrese de haber creado un IAM rol con los permisos necesarios para invocar Intelligent Ranking Amazon Kendra . El siguiente es un ejemplo de una IAM política que permite el uso de la `Rescore` API para un plan de ejecución de rescore específico:

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra-ranking:Rescore",
            "Resource": "arn:aws:kendra-ranking:${Region}:${Account}:rescore-
execution-plan/${RescoreExecutionPlanId}"
        }
    ]
}
```

Ejemplo de docker-compose.yml

Un ejemplo de un archivo docker-compose.yml que utiliza OpenSearch 2.4.0 o una versión posterior con el complemento Intelligent Ranking y Dashboards 2.4.0 o una versión posterior. OpenSearch

```
version: '3'
networks:
  opensearch-net:
volumes:
  <volume-name>:
services:
  opensearch-node:
    image: <Docker image tag name of OpenSearch with Intelligent Ranking plugin>
    container_name: opensearch-node
```

```
    environment:
      - cluster.name=opensearch-cluster
      - node.name=opensearch-node
      - discovery.type=single-node
      - kendra_intelligent_ranking.service.endpoint=https://kendra-
ranking.<region>.api.aws
      - kendra_intelligent_ranking.service.region=<region>
      - kendra_intelligent_ranking.service.execution_plan_id=<rescore-execution-plan-
id>
    ulimits:
      memlock:
        soft: -1
        hard: -1
      nofile:
        soft: 65536
        hard: 65536
    ports:
      - 9200:9200
      - 9600:9600
    networks:
      - opensearch-net
  volumes:
    <docker-volume-name>:/usr/share/opensearch/data
  opensearch-dashboard:
   image: opensearchproject/opensearch-dashboards:<your-version>
   container_name: opensearch-dashboards
   ports:
     - 5601:5601
   environment:
     OPENSEARCH_HOSTS: '["https://opensearch-node:9200"]'
   networks:
     - opensearch-net
```

Ejemplo de un Dockerfile y creación de una imagen

Un ejemplo de uso de la versión 2.4.0 o posterior con el complemento Intelligent `Dockerfile` Ranking. OpenSearch

```
FROM opensearchproject/opensearch:<your-version>
RUN /usr/share/opensearch/bin/opensearch-plugin install --batch  https://github.com/
opensearch-project/search-processor/releases/download/<your-version>/search-
processor.zip
```

Creación de una imagen de Docker OpenSearch con el complemento Intelligent Ranking.

```
docker build --tag=<Docker image tag name of OpenSearch with Intelligent Ranking
 plugin>
```

Interactuar con el complemento de búsqueda inteligente

Una vez que lo hayas configurado OpenSearch (autogestionado) con el plugin Intelligent Ranking, podrás interactuar con el plugin mediante comandos curl o bibliotecas OpenSearch cliente. Las credenciales predeterminadas para acceder OpenSearch con el complemento Intelligent Ranking son el nombre de usuario «admin» y la contraseña «admin».

Para aplicar la configuración del complemento Intelligent Ranking a un OpenSearch índice:

Curl

```
curl -XPUT "https://localhost:9200/<your-docs-index>/_settings" -u 'admin:admin' --
insecure -H 'Content-Type: application/json' -d'
{
  "index": {
    "plugin" : {
      "searchrelevance" : {
        "result_transformer" : {
          "kendra_intelligent_ranking": {
              "order": 1,
              "properties": {
                "title_field": "title_field_name_here",
                "body_field": "body_field_name_here"
              }
          }
        }
      }
    }
  }
}
'
```

Python

```
pip install opensearch-py

from opensearchpy import OpenSearch
```

```
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

setting_body = {
    "index": {
        "plugin" : {
            "searchrelevance" : {
                "result_transformer" : {
                    "kendra_intelligent_ranking": {
                            "order": 1,
                            "properties": {
                                "title_field": "title_field_name_here",
                                "body_field": "body_field_name_here"
                            }
                    }
                }
            }
        }
    }
}

response = client.indices.put_settings(index_name, body=setting_body)
```

Debe incluir el nombre del campo de texto principal que desea utilizar para volver a clasificar, como un campo de cuerpo del documento o de contenido del documento. También puede incluir otros campos de texto, como el título del documento o el resumen del documento.

Ahora puede realizar cualquier consulta y los resultados se clasifican mediante el complemento Intelligent Ranking.

Curl

```
curl -XGET "https://localhost:9200/<your-docs-index>/_search?pretty" -u
 'admin:admin' --insecure -H 'Content-Type: application/json' -d'
{
  "query" : {
    "match" : {
      "body_field_name_here": "intelligent systems"
    }
  }
}
'
```

Python

