aws

Guia do Desenvolvedor

Amazon Kendra

Amazon Kendra: Guia do Desenvolvedor

Table of Contents

..... xiii
O que é o Amazon Kendra? 1
Consulta no Amazon Kendra 1
Benefícios do Amazon Kendra 2
Amazon Kendra Edições 2
Definição de preços do Amazon Kendra 4
Você é um usuário iniciante do Amazon Kendra? 4
Como funciona o Amazon Kendra 6
Índice 7
Usando campos de documentos Amazon Kendra reservados ou comuns 7
Pesquisando índices 9
Documentos 9
Tipos ou formatos de documentos 9
Atributos ou campos do documento 12
Fontes de dados 15
Consultas 17
Tags 18
Marcar recursos 19
Restrições de tags 19
Configuração do Amazon Kendra 20
Inscreva-se para AWS 20
Regiões e endpoints 21
Configurando o AWS CLI 21
Configurando os AWS SDKs 22
IAM funções de acesso para Amazon Kendra 23
IAM funções para índices 23
IAM funções para a BatchPutDocument API 27
IAM funções para fontes de dados 29
Função de nuvem privada virtual (VPC) IAM 121
IAM funções para perguntas frequentes (FAQs) 123
IAM funções para sugestões de consulta 124
IAM funções para mapeamento principal de usuários e grupos 126
IAM funções para AWS IAM Identity Center 128
IAM funções para Amazon Kendra experiências 129

IAM funções para enriquecimento personalizado de documentos 132
Implantação de Amazon Kendra 136
Visão geral 137
Pré-requisitos 137
Configurar o exemplo 138
Página de pesquisa principal 139
Componente de pesquisa 139
Componente de resultados 139
Componente de facetas 139
Componente de paginação 140
Implantação de uma aplicação de pesquisa sem código 140
Como a pesquisa do Experience Builder funciona 140
Projete e ajuste a experiência de pesquisa 141
Fornecer acesso à sua página de pesquisa 142
Configurando uma experiência de pesquisa 143
Ajustar a capacidade 148
Visualizar a capacidade 149
Adicionar e remover capacidade 149
Amazon Kendra Capacidade de classificação inteligente 150
Capacidade para sugestões de consulta 150
Amazon Kendra capacidade de experiência 151
Capacidade para experiência de pesquisa 151
Expansão de consultas adaptável 151
Conceitos básicos 152
Pré-requisitos 152
Inscreva-se para um Conta da AWS 152
Criar um usuário com acesso administrativo 153
Amazon Kendra recursos: AWS CLI, SDK, console 154
Introdução ao Amazon Kendra console 160
Conceitos básicos (AWS CLI) 161
Conceitos básicos (SDK para Python) 163
Conceitos básicos (SDK for Java) 166
Conceitos básicos do S3 (console) 170
Introdução ao MySQL (console) 172
Introdução a uma fonte de identidade do IAM Identity Center (console) 174
Para alterar uma fonte de identidade do IAM Identity Center 177

Criar um índice 179
Como adicionar documentos diretamente a um índice com o upload em lote. 184
Adicionar documentos com a BatchPutDocument API 185
Adicionar documentos de um bucket do S3 187
Adicionar perguntas frequentes a um índice 190
Criação de campos de índice para um arquivo de perguntas frequentes 191
Arquivo CSV básico 192
Arquivo CSV personalizado 192
Arquivo JSON 194
Usando seu arquivo de perguntas frequentes 196
Arquivos de perguntas frequentes em idiomas diferentes do inglês 198
Criação de campos de documentos personalizados 198
Atualização de campos de documentos personalizados 199
Controle do acesso de usuários a documentos por token 202
Usando o OpenID 203
Como usar um JSON Web Token (JWT) com uma senha compartilhada 206
Usando um JSON Web Token (JWT) com uma chave pública 209
Usar o JSON 213
Criar um conector de fonte de dados 216
Definindo um cronograma de atualização 217
Configurações de idioma 217
Conectores de fontes de dados 218
Esquemas de modelos de fonte de dados 219
Adobe Experience Manager 605
Alfresco 615
Aurora (MySQL) 624
Aurora (PostgreSQL) 633
Amazon FSx (Windows) 641
Amazon FSx (EM UM NetApp TOQUE) 650
Amazon RDS/Aurora 659
Amazon RDS (Microsoft SQL Server) 668
Amazon RDS (MySQL) 677
Amazon RDS (Oracle) 685
Amazon RDS (PostgreSQL) 694
Amazon S3 702
Amazon Kendra Rastreador da Web 720

Amazon WorkDocs 742
Box 747
Confluence 755
Conectores de fontes de dados personalizados 776
Dropbox 785
Drupal 794
GitHub 804
Gmail 816
Google Drive 825
IBM DB2 844
Jira 853
Microsoft Exchange 860
Microsoft OneDrive 869
Microsoft SharePoint 885
Microsoft SQL Server 921
Microsoft Teams 929
Microsoft Yammer 941
MySQL 949
Oracle Database 957
PostgreSQL 966
Quip 974
Salesforce 981
ServiceNow 999
Slack 1019
Zendesk 1030
Mapeando campos de fontes de dados 1038
Usando campos de documentos Amazon Kendra reservados ou comuns 7
Adicionar documentos em outros idiomas além do inglês 1044
Configurando Amazon Kendra para usar um Amazon VPC 1046
Configurando Amazon VPC 1047
Conectando-se a Amazon VPC 1050
Conectar-se a um banco de dados 1051
Solução de problemas de conexão VPC 1054
Excluindo um índice, uma fonte de dados ou documentos carregados em lote 1057
Excluir um índice 1057
Excluir uma fonte de dados 1058

Excluindo documentos carregados em lote ... 1060
Enriquecendo seus documentos durante a absorção ... 1062
Como funciona o enriquecimento personalizado de documentos ... 1062
Operações básicas para alterar metadados ... 1063
Funções do Lambda: extrair e alterar metadados ou conteúdo ... 1071
Contratos de dados para funções do Lambda ... 1080
Formato de documento estruturado ... 1082
Exemplo de uma função do Lambda que adere aos contratos de dados ... 1083
Pesquisando um índice ... 1086
Consultar um índice ... 1086
Pré-requisitos ... 1087
Pesquisar um índice (console) ... 1088
Pesquisar um índice (SDK) ... 1088
Pesquisar um índice (Postman) ... 1090
Pesquisar com sintaxe de consulta avançada ... 1092
Pesquisando em idiomas ... 1097
Recuperar passagens ... 1101
Navegar em um índice ... 1104
Apresentar resultados da pesquisa ... 1107
Pesquisa tabular de HTML ... 1111
Sugestões de consulta ... 1115
Sugestões de consulta usando o histórico de consultas ... 1116
Sugestões de consulta usando campos do documento ... 1122
Bloqueie determinadas consultas ou conteúdos de campos do documento de sugestões ... 1127
Corretor ortográfico de consulta ... 1132
Usar o corretor ortográfico de consulta com limites padrão ... 1133
Filtragem e pesquisa de facetas ... 1134
Facetas ... 1135
Usando atributos do documento para filtrar os resultados da pesquisa ... 1139
Filtrando os atributos de cada documento nos resultados da pesquisa ... 1140
Filtragem no contexto do usuário ... 1141
Filtragem por token de usuário ... 1142
Filtragem por ID de usuário e grupo ... 1143
Filtrando por atributo do usuário ... 1144
Filtragem de contexto do usuário para documentos adicionados diretamente a um índice ... 1146
Filtragem de contexto do usuário para perguntas frequentes ... 1146

Filtragem de contexto do usuário para fontes de dados 1146
Respostas de consulta e tipos de resposta 1165
Respostas de consulta 1165
Tipos de resposta 1169
Ajuste e classificação de respostas 1173
Ajuste de respostas 1174
Classificando respostas 1175
Reduzir/expandir os resultados da consulta 1177
Reduzir os resultados 1179
Escolher um documento primário usando a ordem de classificação 1179
Estratégia-chave do documento ausente 1180
Expandir os resultados 1180
Interações com outros Amazon Kendra recursos 1180
Ajuste de consulta 1182
Ajuste de relevância no nível do índice 1183
Ajuste de relevância no nível da consulta 1184
Obter informações com a análise de pesquisa 1186
Métricas para pesquisa 1186
Taxa de cliques 1187
Taxa de cliques zero 1187
Taxa de resultados da pesquisa 1188
Taxa de resposta instantânea 1188
Principais consultas 1188
Principais consultas com zero cliques 1188
Principais consultas com zero resultados de pesquisa 1189
Documentos mais clicados 1189
Total de consultas 1190
Total de documentos 1190
Exemplo de recuperação de dados métricos 1190
De métricas a informações acionáveis 1192
Visualizar e relatar as análises de pesquisa 1192
Gráfico de consultas totais 1193
Gráfico de taxas de cliques 1193
Gráfico de taxa de zero cliques 1193
Gráfico de taxa de resultados de pesquisa zero 1194
Gráfico de taxa de resposta instantânea 1194