```
from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

query = {
  'size': 10,
  "query" : {
    "match" : {
      "body_field_name_here": "intelligent systems"
    }
  }
```

```
}

response = client.search(
    body = query,
    index = index_name
)

print('\nSearch results:')
print(response)
```

Para eliminar la configuración del complemento Intelligent Ranking de un OpenSearch índice:

Curl

```
curl -XPUT "http://localhost:9200/<your-docs-index>/_settings" -H 'Content-Type:
 application/json' -d'
{
  "index": {
    "plugin": {
      "searchrelevance": {
        "result_transformer": {
          "kendra_intelligent_ranking.*": null
        }
      }
    }
  }
}
'
```

Python

```
from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
```

```
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

setting_body = {
  "index": {
    "plugin": {
      "searchrelevance": {
        "result_transformer": {
          "kendra_intelligent_ranking.*": null
        }
      }
    }
  }
}

response = client.indices.put_settings(index_name, body=setting_body)
```

Para probar el complemento Intelligent Ranking en una consulta determinada o para probarlo en determinados campos del cuerpo y del título:

Curl

```
curl -XGET "https://localhost:9200/<your-docs-index>/_search?pretty" -u
 'admin:admin' --insecure -H 'Content-Type: application/json' -d'
{
  "query": {
    "multi-match": {
      "query": "intelligent systems",
      "fields": ["body_field_name_here", "title_field_name_here"]
    }
  },
  "size": 25,
  "ext": {
    "search_configuration": {
      "result_transformer": {
        "kendra_intelligent_ranking": {
          "order": 1,
          "properties": {
```

```
            "title_field": "title_field_name_here",
            "body_field": "body_field_name_here"
          }
        }
      }
    }
  }
}
'
```

Python

```
from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

# Index settings null for kendra_intelligent_ranking

query = {
  "query": {
    "multi_match": {
      "query": "intelligent systems",
      "fields": ["body_field_name_here", "title_field_name_here"]
    }
  },
  "size": 25,
  "ext": {
    "search_configuration": {
      "result_transformer": {
```

```
        "kendra_intelligent_ranking": {
          "order": 1,
          "properties": {
            "title_field": "title_field_name_here",
            "body_field": "body_field_name_here"
          }
        }
      }
    }
  }
}

response = client.search(
    body = query,
    index = index_name
)