Envio de feedback para aprendizado incremental 1195
Usando a Amazon Kendra JavaScript biblioteca para enviar feedback 1197
Etapa 1: inserir uma tag de script em seu aplicativo Amazon Kendra de pesquisa 1197
Etapa 2: adicionar o token de comentários aos resultados da pesquisa 1199
Etapa 3: testar o script de comentários 1200
Usando a Amazon Kendra API para enviar feedback 1200
Adicionar sinônimos personalizados a um índice 1204
Criando um arquivo de dicionário de sinônimos 1206
Adicionar um dicionário de sinônimos a um índice 1208
Atualizando um dicionário de sinônimos 1213
Atualizando um dicionário de sinônimos 1217
Destaques nos resultados da pesquisa 1219
Tutorial: consulte criando uma solução de pesquisa inteligente. 1220
Pré-requisitos 1221
Etapa 1: adicionar documentos 1222
Baixar o conjunto de dados de amostra 1223
Como criar um bucket do Amazon S3 1224
Criação de pastas de dados e metadados no bucket do S3 1227
Carregue os dados de entrada 1230
Etapa 2: detectar entidades 1232
executando um trabalho de análise de entidades no Amazon Comprehend 1232
Etapa 3: Formatar os metadados 1241
Baixando e extraindo a saída do Amazon Comprehend 1242
Carregando a saída no bucket do S3 1245
Conversão da saída para o formato de metadados do Amazon Kendra 1247
Como limpar o bucket do Amazon S3 1252
Etapa 4: criar um índice de e ingerir os metadados. 1254
Criar um índice do Amazon Kendra 1254
Atualizar o perfil do IAM para acessar o Amazon S3 1262
Criação de campos de índice de pesquisa personalizados do Amazon Kendra 1266
Adicionar um bucket do Amazon S3 como fonte de dados para o índice 1271
Sincronizar o índice do Amazon Kendra 1275
Etapa 5: consultar o índice 1278
Consulte o índice do Amazon Kendra 1279
Filtrar os resultados de pesquisa 1285
Etapa 5: limpar 1289

Como limpar os arquivos 1289
..... 1290
Monitorar e registrar em log 1292
Monitorando índices 1292
Monitorar chamadas de API do Amazon Kendra com o CloudTrail 1296
Informações sobre o Amazon Kendra no CloudTrail 1296
Exemplo: entradas de arquivo de log do Amazon Kendra 1297
Monitorar chamadas de API do Amazon Kendra Intelligent Ranking com o CloudTrail 1298
Informações do Amazon Kendra Intelligent Ranking no CloudTrail 1299
Exemplo: entradas de arquivos de log do Amazon Kendra Intelligent Ranking 1300
Monitoramento do Amazon Kendra com o CloudWatch 1301
Visualizar métricas do Amazon Kendra 1301
Criar um alarme 1302
Métricas do CloudWatch para trabalhos de sincronização de índices 1303
Métricas para fontes de dados do Amazon Kendra 1305
Métricas para documentos indexados 1307
Monitoramento do Amazon Kendra com o CloudWatch Logs 1308
Fluxos de log da fonte de dados 1309
Fluxos de log de documentos 1310
Segurança 1312
Proteção de dados 1313
Criptografia em repouso 1314
Criptografia em trânsito 1314
Gerenciamento de chaves 1314
Enpoints da VPC (AWS PrivateLink) 1315
Considerações sobre os endpoints VPC Amazon Kendra e Amazon Kendra Intelligent Ranking 1315
Criação de uma interface VPC endpoint para Amazon Kendra e Amazon Kendra Intelligent Ranking 1315
Criação de uma política de VPC endpoint para Amazon Kendra e Amazon Kendra Intelligent Ranking 1316
Gerenciamento de identidade e acesso 1317
Público 1318
Autenticando com identidades 1319
Gerenciando acesso usando políticas 1322
Como o Amazon Kendra funciona com o IAM 1324

Exemplos de políticas baseadas em identidade ... 1330
AWS políticas gerenciadas ... 1336
Solução de problemas ... 1341
Melhores práticas de segurança ... 1343
Aplicação do princípio de privilégio mínimo ... 1343
Permissões de controle de acesso por perfil (RBAC) ... 1343
Registrar em log e monitorar no Amazon Kendra ... 1343
Validação de conformidade ... 1344
Resiliência ... 1345
Segurança da infraestrutura ... 1346
Análise de configuração e vulnerabilidade ... 1346
Cotas ... 1348
Regiões compatíveis da ... 1348
Cotas ... 1348
Cotas de índice ... 1348
Cotas de conectores de fonte de dados ... 1349
Perguntas frequentes sobre cotas ... 1350
Cotas do dicionário de sinônimos ... 1351
Amazon Kendra cotas de experiência ... 1351
Cotas de resultados de consulta e pesquisa ... 1352
Cotas de sugestões de consulta ... 1353
Cotas de documentos ... 1355
Cotas de resultados de pesquisa em destaque ... 1356
Rescore/reclassifique as cotas dos resultados da pesquisa ... 1357
Solução de problemas ... 1359
Solucionar problemas de origens de dados ... 1359
Meus documentos não foram indexados ... 1359
Meu trabalho de sincronização falhou ... 1360
Meu trabalho de sincronização está incompleto ... 1360
Meu trabalho de sincronização foi bem-sucedido, mas não há documentos indexados ... 1361
Estou enfrentando problemas de formato de arquivo ao sincronizar minha fonte de dados 1362
Quero gerar um relatório de status de sincronização para meus documentos ... 1362
Quanto tempo demora a sincronização de uma fonte de dados? ... 1363
Qual é a cobrança pela sincronização de uma fonte de dados? ... 1363
Estou recebendo um erro Amazon EC2 de autorização ... 1363
Não consigo usar links de índice de pesquisa para abrir meus Amazon S3 objetos ... 1363

Estou recebendo uma mensagem de erro AccessDenied Ao usar o arquivo de certificado SSL 1364
Estou recebendo um erro de autorização ao usar uma fonte SharePoint de dados 1364
Meu índice não rastreia documentos da minha fonte de dados do Confluence 1364
Solucionar problemas de resultados da pesquisa de documentos 1365
Meus resultados de pesquisa não são relevantes para minha consulta de pesquisa 1365
Por que eu só vejo 100 resultados? 1365
Por que os documentos que eu espero ver estão ausentes? 1366
Por que vejo documentos que têm uma política de ACL? 1366
Solução de problemas gerais 1366
Intelligent Ranking do Amazon Kendra 1368
Classificação inteligente para autogestão OpenSearch 1368
Como funciona o plug-in de pesquisa inteligente 1368
Configurando o plug-in de pesquisa inteligente 1369
Interagindo com o plug-in de pesquisa inteligente 1375
Comparando OpenSearch resultados com Amazon Kendra resultados 1381
Classificando semanticamente os resultados de um serviço de pesquisa 1382
Histórico do documento 1392
Referência de API 1409
Glossário do AWS 1410
..... mcdxi

O que é o Amazon Kendra?

O Amazon Kendra é um serviço de pesquisa inteligente que usa processamento de linguagem natural e algoritmos avançados de machine learning para retornar respostas específicas às perguntas de pesquisa de seus dados.

Ao contrário da pesquisa tradicional baseada em palavras-chave, o Amazon Kendra usa seus recursos de compreensão semântica e contextual para decidir se um documento é relevante para uma consulta de pesquisa. Ele retorna respostas específicas às perguntas, oferecendo aos usuários uma experiência próxima à interação com um especialista humano.

> **Note**
>
> Você também poderá usar os recursos do Amazon Kendra de pesquisa semântica para reclassificar os resultados de outro serviço de pesquisa. Consulte Classificação inteligente do Amazon Kendra para obter mais detalhes.

Com o Amazon Kendra, você poderá criar uma experiência de pesquisa unificada conectando vários repositórios de dados a um índice e ingerindo e crawling documentos. Poderá usar os metadados do documento para criar uma experiência de pesquisa personalizada e rica em recursos para seus usuários, ajudando-os a encontrar com eficiência as respostas certas para suas consultas.

O que é Amazon Kendra?

Consulta no Amazon Kendra

Poderá perguntar ao Amazon Kendra os seguintes tipos de consultas:

Perguntas factóides – Perguntas simples sobre quem, o quê, quando ou onde, por exemplo, Onde fica o centro de serviços mais próximo de Seattle? As perguntas factóides têm respostas baseadas em fatos que podem ser retornadas como uma única palavra ou frase. A resposta é obtida de um Perguntas frequentes ou de seus documentos indexados.

Perguntas descritivas – Perguntas em que a resposta pode ser uma frase, uma passagem ou um documento inteiro. Por exemplo, Como faço para conectar meu Echo Plus à minha rede? Ou: Como faço para obter benefícios fiscais para famílias de baixa renda?

Perguntas de palavra-chave e linguagem natural – Perguntas que incluem conteúdo conversacional complexo em que o significado pode não estar claro. Por exemplo, discurso principal. Quando o Amazon Kendra encontra uma palavra como “endereço”, que tem vários significados contextuais, ela infere corretamente o significado por trás da consulta de pesquisa e retorna informações relevantes.

Benefícios do Amazon Kendra

O Amazon Kendra é altamente escalável, consegue atender às demandas de desempenho, está totalmente integrado a outros serviços da AWS, como Amazon S3 e Amazon Lex, e oferece segurança de nível corporativo. Alguns dos benefícios de usar o Amazon Kendra incluem:

Simplicidade – O Amazon Kendra fornece um console e uma API para gerenciar os documentos que você deseja pesquisar. Você poderá usar uma API de pesquisa simples para integrar o Amazon Kendra aos aplicativos do seu cliente, como sites ou aplicativos móveis.

Conectividade – O Amazon Kendra pode se conectar a repositórios de dados ou fontes de dados de terceiros, como o Microsoft SharePoint. Você poderá facilmente indexar e pesquisar seus documentos usando sua fonte de dados.

Precisão – Diferentemente dos serviços de pesquisa tradicionais que usam pesquisas por palavra-chave, o Amazon Kendra tentam entender o contexto da pergunta e retorna a palavra, o trecho ou o documento mais relevante para sua consulta. o Amazon Kendra usa machine learning para melhorar os resultados da pesquisa ao longo do tempo.

Segurança — O Amazon Kendra oferece uma experiência de pesquisa corporativa altamente segura. Os resultados da pesquisa refletem o modelo de segurança da sua organização e podem ser filtrados com base no acesso do usuário ou grupo aos documentos. Os clientes são responsáveis por autenticar e autorizar o acesso do usuário.

Amazon Kendra Edições

O Amazon Kendra tem duas versões: Edição para desenvolvedores e Edição Enterprise. A tabela a seguir aprofunda-se nos atributos e nas diferenças entre as duas.

Amazon Kendra Edição para desenvolvedores	Amazon Kendra Edição Enterprise
A Edição para desenvolvedores do Amazon Kendra oferece todos os recursos do Amazon Kendra a um custo menor.	A Edição Enterprise do Amazon Kendra fornece todos os recursos do Amazon Kendra e foi projetada para contextos de produção.

Amazon Kendra Edição para desenvolvedores	Amazon Kendra Edição Enterprise
Caso de uso ideal • Explorando como o Amazon Kendra indexa seus documentos • Experimentar os recursos • Desenvolver os aplicativos que usam o Amazon Kendra	Caso de uso ideal • Indexando toda a sua biblioteca de documentos corporativos • Implantando seu aplicativo em um ambiente de produção
Recursos • Um nível gratuito com 750 horas de uso incluídas • Até 5 índices com até 5 fontes de dados cada • 10.000 documentos ou 3 GB de texto extraído • Aproximadamente 4.000 consultas por dia ou 0,05 consultas por segundo • Funciona em 1 zona de disponibilidade (AZ) – consulte Zonas de disponibilidade (data centers em regiões da AWS)	Recursos • Até 5 índices com até 50 fontes de dados cada • 100.000 documentos ou 30 GB de texto extraído • Aproximadamente 8.000 consultas por dia ou 0,1 consulta por segundo • Funciona em 3 zonas de disponibilidade (AZ) — consulte Zonas de disponibilidade (data centers em regiões da AWS) **Note** Aumente essa cota usando o console do Service Quotas.
Limitações • Não é para aplicações de produção • Sem garantias de latência ou disponibilidade	Limitações • Nenhum

Note

Para obter uma lista de regiões, endpoints e cotas de serviço com suporte pelo Amazon Kendra, consulte endpoints e cotas.

Definição de preços do Amazon Kendra

Você poderá começar gratuitamente com a Edição para desenvolvedores do Amazon Kendra, que oferece até 750 horas de uso nos primeiros 30 dias.

Depois que seu teste expirar, você será cobrado por todos os índices do Amazon Kendra provisionados, mesmo que estejam vazios e nenhuma consulta seja executada. Depois que o teste expirar, haverá cobranças adicionais pela verificação e sincronização de documentos usando as fontes de dados do Amazon Kendra.

Para obter uma lista completa de cobranças e preços , consulte Definição de preço do Amazon Kendra.

Você é um usuário iniciante do Amazon Kendra?

Se você estiver usando o Amazon Kendra pela primeira vez, recomendamos que leia as seções a seguir nesta ordem:

1	2	3	4	5	6
Como funciona o Amazon Kendra	Conceitos básicos	Criar um índice	Como adicionar documentos diretamente a um índice com o upload em lote.	Criar um conector de fonte de dados	Pesquisando um índice
Apresenta os componentes do Amazon Kendra e descreve como você os usa para criar uma solução de pesquisa.	Explica como configurar sua conta e testar a API de pesquisa do Amazon Kendra.	Explica como usar o Amazon Kendra para criar um índice de pesquisa e adicionar fontes de dados para	Explica como adicionar documentos diretamente a um índice do Amazon Kendra.	Explica como adicionar documentos do seu repositório de dados a um índice do Amazon Kendra.	Explica como usar a API de pesquisa do Amazon Kendra para pesquisar um índice.

1	2	3	4	5	6
Como funciona o Amazon Kendra	Conceitos básicos	Criar um índice	Como adicionar documentos diretamente a um índice com o upload em lote.	Criar um conector de fonte de dados	Pesquisando um índice
		sincronizar seus documentos.			

Como funciona o Amazon Kendra

Amazon Kendra fornece funcionalidade de pesquisa para seu aplicativo. Ele indexa seus documentos diretamente ou do repositório de documentos de terceiros e fornece informações relevantes de forma inteligente para os usuários. Você pode usar Amazon Kendra para criar um índice atualizável de documentos de vários tipos. Para obter uma lista dos tipos de documentos suportados pelo, Amazon Kendra consulte Tipos de documentos.

Amazon Kendra se integra a outros serviços. Por exemplo, você pode potencializar os bots de Amazon Lex bate-papo com a Amazon Kendra pesquisa para fornecer respostas úteis às perguntas dos usuários. Você pode usar um Amazon Simple Storage Service bucket como fonte de dados Amazon Kendra para se conectar e indexar seus documentos. E você pode configurar políticas de acesso ou permissões aos recursos usando AWS Identity and Access Management.

Amazon Kendra tem os seguintes componentes:

- Um índice que contém os documentos e os torna pesquisáveis.
- Uma fonte de dados que armazena os documentos e Amazon Kendra se conecta a. Você pode sincronizar automaticamente uma fonte de dados com um Amazon Kendra índice para que seu índice permaneça atualizado com o repositório de origem.
- Uma API de adição de documentos que adiciona documentos diretamente a um índice.

Você pode usar Amazon Kendra por meio do console ou da API. É possível criar, editar e excluir índices. A exclusão de um índice exclui todos os conectores da fonte de dados e exclui permanentemente todas as informações do documento. Amazon Kendra

Tópicos

- Índice
- Documentos
- Fontes de dados
- Consultas
- Tags

Índice

Um índice contém o conteúdo dos documentos e é estruturado de forma a tornar os documentos pesquisáveis. A forma como você adiciona documentos ao índice depende de como você armazena os documentos.

- Se você armazenar seus documentos em algum tipo de repositório, como um Amazon S3 bucket ou um SharePoint site da Microsoft, você usa um conector de fonte de dados para indexar seus documentos do seu repositório.
- Se você não armazena seus documentos em um repositório, você usa a BatchPutDocumentAPI para indexar diretamente seus documentos.
- Para perguntas e respostas das perguntas frequentes, que devem ser armazenadas em um bucket do Amazon Kendra (Amazon S3), carregue elas do bucket

Você pode criar índices com o Amazon Kendra console AWS CLI, o ou um AWS SDK. Para obter informações sobre os tipos de documentos que podem ser indexados, consulte Tipos de documentos.

Usando campos de documentos Amazon Kendra reservados ou comuns

Com a UpdateIndex API, você pode criar campos reservados ou comuns usando `DocumentMetadataConfigurationUpdates` e especificando o nome do campo de índice Amazon Kendra reservado para mapear para seu atributo de documento/nome de campo equivalente. Você também pode criar campos personalizados. Se você usa um conector de fonte de dados, a maioria inclui mapeamentos de campo que mapeiam os campos do documento da fonte de dados para campos de Amazon Kendra índice. Se usar o console, atualize os campos selecionando a fonte de dados, a ação de edição e, em seguida, prosseguindo para a seção de mapeamentos de campo para configurar a fonte de dados.