print('\nSearch results:')
print(response)
```

Comparar OpenSearch los resultados con Amazon Kendra los resultados

Puede comparar los resultados clasificados side-by-side OpenSearch (autogestionados) con Amazon Kendra los resultados reclasificados. OpenSearch La versión 2.4.0 y posteriores de Dashboards ofrecen side-by-side resultados para que puedas comparar la OpenSearch clasificación de los documentos con la forma en que el plugin clasifica los documentos para una consulta de búsqueda. Amazon Kendra

Antes de poder comparar los resultados OpenSearch clasificados con los Amazon Kendra reclasificados, asegúrate de que tus OpenSearch paneles estén respaldados por un OpenSearch servidor con el complemento Intelligent Ranking. Puede configurarlo utilizando Docker y un script Bash de inicio rápido. Consulte Configuración del complemento de búsqueda inteligente.

A continuación, se describe cómo comparar OpenSearch y Amazon Kendra buscar los resultados en OpenSearch los paneles de control. Para obtener más información, consulte la OpenSearchdocumentación.

Comparación de los resultados de búsqueda en los OpenSearch paneles

1. Abre http://localhost:5601 e inicia sesión en OpenSearch Dashboards. Las credenciales predeterminadas son el nombre de usuario “admin” y la contraseña “admin”.

2. Selecciona Relevancia de búsqueda en los OpenSearch complementos del menú de navegación.
3. Escriba el texto de búsqueda en la barra de búsqueda.
4. Seleccione su índice para la consulta 1 e introduzca una consulta en la OpenSearch consulta DSL. Puede usar la variable `%SearchText%` para hacer referencia al texto de búsqueda que ingresó en la barra de búsqueda. Para ver un ejemplo de esta consulta, consulte OpenSearch la documentación. Los resultados devueltos para esta consulta son los OpenSearch resultados sin utilizar el complemento Intelligent Ranking.
5. Seleccione el mismo índice para la consulta 2 e introduzca la misma consulta en la OpenSearch consulta DSL. Además, incluya la extensión `kendra_intelligent_ranking` y especifique la extensión obligatoria `body_field` para clasificarla. También puede especificar el campo de título, pero el campo de cuerpo es obligatorio. Para ver un ejemplo de esta consulta, consulte OpenSearch la documentación. Los resultados devueltos para esta consulta son los resultados Amazon Kendra reclasificados mediante el complemento Intelligent Ranking. El complemento clasifica hasta 25 resultados.
6. Seleccione Buscar para ver y comparar los resultados.

Clasificación semántica de los resultados de un servicio de búsqueda

Amazon Kendra Intelligent Ranking utiliza Amazon Kendra las capacidades de búsqueda semántica para volver a clasificar los resultados de un servicio de búsqueda. Para ello, tiene en cuenta el contexto de la consulta de búsqueda, además de toda la información disponible en los documentos del servicio de búsqueda. Amazon Kendra La clasificación inteligente puede mejorar la coincidencia sencilla de palabras clave.

La CreateRescoreExecutionPlanAPI crea un recurso de clasificación Amazon Kendra inteligente que se utiliza para aprovisionar la API Rescore. La `Rescore` API reclasifica los resultados de búsqueda de un servicio de búsqueda, por ejemplo OpenSearch (autogestionado).

Cuando llama a `CreateRescoreExecutionPlan`, establece las unidades de capacidad necesarias para volver a clasificar los resultados de un servicio de búsqueda. Si no necesita más unidades de capacidad que la unidad única por defecto, no la cambie. Proporcione solo un nombre para su plan de ejecución de rescore. Puede establecer hasta 1000 unidades adicionales. Para obtener información sobre lo que incluye una unidad de capacidad única, consulte Ajustar la capacidad. Una

vez que Amazon Kendra aprovisione Intelligent Ranking, se le cobrará por hora en función de las unidades de capacidad establecidas. Consulte la información sobre los niveles y precios gratuitos.

Cuando llama a `CreateRescoreExecutionPlan`, se genera un identificador del plan de ejecución de rescore que se devuelve en la respuesta. La API `Rescore` usa el ID del plan de ejecución de rescore para volver a clasificar los resultados de un servicio de búsqueda según la capacidad que haya establecido. Incluya el ID del plan de ejecución de rescore en los archivos de configuración de su servicio de búsqueda. Por ejemplo, si usa OpenSearch (autogestionado), incluye el ID del plan de ejecución de rescore en su archivo docker-compose.yml u opensearch.yml; consulte Clasificación inteligente de los resultados (autoservicio). OpenSearch

También se genera un nombre de recurso de Amazon (ARN) en la respuesta cuando se llama a `CreateRescoreExecutionPlan`. Puede usar este ARN para crear una política de permisos en AWS Identity and Access Management (IAM) para restringir el acceso de los usuarios a un ARN específico para un plan de ejecución de rescore específico. Si desea ver un ejemplo de una IAM política que permite usar la `Rescore` API para un plan de ejecución de rescore específico, consulte Amazon Kendra Intelligent Ranking for self-management. OpenSearch

A continuación se muestra un ejemplo de cómo crear un plan de ejecución de rescore con unidades de capacidad configuradas en 1.

CLI

```
aws kendra-ranking create-rescore-execution-plan \
  --name MyRescoreExecutionPlan \
  --capacity-units '{"RescoreCapacityUnits":1}'

Response:

{
    "Id": "<rescore execution plan ID>",
    "Arn": "arn:aws:kendra-ranking:<region>:<account-id>:rescore-execution-plan/
<rescore-execution-plan-id>"
}
```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time
```

```
kendra_ranking = boto3.client("kendra-ranking")

print("Create a rescore execution plan.")

# Provide a name for the rescore execution plan
name = "MyRescoreExecutionPlan"
# Set your required additional capacity units
# Don't set capacity units if you don't require more than 1 unit given by default
capacity_units = 1

try:
    rescore_execution_plan_response = kendra_ranking.create_rescore_execution_plan(
        Name = name,
        CapacityUnits = {"RescoreCapacityUnits":capacity_units}
    )

    pprint.pprint(rescore_execution_plan_response)

    rescore_execution_plan_id = rescore_execution_plan_response["Id"]

    print("Wait for Amazon Kendra to create the rescore execution plan.")

    while True:
        # Get the details of the rescore execution plan, such as the status
        rescore_execution_plan_description =
 kendra_ranking.describe_rescore_execution_plan(
            Id = rescore_execution_plan_id
        )
        # When status is not CREATING quit.
        status = rescore_execution_plan_description["Status"]
        print(" Creating rescore execution plan. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
import java.util.concurrent.TimeUnit;

import software.amazon.awssdk.services.kendraranking.KendraRankingClient;
import
 software.amazon.awssdk.services.kendraranking.model.CapacityUnitsConfiguration;
import
 software.amazon.awssdk.services.kendraranking.model.CreateRescoreExecutionPlanRequest;
import
 software.amazon.awssdk.services.kendraranking.model.CreateRescoreExecutionPlanResponse;
import
 software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanRequest;
import
 software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanResponse;
import
 software.amazon.awssdk.services.kendraranking.model.RescoreExecutionPlanStatus;

public class CreateRescoreExecutionPlanExample {

  public static void main(String[] args) throws InterruptedException {

    String rescoreExecutionPlanName = "MyRescoreExecutionPlan";
    int capacityUnits = 1;

    KendraRankingClient kendraRankingClient = KendraRankingClient.builder().build();

    System.out.println(String.format("Creating a rescore execution plan named %s",
 rescoreExecutionPlanName));

    CreateRescoreExecutionPlanResponse createResponse =
 kendraRankingClient.createRescoreExecutionPlan(
        CreateRescoreExecutionPlanRequest.builder()
            .name(rescoreExecutionPlanName)
            .capacityUnits(
                CapacityUnitsConfiguration.builder()
                    .rescoreCapacityUnits(capacityUnits)
                    .build()
            )
            .build()
    );

    String rescoreExecutionPlanId = createResponse.id();
```

```
    System.out.println(String.format("Waiting for rescore execution plan with id %s
 to finish creating.", rescoreExecutionPlanId));
    while (true) {
      DescribeRescoreExecutionPlanResponse describeResponse =
 kendraRankingClient.describeRescoreExecutionPlan(
          DescribeRescoreExecutionPlanRequest.builder()
              .id(rescoreExecutionPlanId)
              .build()
      );
      RescoreExecutionPlanStatus rescoreExecutionPlanStatus =
 describeResponse.status();
      if (rescoreExecutionPlanStatus != RescoreExecutionPlanStatus.CREATING) {
        break;
      }
      TimeUnit.SECONDS.sleep(60);
    }

    System.out.println("Rescore execution plan creation is complete.");
  }
}
```

A continuación se muestra un ejemplo de actualización de un plan de ejecución rescore para establecer las unidades de capacidad en 2.

CLI

```
aws kendra-ranking update-rescore-execution-plan \
  --id <rescore execution plan ID> \
  --capacity-units '{"RescoreCapacityUnits":2}'
```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint
import time

kendra_ranking = boto3.client("kendra-ranking")

print("Update a rescore execution plan.")

# Provide the ID of the rescore execution plan
```

```
id = <rescore execution plan ID>
# Re-set your required additional capacity units
capacity_units = 2

try:
    kendra_ranking.update_rescore_execution_plan(
        Id = id,
        CapacityUnits = {"RescoreCapacityUnits":capacity_units}
    )

    print("Wait for Amazon Kendra to update the rescore execution plan.")

    while True:
        # Get the details of the rescore execution plan, such as the status
        rescore_execution_plan_description =
 kendra_ranking.describe_rescore_execution_plan(
            Id = id
        )
        # When status is not UPDATING quit.
        status = rescore_execution_plan_description["Status"]
        print(" Updating rescore execution plan. Status: "+status)
        time.sleep(60)
        if status != "UPDATING":
            break

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
import java.util.concurrent.TimeUnit;

import software.amazon.awssdk.services.kendraranking.KendraRankingClient;
import
 software.amazon.awssdk.services.kendraranking.model.CapacityUnitsConfiguration;
import
 software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanRequest;
import
 software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanResponse;
import
 software.amazon.awssdk.services.kendraranking.model.RescoreExecutionPlanStatus;
```

```
import
 software.amazon.awssdk.services.kendraranking.model.UpdateRescoreExecutionPlanRequest;
import
 software.amazon.awssdk.services.kendraranking.model.UpdateRescoreExecutionPlanResponse;

public class UpdateRescoreExecutionPlanExample {

  public static void main(String[] args) throws InterruptedException {

    String rescoreExecutionPlanId = <rescore execution plan ID>;
    int newCapacityUnits = 2;

    KendraRankingClient kendraRankingClient = KendraRankingClient.builder().build();

    System.out.println(String.format("Updating a rescore execution plan named %s",
 rescoreExecutionPlanId));

    UpdateRescoreExecutionPlanResponse updateResponse =
 kendraRankingClient.updateRescoreExecutionPlan(
        UpdateRescoreExecutionPlanRequest.builder()
            .id(rescoreExecutionPlanId)
            .capacityUnits(
                CapacityUnitsConfiguration.builder()
                    .rescoreCapacityUnits(newCapacityUnits)
                    .build()
            )
            .build()
    );