Você pode configurar o objeto `Search` para definir um campo como exibível, facetável, pesquisável e classificável. Configure o objeto `Relevance` para definir a ordem de classificação, a duração do aumento ou o período de tempo de um campo a ser aplicado ao aumento, à atualização, ao valor de importância e aos valores de importância mapeados para valores de campo específicos. Se usar o console, defina as configurações de pesquisa de um campo selecionando a opção de faceta no menu de navegação. Para definir o ajuste de relevância, selecione a opção de pesquisar o índice no menu de navegação, insira uma consulta e use as opções do painel lateral para ajustar a relevância da pesquisa. Você não pode alterar o tipo de campo depois de criar o campo.

Amazon Kendra tem os seguintes campos de documento reservados ou comuns que você pode usar:

- `_authors`: uma lista de um ou mais autores responsáveis pelo conteúdo do documento.
- `_category`: uma categoria que coloca um documento em um grupo específico.
- `_created_at`: a data e a hora no formato ISO 8601 em que o documento foi criado. Por exemplo, 2012-03-25T12:30:10+01:00 é o formato de data e hora ISO 8601 para 25 de março de 2012 às 12h30 (mais 10 segundos) no horário da Europa Central.
- `_data_source_id`: o identificador da fonte de dados que contém o documento.
- `_document_body`: o conteúdo do documento de trabalho.
- `_document_id`: o identificador exclusivo de cada documento.
- `_document_title`: o título do documento.
- `_excerpt_page_number`: o número da página em um arquivo PDF em que o trecho do documento aparece. Se o índice foi criado antes de 8 de setembro de 2020, você deve reindexar os documentos antes de poder usar esse atributo.
- `_faq_id`: se for um documento do tipo pergunta e resposta (Perguntas frequentes), um identificador exclusivo para as Perguntas frequentes.
- `_file_type`: o tipo de arquivo do documento, como pdf ou doc.
- `_last_updated_at`: a data e a hora no formato ISO 8601 em que o documento foi atualizado pela última vez. Por exemplo, 2012-03-25T12:30:10+01:00 é o formato de data e hora ISO 8601 para 25 de março de 2012 às 12h30 (mais 10 segundos) no horário da Europa Central.
- `_source_uri`: o URI em que o documento está disponível. Por exemplo, o URI do documento no site da empresa.
- `_version`: um identificador para a versão específica de um documento.
- `_view_count`: o número de vezes que o documento foi visualizado.
- `_language_code`(String): o código de um idioma que se aplica ao documento. O padrão é inglês se você não especificar um idioma. Para obter mais informações sobre os idiomas suportados, incluindo os códigos, consulte Adicionar documentos em outros idiomas além do inglês.

Para campos personalizados, você cria esses campos usando `DocumentMetadataConfigurationUpdates` com a API `UpdateIndex`, assim como faz ao criar um campo reservado ou comum. Você deve definir o tipo de dados apropriado para o campo personalizado. Se usar o console, atualize os campos selecionando a fonte de dados, a ação de edição e, em seguida, prosseguindo para a seção de mapeamentos de campo para configurar a

fonte de dados. Algumas fontes de dados não oferecem suporte à adição de novos campos ou campos personalizados. Você não pode alterar o tipo de campo depois de criar o campo.

Estes são os tipos que podem ser definidos em campos personalizados:

- Data
- Número
- String
- Lista de strings

Se você adicionou documentos ao índice usando a BatchPutDocumentAPI, `Attributes` lista os campos/atributos dos seus documentos e cria campos usando o objeto. `DocumentAttribute`

Para documentos indexados de uma fonte de Amazon S3 dados, você cria campos usando um arquivo de metadados JSON que inclui as informações dos campos.

Ao usar um banco de dados compatível como fonte de dados, poderá configurae as campos usando a opção de mapeamentos de campo.

Pesquisando índices

Depois de criar um índice, comece a pesquisar os documentos. Para obter mais informações, consulte Pesquisar índices.

Documentos

Esta seção explica como Amazon Kendra indexa os diversos formatos de documentos suportados e os diferentes campos/atributos dos documentos.

Tópicos

- Tipos ou formatos de documentos
- Atributos ou campos do documento

Tipos ou formatos de documentos

Amazon Kendra oferece suporte a tipos ou formatos de documentos populares, como PDF, HTML PowerPoint, Word e muito mais. Um índice pode conter vários formatos de documento.

Amazon Kendra extrai o conteúdo dentro dos documentos para tornar os documentos pesquisáveis. Os documentos são analisados de forma a otimizar a pesquisa no texto extraído e em qualquer conteúdo tabular (tabelas HTML) dentro dos documentos. Isso significa estruturar os documentos em campos ou atributos que são usados para pesquisa. Os metadados do documento, como a data da última modificação, podem ser campos úteis para pesquisa.

Os documentos podem ser organizados em linhas e colunas. Por exemplo, cada documento é uma linha e cada campo/atributo do documento, como o título e o conteúdo do corpo, é uma coluna. Por exemplo, se você usa um banco de dados como fonte de dados, os dados devem ser estruturados ou organizados em linhas e colunas.

Você pode adicionar documentos ao índice das seguintes formas:

- API BatchPutDocument
- Conector da fonte de dados

Se quiser adicionar um arquivo de perguntas frequentes, use a CreateFaqAPI para adicionar o arquivo armazenado em um Amazon S3 bucket. É possível escolher entre um formato CSV básico, um formato CSV que inclua atributos personalizados em um cabeçalho e um formato JSON que inclua campos personalizados. O formato padrão é CSV básico.

Veja a seguir informações sobre cada formato de documento compatível e como Amazon Kendra trata cada formato ao indexar documentos.

Formato do documento	Tratado como	Como o documento é tratado	Estrutura original
Formato de documento portátil (PDF)	HTML	Convertido em HTML e, em seguida, o conteúdo é extraído.	Não estruturado
HyperText Linguagem de marcação (HTML)	HTML	As tags HTML são filtradas para extrair conteúdo. O conteúdo deve estar entre as tags principais de HTML início e fim	Semiestruturado

Formato do documento	Tratado como	Como o documento é tratado	Estrutura original
		(<HTML>content</HTML>).	
Linguagem de marcação extensível (XML)	XML	As tags HTML são filtradas para extrair conteúdo.	Semiestruturado
Transformação de linguagem de folha de estilo extensível (XSLT)	XSLT	As tags HTML são filtradas para extrair conteúdo.	Semiestruturado
MarkDown (Maryland)	Texto sem formatação	O conteúdo é extraído com a MarkDown sintaxe incluída.	Semiestruturado
CSV (valores separados por vírgula)	CSV	Conteúdo extraído de cada célula, com um único arquivo tratado como um único resultado de documento.	Estruturado para arquivos de perguntas frequentes, caso contrário, semiestruturado
Microsoft Excel (XLS e XLSX)	XLS e XLSX	Conteúdo extraído de cada célula, com um único arquivo tratado como um único resultado de documento.	Semiestruturado
JavaScript Notação de objeto (JSON)	Texto sem formatação	O conteúdo é extraído com a sintaxe MarkDown incluída.	Semiestruturado

Formato do documento	Tratado como	Como o documento é tratado	Estrutura original
Formato Rich Text (RTF)	RTF	A sintaxe RTF é filtrada para extrair conteúdo.	Semiestruturado
Microsoft PowerPoint (PPT)	ppt	Somente o conteúdo de texto é extraído dos PowerPoint slides para pesquisa. Imagens e outros conteúdos não são extraídos.	Não estruturado
Microsoft Word (DOCX)	DOCX	Somente o conteúdo de texto é extraído das páginas do Word para pesquisa. Imagens e outros conteúdos não são extraídos.	Não estruturado
Texto sem formatação (TXT)	TXT	Todo o texto no documento de texto é extraído.	Não estruturado

Atributos ou campos do documento

Um documento tem atributos ou campos associados a ele. Os campos de um documento são as propriedades de um documento ou o que está contido na estrutura de um documento. Por exemplo, cada um dos documentos pode conter título, corpo do texto e autor. Você também pode adicionar campos personalizados para documentos específicos. Por exemplo, se o índice pesquisar documentos fiscais, você poderá especificar um campo personalizado para o tipo de documento fiscal, como W-2, 1099 e assim por diante.

Antes de usar um campo de documento em uma consulta, ele deve ser mapeado para um campo de índice. Por exemplo, o campo do título pode ser mapeado para o campo `_document_title`. Para obter mais informações, consulte Mapear campos. Para adicionar um novo campo, você deve criar um campo de índice para o qual mapear o campo. Você cria campos de índice usando o console ou usando a UpdateIndexAPI.

Você pode usar os campos do documento para filtrar respostas e criar resultados de pesquisa facetados. Por exemplo, você pode filtrar uma resposta para retornar somente uma versão específica de um documento ou filtrar pesquisas para retornar somente documentos fiscais do tipo 1099 que correspondam ao termo de pesquisa. Para obter mais informações, consulte Filtrar e pesquisar por facetas.

Você também pode usar os campos do documento para ajustar manualmente a resposta da consulta. Por exemplo, você pode optar por aumentar a importância do campo do título para aumentar o peso Amazon Kendra atribuído ao campo ao determinar quais documentos devem ser retornados na resposta. Para obter mais informações, consulte Ajustar a relevância da pesquisa.

Se você estiver adicionando um documento diretamente a um índice, especifique os campos no parâmetro de entrada do documento para a BatchPutDocumentAPI. Você especifica os valores do campo personalizado em uma matriz de DocumentAttributeobjetos. Se você estiver usando uma fonte de dados, o método usado para adicionar os campos do documento dependerá da fonte de dados. Para obter mais informações, consulte Mapear campos de fonte de dados.

Usando campos de documentos Amazon Kendra reservados ou comuns

Com a UpdateIndex API, você pode criar campos reservados ou comuns usando `DocumentMetadataConfigurationUpdates` e especificando o nome do campo de índice Amazon Kendra reservado para mapear para seu atributo de documento/nome de campo equivalente. Você também pode criar campos personalizados. Se você usa um conector de fonte de dados, a maioria inclui mapeamentos de campo que mapeiam os campos do documento da fonte de dados para campos de Amazon Kendra índice. Se usar o console, atualize os campos selecionando a fonte de dados, a ação de edição e, em seguida, prosseguindo para a seção de mapeamentos de campo para configurar a fonte de dados.

Você pode configurar o objeto `Search` para definir um campo como exibível, facetável, pesquisável e classificável. Configure o objeto `Relevance` para definir a ordem de classificação, a duração do aumento ou o período de tempo de um campo a ser aplicado ao aumento, à atualização, ao valor de importância e aos valores de importância mapeados para valores de campo específicos. Se usar

o console, defina as configurações de pesquisa de um campo selecionando a opção de faceta no menu de navegação. Para definir o ajuste de relevância, selecione a opção de pesquisar o índice no menu de navegação, insira uma consulta e use as opções do painel lateral para ajustar a relevância da pesquisa. Você não pode alterar o tipo de campo depois de criar o campo.

Amazon Kendra tem os seguintes campos de documento reservados ou comuns que você pode usar:

- `_authors`: uma lista de um ou mais autores responsáveis pelo conteúdo do documento.
- `_category`: uma categoria que coloca um documento em um grupo específico.
- `_created_at`: a data e a hora no formato ISO 8601 em que o documento foi criado. Por exemplo, 2012-03-25T12:30:10+01:00 é o formato de data e hora ISO 8601 para 25 de março de 2012 às 12h30 (mais 10 segundos) no horário da Europa Central.
- `_data_source_id`: o identificador da fonte de dados que contém o documento.
- `_document_body`: o conteúdo do documento de trabalho.
- `_document_id`: o identificador exclusivo de cada documento.
- `_document_title`: o título do documento.
- `_excerpt_page_number`: o número da página em um arquivo PDF em que o trecho do documento aparece. Se o índice foi criado antes de 8 de setembro de 2020, você deve reindexar os documentos antes de poder usar esse atributo.
- `_faq_id`: se for um documento do tipo pergunta e resposta (Perguntas frequentes), um identificador exclusivo para as Perguntas frequentes.
- `_file_type`: o tipo de arquivo do documento, como pdf ou doc.
- `_last_updated_at`: a data e a hora no formato ISO 8601 em que o documento foi atualizado pela última vez. Por exemplo, 2012-03-25T12:30:10+01:00 é o formato de data e hora ISO 8601 para 25 de março de 2012 às 12h30 (mais 10 segundos) no horário da Europa Central.
- `_source_uri`: o URI em que o documento está disponível. Por exemplo, o URI do documento no site da empresa.
- `_version`: um identificador para a versão específica de um documento.
- `_view_count`: o número de vezes que o documento foi visualizado.
- `_language_code`(String): o código de um idioma que se aplica ao documento. O padrão é inglês se você não especificar um idioma. Para obter mais informações sobre os idiomas suportados, incluindo os códigos, consulte Adicionar documentos em outros idiomas além do inglês.

Para campos personalizados, você cria esses campos usando `DocumentMetadataConfigurationUpdates` com a API `UpdateIndex`, assim como faz ao criar um campo reservado ou comum. Você deve definir o tipo de dados apropriado para o campo personalizado. Se usar o console, atualize os campos selecionando a fonte de dados, a ação de edição e, em seguida, prosseguindo para a seção de mapeamentos de campo para configurar a fonte de dados. Algumas fontes de dados não oferecem suporte à adição de novos campos ou campos personalizados. Você não pode alterar o tipo de campo depois de criar o campo.

Estes são os tipos que podem ser definidos em campos personalizados:

- Data
- Número
- String
- Lista de strings

Se você adicionou documentos ao índice usando a BatchPutDocumentAPI, `Attributes` lista os campos/atributos dos seus documentos e cria campos usando o objeto. `DocumentAttribute`

Para documentos indexados de uma fonte de Amazon S3 dados, você cria campos usando um arquivo de metadados JSON que inclui as informações dos campos.

Ao usar um banco de dados compatível como fonte de dados, poderá configurae as campos usando a opção de mapeamentos de campo.

Fontes de dados

Uma fonte de dados é um repositório de dados ou um local que Amazon Kendra se conecta e indexa seus documentos ou conteúdo. Por exemplo, você pode configurar Amazon Kendra para se conectar à Microsoft SharePoint para rastrear e indexar seus documentos armazenados nessa fonte. Você também pode indexar páginas da Web fornecendo os URLs Amazon Kendra para rastreamento. Você pode sincronizar automaticamente uma fonte de dados com um Amazon Kendra índice para que documentos adicionados, atualizados ou excluídos na fonte de dados também sejam adicionados, atualizados ou excluídos do índice.

As fontes de dados compatíveis são:

- Adobe Experience Manager

- Alfresco
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (Windows)
- Amazon FSx (EM UM NetApp TOQUE)
- Fontes de dados do banco de dados
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oráculo)
- Amazon RDS (PostgreSQL)
- Amazon S3 baldes
- Amazon Kendra Rastreador da Web
- Amazon WorkDocs
- Box
- Confluence
- Fontes de dados personalizadas
- Dropbox
- Drupal
- GitHub
- Gmail
- Google Workspace Drives
- 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 obter uma lista dos tipos ou formatos de documentos suportados pelo, Amazon Kendra consulte Tipos de documentos. Você deve primeiro criar um índice antes de criar um conector de fonte de dados para indexar os documentos a partir da sua fonte de dados.

> **Note**
>
> Para criar um índice de documentos, você não precisa usar uma fonte de dados. Você pode adicionar documentos diretamente a um índice com o upload em lote. Para obter mais informações, consulte Adicionar documentos diretamente a um índice.

Para ver um passo a passo sobre como usar o Amazon Kendra console, a AWS CLI ou os SDKs, consulte Introdução.

Consultas

Para obter respostas, os usuários consultam um índice. Os usuários podem usar linguagem natural em suas consultas. A resposta contém informações, como o título, um trecho do texto e a localização dos documentos no índice que fornecem a melhor resposta.

Amazon Kendra usa todas as informações que você fornece sobre seus documentos, não apenas o conteúdo dos documentos, para determinar se um documento é relevante para a consulta. Por exemplo, se seu índice contém informações sobre quando os documentos foram atualizados pela última vez, você pode pedir Amazon Kendra para atribuir uma relevância maior aos documentos que foram atualizados mais recentemente.

Uma consulta também pode conter critérios sobre como filtrar a resposta para que Amazon Kendra retorne somente documentos que satisfaçam os critérios do filtro. Por exemplo, ao criar um campo de índice chamado departamento, você pode filtrar a resposta para que somente documentos com o campo departamento definido como legal sejam retornados. Para obter mais informações, consulte Filtrar pesquisa.

Você pode influenciar os resultados de uma consulta ajustando a relevância de campos individuais no índice. O ajuste muda a importância de um campo nos resultados. Por exemplo, se você aumentar a importância de documentos com a categoria nova, é mais provável que documentos com essa categoria sejam incluídos na resposta. Para obter mais informações, consulte Ajustar a relevância da pesquisa e Ajustar as respostas.

Para obter mais informações sobre o uso de consultas, consulte Pesquisar em um índice.

Tags

Gerencie índices, fontes de dados e perguntas frequentes atribuindo tags ou rótulos. Você pode usar tags para categorizar seus Amazon Kendra recursos de várias maneiras. Por exemplo, por finalidade, por proprietário, por aplicativo ou por qualquer combinação. Cada tag consiste em uma chave e um valor, ambos definidos por você.