    System.out.println(String.format("Waiting for rescore execution plan with id %s
 to finish updating.", rescoreExecutionPlanId));
    while (true) {
      DescribeRescoreExecutionPlanResponse describeResponse =
 kendraRankingClient.describeRescoreExecutionPlan(
          DescribeRescoreExecutionPlanRequest.builder()
              .id(rescoreExecutionPlanId)
              .build()
      );
      RescoreExecutionPlanStatus rescoreExecutionPlanStatus =
 describeResponse.status();
      if (rescoreExecutionPlanStatus != RescoreExecutionPlanStatus.UPDATING) {
        break;
      }
      TimeUnit.SECONDS.sleep(60);
```

```
    }

    System.out.println("Rescore execution plan update is complete.");
  }
}
```

A continuación se muestra un ejemplo de utilización de la API `Rescore`.

CLI

```
aws kendra-ranking rescore \
  --rescore-execution-plan-id <rescore execution plan ID> \
  --search-query "intelligent systems" \
  --documents "[{\"Id\": \"DocId1\",\"Title\": \"Smart systems\", \"Body\":
 \"intelligent systems in everyday life\",\"OriginalScore\": 2.0}, {\"Id\":
 \"DocId2\",\"Title\": \"Smarter systems\", \"Body\": \"living with intelligent
 systems\",\"OriginalScore\": 1.0}]"
```

Python

```
import boto3
from botocore.exceptions import ClientError
import pprint

kendra_ranking = boto3.client("kendra-ranking")

print("Use the Rescore API.")

# Provide the ID of the rescore execution plan
id = <rescore execution plan ID>
# The search query from the search service
query = "intelligent systems"
# The list of documents for Intelligent Ranking to rescore
document_list = [
    {"Id": "DocId1", "Title": "Smart systems", "Body": "intelligent systems in
 everyday life", "OriginalScore": 2.0},
    {"Id": "DocId2", "Title": "Smarter systems", "Body": "living with intelligent
 systems", "OriginalScore": 1.0}
]

try:
    rescore_response = kendra_ranking.rescore(
```

```
        rescore_execution_plan_id = id,
        search_query = query,
        documents = document_list
    )

    print(rescore_response["RescoreId"])
    print(rescore_resposne["ResultItems"])

except ClientError as e:
        print("%s" % e)

print("Program ends.")
```

Java

```
import java.util.ArrayList;
import java.util.List;

import software.amazon.awssdk.services.kendraranking.KendraRankingClient;
import software.amazon.awssdk.services.kendraranking.model.RescoreRequest;
import software.amazon.awssdk.services.kendraranking.model.RescoreResponse;
import software.amazon.awssdk.services.kendraranking.model.Document;

public class RescoreExample {

  public static void main(String[] args) {

    String rescoreExecutionPlanId = <rescore execution plan ID>;
    String query = "intelligent systems";

    List<Document> documentList = new ArrayList<>();
    documentList.add(
        Document.builder()
            .id("DocId1")
            .originalScore(2.0F)
            .body("intelligent systems in everyday life")
            .title("Smart systems")
            .build()
    );
    documentList.add(
        Document.builder()
            .id("DocId2")
            .originalScore(1.0F)
```

```
            .body("living with intelligent systems")
            .title("Smarter systems")
            .build()
    );

    KendraRankingClient kendraRankingClient = KendraRankingClient.builder().build();

    RescoreResponse rescoreResponse = kendraRankingClient.rescore(
        RescoreRequest.builder()
            .rescoreExecutionPlanId(rescoreExecutionPlanId)
            .searchQuery(query)
            .documents(documentList)
            .build()
    );