As tags ajudam a:

- Identifique e organize seus AWS recursos. Muitos AWS serviços oferecem suporte à marcação, então você pode atribuir a mesma tag a recursos em serviços diferentes para indicar que os recursos estão relacionados. Por exemplo, você pode marcar um índice e o Amazon Lex bot que usa o índice com a mesma tag.
- Alocar custos. Você ativa as tags no AWS Billing and Cost Management painel. AWS usa tags para categorizar seus custos e entregar um relatório mensal de alocação de custos para você. Para obter mais informações, consulte Alocação e marcação de custos em Sobre o AWS Billing and Cost Management.
- Controle o acesso aos seus atributos. Você pode usar tags em políticas do AWS Identity and Access Management (IAM) para controlar o acesso aos recursos do Amazon Kendra . Você pode anexar essas políticas a uma IAM função ou usuário para ativar o controle de acesso baseado em tags. Para obter mais informações, consulte Autorização baseada em tags.

Você pode criar e gerenciar tags usando o AWS Management Console, o AWS Command Line Interface (AWS CLI) ou a Amazon Kendra API.

Marcar recursos

Se você estiver usando o Amazon Kendra console, poderá marcar recursos ao criá-los ou adicioná-los posteriormente. Também é possível usar o console para atualizar ou remover tags.

Se você estiver usando o AWS Command Line Interface (AWS CLI) ou a Amazon Kendra API, use as seguintes operações para gerenciar tags para seus recursos:

- CreateDataSource—Aplique tags ao criar uma fonte de dados.
- CreateFaq—Aplique tags ao criar uma FAQ.
- CreateIndex—Aplique tags ao criar um índice.
- ListTagsForResource—Visualize as tags associadas a um recurso.
- TagResource—Adicione e modifique tags para um recurso.
- UntagResource—Remover tags de um recurso.

Restrições de tags

As restrições a seguir se aplicam às tags nos Amazon Kendra recursos:

- Número máximo de tags: 50
- Tamanho máximo da chave: 128 caracteres
- Tamanho máximo do valor: 256 caracteres
- Caracteres válidos de chave e valor a-z, A-Z, 0-9, espaço e os seguintes caracteres: _ . : / = + - e @
- As chaves e os valores diferenciam letras maiúsculas de minúsculas
- Não use `aws:` como um prefixo para chaves, pois ele é reservado para uso da AWS

Configuração do Amazon Kendra

Antes de usar o Amazon Kendra, é necessário ter uma conta da Amazon Web Services (AWS). Depois de ter uma AWS conta, você pode acessar o Amazon Kendra por meio do console Amazon Kendra, AWS Command Line Interface do () ou dos SDKs. AWS CLI AWS

Este guia inclui exemplos para AWS CLI Java e Python.

Tópicos

- Inscreva-se para AWS
- Regiões e endpoints
- Configurando o AWS CLI
- Configurando os AWS SDKs

Inscreva-se para AWS

Quando você se inscreve no Amazon Web Services (AWS), sua conta é automaticamente cadastrada em todos os serviços AWS, incluindo o Amazon Kendra. Você será cobrado apenas pelos serviços que usar.

Se você já tiver uma AWS conta, vá para a próxima tarefa. Se você ainda não possuir uma conta da AWS , use o procedimento a seguir para criar uma.

Para se inscrever em AWS

1. Abra https://aws.amazon.com e escolha Criar uma AWS conta.
2. Siga as instruções na tela para concluir a criação da conta. Anote o número de 12 dígitos da conta da AWS . Parte do procedimento de cadastro envolve uma chamada telefônica e a digitação de um PIN usando o teclado do telefone.
3. Crie um usuário administrador AWS Identity and Access Management (IAM). Consulte Criar seu primeiro grupo e usuário do IAM no Guia do usuário do AWS Identity and Access Management para obter instruções.

Regiões e endpoints

Um endpoint é um URL que é o ponto de entrada para um serviço da Web. Cada endpoint está associado a uma AWS região específica. Se você usa uma combinação do console Amazon Kendra, do e dos SDKs AWS CLI do Amazon Kendra, preste atenção às suas regiões padrão, pois todos os componentes do Amazon Kendra de uma determinada campanha (índice, consulta etc.) devem ser criados na mesma região. Para obter uma lista de todas as regiões e endpoints com suporte pelo o Amazon Kendra, consulte Regiões e endpoints.

Configurando o AWS CLI

A AWS Command Line Interface (AWS CLI) é uma ferramenta unificada para desenvolvedores para gerenciar AWS serviços, incluindo o Amazon Kendra. Recomendamos que você a instale.

1. Para instalar o AWS CLI, siga as instruções em Instalando a interface de linha de AWS comando no Guia do usuário da interface de linha de AWS comando.
2. Para configurar o AWS CLI e configurar um perfil para chamar o AWS CLI, siga as instruções em Configurando o no Guia do AWS CLI usuário da interface de linha de AWS comando.
3. Para confirmar se o AWS CLI perfil está configurado corretamente, execute o seguinte comando:

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

   Se o seu perfil foi configurado corretamente, você verá uma saída semelhante a esta:

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

4. Para verificar se o AWS CLI está configurado para uso com o Amazon Kendra, execute os seguintes comandos:

   ```
   aws kendra help
   ```

   Se AWS CLI estiver configurado corretamente, você verá uma lista dos AWS CLI comandos compatíveis com o Amazon Kendra, o tempo de execução do Amazon Kendra e os eventos do Amazon Kendra.

Configurando os AWS SDKs

Baixe e instale os AWS SDKs que você deseja usar. Este guia fornece exemplos para Python. Para obter informações sobre outros AWS SDKs, consulte Ferramentas para Amazon Web Services.

O pacote para o SDK do Python é chamado de Boto3.

Antes de executar os comandos Python abaixo, você deverá primeiro baixar e instalar o Python 3.6 ou posterior no sistema operacional. O suporte para Python 3.5 e versões anteriores está obsoleto. Se não tiver pip incluído no seu diretório de scripts do Python, você poderá baixar o get-pip.py e armazená-lo no seu diretório de scripts. Você também poderá definir seu diretório Python como um Caminho ou variável de ambiente usando um 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 o Boto3, você deve configurar as credenciais de autenticação para sua AWS conta usando o console do IAM.

IAM funções de acesso para Amazon Kendra

Quando você cria um índice, fonte de dados ou perguntas frequentes, Amazon Kendra precisa acessar os AWS recursos necessários para criar o Amazon Kendra recurso. Você deve criar uma política AWS Identity and Access Management (IAM) antes de criar o Amazon Kendra recurso. Ao chamar a operação, forneça o nome do recurso da Amazon (ARN) da função com a política em anexo. Por exemplo, se você estiver chamando a BatchPutDocumentAPI para adicionar documentos de um Amazon S3 bucket, você fornece Amazon Kendra uma função com uma política que tem acesso ao bucket.

Você pode criar uma nova IAM função no Amazon Kendra console ou escolher uma função IAM existente para usar. O console exibe funções que têm a string “kendra” ou “Kendra” no nome da função.

Os tópicos a seguir fornecem detalhes das políticas necessárias. Se você criar IAM funções usando o Amazon Kendra console, essas políticas serão criadas para você.

Tópicos

- IAM funções para índices
- IAM funções para a BatchPutDocument API
- IAM funções para fontes de dados
- Função de nuvem privada virtual (VPC) IAM
- IAM funções para perguntas frequentes (FAQs)
- IAM funções para sugestões de consulta
- IAM funções para mapeamento principal de usuários e grupos
- IAM funções para AWS IAM Identity Center
- IAM funções para Amazon Kendra experiências
- IAM funções para enriquecimento personalizado de documentos

IAM funções para índices

Ao criar um índice, você deve fornecer uma IAM função com permissão para gravar em um Amazon CloudWatch. Você também deve fornecer uma política de confiança que Amazon Kendra permita assumir a função. A seguir estão as políticas que devem ser fornecidas.

IAM funções para índices

Uma política de função para Amazon Kendra permitir o acesso a um 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:*"
        }
    ]
}
```

Uma política de funções para Amazon Kendra permitir o acesso AWS Secrets Manager. Se você estiver usando o contexto do usuário Secrets Manager como um local chave, poderá usar a política a seguir.

```
{
   "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"
               ]
            }
         }
      }
   ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para a BatchPutDocument API

> **Warning**
>
> Amazon Kendra não usa uma política de bucket que conceda permissões a um Amazon Kendra principal para interagir com um bucket do S3. Em vez disso, ele usa as funções do IAM . Certifique-se de que isso Amazon Kendra não esteja incluído como membro confiável em sua política de bucket para evitar problemas de segurança de dados ao conceder permissões acidentalmente a diretores arbitrários. No entanto, você pode adicionar uma política de bucket para usar um Amazon S3 bucket em contas diferentes. Para obter mais informações, consulte Políticas para uso do Amazon S3 em todas as contas. Para obter mais informações sobre as funções do IAM para fontes de dados do S3, consulte as funções do IAM.

Ao usar a BatchPutDocumentAPI para indexar documentos em um Amazon S3 bucket, você deve fornecer Amazon Kendra uma IAM função com acesso ao bucket. Você também deve fornecer uma política de confiança que Amazon Kendra permita assumir a função. Se os documentos no bucket estiverem criptografados, você deverá fornecer permissão para usar a chave mestra do AWS KMS cliente (CMK) para descriptografar os documentos.

IAM funções para a BatchPutDocument API

Uma política de função necessária para Amazon Kendra permitir o acesso a um Amazon S3 bucket.

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

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

É recomendável que você inclua `aws:sourceAccount` e `aws:sourceArn` na política de confiança. Isso limita as permissões e verifica com segurança se `aws:sourceAccount` `aws:sourceArn` elas são as mesmas fornecidas na política de IAM função da `sts:AssumeRole` ação. Isso impede que entidades não autorizadas acessem suas IAM funções e suas permissões. Para obter mais informações, consulte o AWS Identity and Access Management guia sobre o problema confuso do deputado.

```
{
    "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/
*"
                }
            }
```

```
        }
    ]
}
```

Uma política de função opcional Amazon Kendra para permitir o uso de uma chave mestra AWS KMS do cliente (CMK) para descriptografar documentos em um bucket. Amazon S3

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

IAM funções para fontes de dados

Ao usar a CreateDataSourceAPI, você deve atribuir Amazon Kendra uma IAM função que tenha permissão para acessar os recursos. As permissões específicas necessárias dependem da fonte de dados.

IAM funções para fontes de dados do Adobe Experience Manager

Ao usar o Adobe Experience Manager, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Adobe Experience Manager.
- Permissão para chamar as APIs públicas necessárias para o conector do Adobe Experience Manager.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Adobe Experience Manager Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados da Alfresco

Ao usar o Alfresco, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Alfresco.
- Permissão para chamar as APIs públicas necessárias para o conector do Alfresco.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados Alfresco Amazon Kendra por meio Amazon VPC de. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados Aurora (MySQL)

Ao usar Aurora (MySQL), você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Aurora (MySQL).
- Permissão para chamar as APIs públicas necessárias para o conector Aurora (MySQL).
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados Aurora (MySQL) por meio de Amazon Kendra . Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados Aurora (PostgreSQL)

Ao usar o Aurora (PostgreSQL), você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Aurora (PostgreSQL).
- Permissão para chamar as APIs públicas necessárias para o conector do Aurora (PostgreSQL).
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados Aurora (PostgreSQL) por meio de. Amazon Kendra Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes Amazon FSx de dados

Ao usar Amazon FSx, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu sistema de Amazon FSx arquivos.
- Permissão para acessar Amazon Virtual Private Cloud (VPC) onde seu sistema de Amazon FSx arquivos reside.
- Permissão para obter o nome de domínio do Active Directory para seu sistema de Amazon FSx arquivos.
- Permissão para chamar as APIs públicas necessárias para o conector do Amazon FSx .

- Permissão para chamar as APIs `BatchPutDocument` e `BatchDeleteDocument` para atualizar o 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}}"
        }
        ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados de banco de dados

Ao usar um banco de dados como fonte de dados, você Amazon Kendra fornece uma função que tem as permissões necessárias para se conectar ao. Isso inclui:

- Permissão para acessar o AWS Secrets Manager segredo que contém o nome de usuário e a senha do site. Para obter mais informações sobre os conteúdos da senha, consulte fontes de dados.
- Permissão para usar a chave mestra AWS KMS do cliente (CMK) para descriptografar o nome de usuário e a senha secreta armazenados pelo. Secrets Manager
- Permissão para usar as operações `BatchPutDocument` e `BatchDeleteDocument` para atualizar o índice.
- Permissão para acessar o Amazon S3 bucket que contém o certificado SSL usado para se comunicar com o site.

> **Note**
>
> Você pode conectar fontes de dados do banco de dados Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
    "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/*"
            ]
        }
    ]
}
```

Há duas políticas opcionais que você pode usar com uma fonte de dados.

Se você criptografou o Amazon S3 bucket que contém o certificado SSL usado para se comunicar com o, forneça uma política para dar Amazon Kendra acesso à chave.

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

```
        }
    ]
}
```

Se você estiver usando uma VPC, forneça uma política que dê Amazon Kendra acesso aos recursos necessários. Consulte as Funções do IAM para fontes de dados e VPC para ver a política necessária.

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados Amazon RDS (Microsoft SQL Server)

Ao usar um conector de fonte de dados Amazon RDS (Microsoft SQL Server), você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua instância de fonte de dados Amazon RDS (Microsoft SQL Server).
- Permissão para chamar as APIs públicas necessárias para o conector da fonte de dados Amazon RDS (Microsoft SQL Server).
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados Amazon RDS (Microsoft SQL Server) Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados Amazon RDS (MySQL)

Ao usar um conector de fonte de dados Amazon RDS (MySQL), você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua instância de fonte de dados Amazon RDS (MySQL).
- Permissão para chamar as APIs públicas necessárias para o conector da Amazon RDS fonte de dados (MySQL).
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados Amazon RDS (MySQL) por meio de Amazon Kendra . Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados Amazon RDS (Oracle)

Ao usar um conector de fonte de dados Amazon RDS Oracle, você fornece uma função com as seguintes políticas.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua instância de fonte de dados Amazon RDS (Oracle).
- Permissão para chamar as APIs públicas necessárias para o conector da fonte de dados Amazon RDS (Oracle).
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados Amazon RDS Oracle Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados Amazon RDS (PostgreSQL)

Ao usar um conector de fonte de dados Amazon RDS (PostgreSQL), você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua instância de fonte de dados Amazon RDS (PostgreSQL).
- Permissão para chamar as APIs públicas necessárias para o conector da fonte de dados do Amazon RDS (PostgreSQL).
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados Amazon RDS (PostgreSQL) por meio de. Amazon Kendra Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes Amazon S3 de dados

> **Warning**
>
> Amazon Kendra não usa uma política de bucket que conceda permissões a um Amazon Kendra principal para interagir com um bucket do S3. Em vez disso, ele usa IAM funções. Certifique-se de que isso Amazon Kendra não esteja incluído como membro confiável em sua política de bucket para evitar problemas de segurança de dados ao conceder permissões acidentalmente a diretores arbitrários. No entanto, você pode adicionar uma política de bucket para usar um bucket do Amazon S3 em contas diferentes. Para obter mais

informação, consulte Políticas para usar Amazon S3 em todas as contas (role a tela para baixo).

Ao usar um Amazon S3 bucket como fonte de dados, você fornece uma função que tem permissão para acessar o bucket e usar as `BatchPutDocument` `BatchDeleteDocument` operações e. Se os documentos no Amazon S3 bucket estiverem criptografados, você deverá fornecer permissão para usar a chave mestra do AWS KMS cliente (CMK) para descriptografar os documentos.

As políticas de função a seguir devem Amazon Kendra permitir assumir uma função. Role mais para baixo para ver uma política de confiança para assumir uma função.

Uma política de função necessária para Amazon Kendra permitir o uso Amazon S3 de um bucket como fonte de dados.

```
{
    "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"
            ]
        }
    ]
}
```

Uma política de função opcional Amazon Kendra para permitir o uso de uma chave mestra AWS KMS do cliente (CMK) para descriptografar documentos em um bucket. Amazon S3

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

Uma política de função opcional Amazon Kendra para permitir o acesso a um Amazon S3 bucket Amazon VPC, usando um e sem ativar AWS KMS ou compartilhar AWS KMS permissões.

```
{
  "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}}"
    }
  ]
}
```

Uma política de função opcional Amazon Kendra para permitir o acesso a um Amazon S3 bucket enquanto usa um Amazon VPC e com AWS KMS as permissões ativadas.

```
{
  "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}}"
    }
  ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

Políticas para usar Amazon S3 em todas as contas

Se seu Amazon S3 bucket estiver em uma conta diferente da conta que você usa para seu Amazon Kendra índice, você pode criar políticas para usá-lo em várias contas.

Uma política de função para usar seu Amazon S3 bucket como fonte de dados quando o bucket está em uma conta diferente do seu Amazon Kendra índice. `s3:PutObject` e `s3:PutObjectAcl` são

opcionais. Você pode usá-los se quiser incluir um [arquivo de configuração para sua lista de controle de acesso](#).

```
{
    "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/*"
        }
```

```
    ]
}
```

Uma política de bucket para permitir que a função da fonte de Amazon S3 dados acesse o Amazon S3 bucket em todas as contas. `s3:PutObject` e `s3:PutObjectAcl` são opcionais. Você pode usá-los se quiser incluir um arquivo de configuração para sua lista de controle de acesso.

```
{
    "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"
        }
    ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

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

IAM funções para fontes de dados do Amazon Kendra Web Crawler

Ao usar o Amazon Kendra Web Crawler, você fornece uma função com as seguintes políticas:

- Permissão para acessar o AWS Secrets Manager segredo que contém as credenciais para se conectar a sites ou a um servidor proxy da web apoiado pela autenticação básica. Para obter mais informações sobre os conteúdos da senha, consulte Usando a fonte de dados do Web Crawler.
- Permissão para usar a chave mestra AWS KMS do cliente (CMK) para descriptografar o nome de usuário e a senha secreta armazenados pelo. Secrets Manager
- Permissão para usar as operações `BatchPutDocument` e `BatchDeleteDocument` para atualizar o índice.
- Se você usa um Amazon S3 bucket para armazenar sua lista de URLs iniciais ou sitemaps, inclua permissão para acessar o bucket. Amazon S3

> **Note**
>
> Você pode conectar uma fonte de dados do Amazon Kendra Web Crawler por meio de Amazon Kendra . Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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"
  }]
}
```

Se você armazenar seus URLs iniciais ou sitemaps em um Amazon S3 bucket, deverá adicionar essa permissão à função.