    System.out.println(rescoreResponse.rescoreId());
    System.out.println(rescoreResponse.resultItems());
  }
}
```

Historial de documentos para Amazon Kendra

- Última actualización de la documentación: 27 de febrero de 2024

En la siguiente tabla se describen los cambios importantes en cada versión de Amazon Kendra. Para obtener notificaciones sobre las actualizaciones de esta documentación, puede suscribirse a la fuente RSS.

Cambio	Descripción	Fecha
Nueva característica	Amazon Kendra ahora admite una versión actualizada del conector de la fuente de GitHub datos. Para obtener más información, consulte GitHub.	27 de febrero de 2024
Nueva característica	Amazon Kendra ahora admite versiones actualizadas del conector de fuente de Amazon FSx datos. Para obtener más información, consulte Amazon FSx (Windows) y Amazon FSx (NetAppONTAP).	8 de febrero de 2024
Nueva característica	Amazon Kendra ahora es compatible con una versión actualizada del conector de fuentes de datos de Slack. Para más información, consulte Slack.	11 de enero de 2024
Nueva característica	Amazon Kendra ahora permite contraer y expandir los resultados de búsqueda. Para obtener más información,	19 de octubre de 2023

	consulte Contraer o expandir los resultados de búsqueda.	
Nueva característica	Amazon Kendra ahora admite un conector de fuente de datos Aurora (MySQL). Para más información, consulte Aurora (MySQL).	28 de septiembre de 2023
Nueva característica	Amazon Kendra ahora admite un conector de fuente de datos Aurora (PostgreSQL). Para obtener información, consulte Aurora (PostgreSQL).	28 de septiembre de 2023
Nueva característica	Amazon Kendra ahora admite un conector de fuente de datos Amazon RDS (MySQL). Para más información, consulte Amazon RDS (MySQL).	28 de septiembre de 2023
Nueva característica	Amazon Kendra ahora admite un conector de fuente de datos Amazon RDS (Microsoft SQL Server). Para más información, consulte Amazon RDS (Microsoft SQL Server).	28 de septiembre de 2023
Nueva característica	Amazon Kendra ahora es compatible con un conector de fuente de datos Amazon RDS (Oracle). Para más información, consulte Amazon RDS (Oracle).	28 de septiembre de 2023

Nueva característica	Amazon Kendra ahora admite un conector de fuente de datos Amazon RDS (PostgreSQL). Para obtener información, consulte Amazon RDS (PostgreSQL).	28 de septiembre de 2023
Nueva característica	Amazon Kendra ahora es compatible con un conector de fuente de datos IBM DB2. Para más información, consulte IBM DB2.	28 de septiembre de 2023
Nueva característica	Amazon Kendra ahora es compatible con un conector de fuente de datos de Microsoft SQL Server. Para más información, consulte Microsoft SQL Server.	28 de septiembre de 2023
Nueva característica	Amazon Kendra ahora es compatible con un conector de fuente de datos MySQL. Para más información, consulte MySQL.	28 de septiembre de 2023
Nueva característica	Amazon Kendra ahora es compatible con un conector de fuente de datos de Oracle Database. Para más información, consulte Oracle Database.	28 de septiembre de 2023
Nueva característica	Amazon Kendra ahora es compatible con un conector de fuente de datos PostgreSQL. Para obtener información, consulte PostgreSQL.	28 de septiembre de 2023

Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Drupal. Para más información, consulte Drupal.	6 de septiembre de 2023
Nueva característica	Recupere pasajes semántica mente relevantes utilizando la API Retrieve de Amazon Kendra para sistemas de generación aumentada de recuperación (RAG).	22 de junio de 2023
Nueva característica	Amazon Kendra ahora es compatible con una versión actualizada del conector de fuentes de datos de Amazon Kendra Web Crawler. Para más información, consulte Amazon Kendra Web Crawler v2.0.	21 de junio de 2023
Ampliación de las regiones	Amazon Kendra ya está disponible en Europa (Londres) (eu-west-2).	5 de junio de 2023
Nueva característica	Amazon Kendra ahora es compatible con una versión actualizada del conector de fuente de datos de Alfresco. Para más información, consulte Alfresco.	16 de mayo de 2023

Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Adobe Experience Manager. Para más información, consulte Adobe Experience Manager.	11 de mayo de 2023
Nueva característica	Amazon Kendra ahora admite la configuración de campos y atributos del documento al llamar. GetQuerySuggestions Ahora puede basar las sugerencias de consulta en el contenido de los campos del documento. Para más información, consulte Sugerencias de consulta.	2 de mayo de 2023
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Gmail. Para más información, consulte Gmail.	13 de abril de 2023
Nueva característica	Amazon Kendra ahora es compatible con una versión actualizada del conector de fuentes OneDrive de datos de Microsoft. Para obtener más información, consulte Microsoft OneDrive v2.0.	3 de abril de 2023

Nueva característica	Mejore la visibilidad de los documentos nuevos o promocione determinados documentos cuando los usuarios escriban determina das consultas mediante los Resultados destacados.	30 de marzo de 2023
Nueva característica	Amazon Kendra ahora es compatible con un conector de fuente de datos actualiza do para Microsoft SharePoint. Para obtener más información, consulte Microsoft SharePoint.	2 de marzo de 2023
Nueva característica	Amazon Kendra ahora es compatible con una versión actualizada del conector de fuentes de datos de Confluenc e. Para más información, consulte Confluence.	1 de marzo de 2023
Ampliación de las regiones	Amazon Kendra ya está disponible en Asia Pacífico (Tokio) (ap-northeast-1).	7 de febrero de 2023
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Microsoft Exchange. Para más información, consulte Microsoft Exchange.	12 de enero de 2023

Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Microsoft Yammer. Para más información, consulte Microsoft Yammer.	12 de enero de 2023
Nueva característica	Amazon Kendra ahora admite la indexación de los tipos de documentos RTF, XML, XSLT, MS_EXCEL, CSV, JSON y MD. Para más información, consulte Tipos de documentos.	11 de enero de 2023
Nueva característica	Amazon Kendra ahora es compatible con una versión actualizada del conector de fuentes de datos. Amazon S3 Para obtener más información, consulte Amazon S3.	10 de enero de 2023
Nueva característica	OpenSearchLos resultados de búsqueda (autogestionados) se pueden clasificar semánticamente mediante la clasificación Amazon Kendra inteligente.	9 de enero de 2023
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Microsoft Teams. Para más información, consulte Microsoft Teams.	5 de enero de 2023

Nueva característica	Amazon Kendra tiene un conector de fuente de datos actualizado para Google Drive. Para más información, consulte Google Drive.	5 de enero de 2023
Nueva característica	Amazon Kendra tiene un conector de fuente de datos actualizado para ServiceNow. Para obtener más información, consulte ServiceNow.	21 de diciembre de 2022
Nueva característica	Amazon Kendra tiene un conector de fuente de datos actualizado para Salesforc e. Para más información, consulte Salesforce.	21 de diciembre de 2022
Ampliación de las regiones	Amazon Kendra ya está disponible en Asia Pacífico (Bombay) (ap-south-1).	14 de diciembre de 2022
Nueva característica	La característica de búsqueda tabular de Amazon Kendra permite buscar y extraer respuestas de tablas incrustad as en documentos HTML.	27 de noviembre de 2022
Nueva característica	Amazon Kendra admite la búsqueda semántica para un conjunto selecto de idiomas.	27 de noviembre de 2022
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Dropbox. Para más informaci ón, consulte Dropbox.	27 de septiembre de 2022

Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Zendesk. Para más información, consulte Zendesk.	17 de agosto de 2022
Nueva característica	El control de acceso a nivel de documento ahora se puede volver a configurar después de indexar los documentos. Para más información, consulte Configuración de control de acceso.	14 de julio de 2022
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Alfresco. Para más información, consulte Alfresco.	30 de junio de 2022
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para GitHub. Para obtener más información, consulte GitHub.	2 de junio de 2022
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Jira. Para más información, consulte Jira.	12 de mayo de 2022
Nueva característica	Las facetas anidadas dentro de una faceta se pueden mostrar en los resultados de la búsqueda. Para más información, consulte Facetas.	5 de mayo de 2022

Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Quip. Para más información, consulte Quip.	19 de abril de 2022
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Box. Para más información, consulte Box.	6 de abril de 2022
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Slack. Para más información, consulte Slack.	14 de marzo de 2022
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para Amazon FSx. Para obtener más información, consulte Amazon FSx.	8 de febrero de 2022
AWS actualizaciones de políticas gestionadas: políticas nuevas	Amazon Kendra se han añadido nuevas políticas AWS gestionadas. Para más información, consulte Políticas administradas de AWS para Amazon Kendra.	3 de enero de 2022

Nueva característica	Amazon Kendra la aplicación de búsqueda se puede implementar con unos pocos clics sin necesidad de ningún código de interfaz. Para más información, consulte Implementar una aplicación de búsqueda sin código.	1 de diciembre de 2021
Nueva característica	Puede enriquecer los metadatos y contenido del documento durante el proceso de ingesta de documentos. Para más información, consulte Personalización de los metadatos del documento durante el proceso de ingesta.	1 de diciembre de 2021
Nueva característica	Amazon Kendra ofrece análisis de búsqueda para obtener información útil sobre su aplicación de búsqueda. Para más información, consulte Obtener información con análisis de búsqueda.	1 de diciembre de 2021
Ampliación de las regiones	Amazon Kendra ahora está disponible en AWS GovCloud (US-West) (us-gov-west-1).	13 de octubre de 2021
Nueva característica	Amazon Kendra ahora puede indexar documentos en varios idiomas y filtrar los resultados de búsqueda por idioma. Consulte Agregar documentos en idiomas distintos del inglés y Buscar en otros idiomas.	7 de octubre de 2021

Nueva característica	Amazon Kendra ahora se integra con el directorio de Identity Center para obtener los niveles de acceso de grupos y usuarios para filtrar el contexto de los usuarios. Consulte Configuración de grupos de usuarios para IAM Identity Center.	6 de octubre de 2021
Nuevo tutorial	Amazon Kendra ahora ofrece un tutorial que explica cómo crear una solución de búsqueda enriquecida con metadatos. Consulte Creación de una solución de búsqueda inteligente.	13 de agosto de 2021
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para. Amazon WorkDocs Para obtener más información, consulte Amazon WorkDocs.	20 de julio de 2021
Nueva característica	Amazon Kendra ahora proporciona un rastreador web para rastrear e indexar páginas web. Para más información, consulte Rastreador web.	17 de junio de 2021
Ampliación de las regiones	Amazon Kendra ahora está disponible en Canadá (Central) (ca-central-1).	16 de junio de 2021

Ampliación de las regiones	Amazon Kendra ahora está disponible en EE. UU. East (Ohio) (us-east-2).	7 de junio de 2021
Nueva característica	Amazon Kendra ahora admite sugerencias de consultas , en las que se sugieren a los usuarios consultas populares relacionadas con su búsqueda. Para más información, consulte Sugerir consultas de búsqueda populares.	27 de mayo de 2021
AWS actualizaciones de políticas gestionadas: nuevas políticas	Amazon Kendra se han añadido nuevas políticas AWS gestionadas. Para más información, consulte Políticas administradas de AWS para Amazon Kendra.	27 de mayo de 2021
Ampliación de las regiones	Amazon Kendra ya está disponible en Asia Pacífico (Singapur) (ap-southeast-1).	5 de mayo de 2021
Nueva característica	Amazon Kendra ahora permite ajustar la relevancia de la búsqueda en la consulta al anular las configuraciones de ajuste establecidas a nivel de índice. Para más información, consulte Ajuste de la relevancia de la búsqueda y Ajuste de las respuestas.	20 de abril de 2021

Nueva característica	Amazon Kendra ahora admite la autenticación OAuth 2.0 y el uso de ServiceNow consultas para seleccionar documento s para su indexación. Para obtener más información, consulte. ServiceNow	1 de abril de 2021
Nueva característica	Amazon Kendra ahora admite el aprendizaje increment al para los documentos de preguntas frecuentes. Para más información, consulte Envío de comentarios para un aprendizaje incremental.	17 de febrero de 2021
Nueva característica	Amazon Kendra ahora admite sinónimos de índice. Para más información, consulte Agregar sinónimos a un índice.	10 de diciembre de 2020
Nueva característica	Amazon Kendra ahora proporciona un conector de base de datos para Google Workspace Drive. Para más información, consulte Uso de un origen de datos de Google Workspace Drive.	8 de diciembre de 2020
Nueva característica	Amazon Kendra ahora incluye una JavaScript bibliotec a en la que es más fácil enviar comentarios sobre las consultas Amazon Kendra. Para más información, consulte Enviar comentarios.	8 de diciembre de 2020

Nueva característica	Amazon Kendra ahora es compatible con el control de acceso de los usuarios basado en fichas. Para más informaci ón, consulte Control de acceso a documentos en un índice.	5 de noviembre de 2020
Nueva característica	El conector de fuentes de datos Amazon Kendra de Confluence ahora funciona con la nube de Confluenc e. Para más información, consulte Uso de un origen de datos de Confluence.	5 de noviembre de 2020
Ampliación de las regiones	Amazon Kendra ya está disponible en Asia Pacífico (Sídney) (ap-southeast-2).	2 de noviembre de 2020
Nueva característica	Amazon Kendra ahora proporciona un conector de fuente de datos para el servidor de Confluence. Para más información, consulte Uso de un origen de datos de Confluence.	26 de octubre de 2020
Nueva característica	Amazon Kendra ahora proporciona una fuente de datos que puede utilizar para generar estadísticas para sus conectores personali zados. Para más información, consulte Uso de un origen de datos personalizado.	21 de octubre de 2020

Nueva característica	Amazon Kendra ahora admite atributos personalizados para las preguntas más frecuentes. Para más información, consulte Agregar preguntas y respuestas.	17 de septiembre de 2020
Nueva característica	Amazon Kendra ahora devuelve las puntuaciones de confianza de los resultados de las consultas. Para obtener más información, consulte QueryResultItem.	15 de septiembre de 2020
Nueva característica	AWS CloudFormation ahora admite Amazon Kendra. Para obtener más información, consulte la referencia de tipos de Amazon Kendra recurso - AWS CloudFormation.	10 de septiembre de 2020
Nueva característica	Amazon Kendra añade soporte para AWS PrivateLink. Para más información, consulte Amazon Kendra y puntos de conexión de VPC de tipo interfaz (AWS PrivateLink).	7 de julio de 2020
Nueva guía	Esta es la primera versión de la Guía para desarrolladores de Amazon Kendra .	11 de mayo de 2020

Referencia de la API

La documentación de referencia de la API ahora es una guía aparte.

Glosario de AWS

Para ver la terminología más reciente de AWS, consulte el Glosario de AWS en la Referencia de Glosario de AWS.

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la version original de inglés, prevalecerá la version en inglés.