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

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes Amazon WorkDocs de dados

Ao usar Amazon WorkDocs, você fornece uma função com as seguintes políticas

- Permissão para verificar a ID do diretório (ID da organização) que corresponde ao repositório do site do Amazon WorkDocs .
- Permissão para obter o nome de domínio do Active Directory que contém o diretório do site do Amazon WorkDocs .
- Permissão para chamar as APIs públicas necessárias para o conector do Amazon WorkDocs .
- Permissão para chamar as APIs `BatchPutDocument` e `BatchDeleteDocument` para atualizar o 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"
      ]
    }
  ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

```
{
```

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

IAM funções para fontes de dados do Box

Ao usar o Box, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Slack.
- Permissão para chamar as APIs públicas necessárias para o conector do Box.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Box Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

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

IAM funções para fontes de dados do Confluence

IAM funções para o Confluence Connector v1.0

Ao usar um conector de fonte de dados do servidor do Confluence, você fornece uma função com as políticas a seguir.

- Permissão para acessar o AWS Secrets Manager segredo que contém as credenciais necessárias para se conectar ao Confluence. Para obter mais informações sobre os conteúdos da senha, consulte Fontes de dados do Confluence.
- Permissão para usar a chave mestra AWS KMS do cliente (CMK) para descriptografar o nome de usuário e a senha secreta armazenados pelo. Secrets Manager
- Permissão para usar as operações `BatchPutDocument` e `BatchDeleteDocument` para atualizar o índice.

> **Note**
>
> Você pode conectar uma fonte de dados do Confluence por meio de Amazon Kendra . Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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"
  }]
}
```

Se você estiver usando uma VPC, forneça uma política que dê Amazon Kendra acesso aos recursos necessários. Consulte as Funções do IAM para fontes de dados e VPC para ver a política necessária.

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

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

IAM funções para o Confluence Connector v2.0

Ao usar um conector de fonte de dados do conector do Confluence v2.0, você fornece uma função com as políticas a seguir.

- Permissão para acessar o AWS Secrets Manager segredo que contém as credenciais de autenticação do Confluence. Para obter mais informações sobre os conteúdos da senha, consulte Fontes de dados do Confluence.
- Permissão para usar a chave mestra AWS KMS do cliente (CMK) para descriptografar o nome de usuário e a senha secreta armazenados pelo. AWS Secrets Manager
- Permissão para usar as operações `BatchPutDocument` e `BatchDeleteDocument` para atualizar o índice.

Você também deve anexar uma política de confiança que Amazon Kendra permita assumir a função.

> **Note**
>
> Você pode conectar uma fonte de dados do Confluence por meio de Amazon Kendra . Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

Uma política de função para permitir Amazon Kendra a conexão com o 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"
    }
  ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Dropbox

Ao usar o Dropbox, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Dropbox.
- Permissão para chamar as APIs públicas necessárias para o conector do Dropbox.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Dropbox Amazon Kendra por meio Amazon VPC de. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
"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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

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

IAM funções para fontes de dados do Drupal

Ao usar o Drupal, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Drupal.
- Permissão para chamar as APIs públicas necessárias para o conector do Drupal.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Drupal por meio de Amazon Kendra . Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

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

IAM funções para fontes GitHub de dados

Ao usar GitHub, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu GitHub.
- Permissão para chamar as APIs públicas necessárias para o GitHub conector.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte GitHub de dados Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

```
    ]
}
```

IAM funções para fontes de dados do Gmail

Ao usar o Gmail, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Gmail.
- Permissão para chamar as APIs públicas necessárias para o conector do Gmailconnector.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Gmail Amazon Kendra por meio Amazon VPC de. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
"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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Google Drive

Ao usar uma fonte de dados do Google Workspace Drive, você Amazon Kendra fornece uma função que tem as permissões necessárias para se conectar ao site. Isso inclui:

- Permissão para obter e decifrar o AWS Secrets Manager segredo que contém o e-mail da conta do cliente, o e-mail da conta do administrador e a chave privada necessários para se conectar ao site do Google Drive. Para obter mais informações sobre os conteúdos da senha, consulte fontes de dados do Google Drive.
- Permissão para usar as BatchDeleteDocumentAPIs BatchPutDocumente.

> **Note**
>
> Você pode conectar uma fonte de dados do Google Drive Amazon Kendra ao Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

A IAM política a seguir fornece as permissões necessárias:

```
{
  "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"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados IBM DB2

Ao usar um conector de fonte de dados do IBM DB2, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua instância de fonte de dados IBM DB2.
- Permissão para chamar as APIs públicas necessárias para o conector da fonte de dados do IBM DB2.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados IBM DB2 Amazon Kendra por meio Amazon VPC de. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Jira

Ao usar o Jira, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Jira.
- Permissão para chamar as APIs públicas necessárias para o conector do Jira.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Jira Amazon Kendra por meio Amazon VPC de. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Microsoft Exchange

Ao usar uma fonte de dados do Microsoft Exchange, você Amazon Kendra fornece uma função que tem as permissões necessárias para se conectar ao site. Isso inclui:

- Permissão para obter e descriptografar o AWS Secrets Manager segredo que contém a ID do aplicativo e a chave secreta necessárias para se conectar ao site do Microsoft Exchange. Para obter mais informações sobre os conteúdos da senha, consulte fontes de dados do Microsoft Exchange.
- Permissão para usar as BatchDeleteDocumentAPIs BatchPutDocumente.

> **Note**
>
> Você pode conectar uma fonte de dados do Microsoft Exchange Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

A IAM política a seguir fornece as permissões necessárias:

```
{
  "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"
  }]
}
```

Se você estiver armazenando a lista de usuários para indexar em um Amazon S3 bucket, também deverá fornecer permissão para usar a `GetObject` operação do S3. O modelo de política do IAM a seguir fornece as permissões necessárias.

```
{
  "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"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de OneDrive dados da Microsoft

Ao usar uma fonte de OneDrive dados da Microsoft, você Amazon Kendra fornece uma função que tem as permissões necessárias para se conectar ao site. Isso inclui:

- Permissão para obter e descriptografar o AWS Secrets Manager segredo que contém o ID do aplicativo e a chave secreta necessários para se conectar ao site. OneDrive Para obter mais informações sobre o conteúdo do segredo, consulte Fontes de OneDrive dados da Microsoft.
- Permissão para usar as BatchDeleteDocumentAPIs BatchPutDocumente.

> **Note**
>
> Você pode conectar uma fonte de OneDrive dados da Microsoft Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

A IAM política a seguir fornece as permissões necessárias:

```
{
  "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"
  }]
}
```

Se você estiver armazenando a lista de usuários para indexar em um Amazon S3 bucket, também deverá fornecer permissão para usar a `GetObject` operação do S3. O modelo de política do IAM a seguir fornece as permissões necessárias.

```
{
  "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"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de SharePoint dados da Microsoft

IAM funções para SharePoint Connector v1.0

Para uma fonte de dados Microsoft SharePoint Connector v1.0, você fornece uma função com as políticas a seguir.

- Permissão para acessar o AWS Secrets Manager segredo que contém o nome de usuário e a senha do SharePoint site. Para obter mais informações sobre o conteúdo do segredo, consulte Fontes de SharePoint dados da Microsoft.
- Permissão para usar a chave mestra AWS KMS do cliente (CMK) para descriptografar o nome de usuário e a senha secreta armazenados pelo. AWS Secrets Manager

- Permissão para usar as operações `BatchPutDocument` e `BatchDeleteDocument` para atualizar o índice.
- Permissão para acessar o Amazon S3 bucket que contém o certificado SSL usado para se comunicar com o SharePoint site.

Você também deve anexar uma política de confiança que Amazon Kendra permita assumir a função.

> **Note**
>
> Você pode conectar uma fonte de SharePoint dados da Microsoft Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
    "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/*"
            ]
        }
    ]
}
```

Se você criptografou o Amazon S3 bucket que contém o certificado SSL usado para se comunicar com o SharePoint site, forneça uma política para dar Amazon Kendra acesso à chave.

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

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para SharePoint Connector v2.0

Para uma fonte de dados Microsoft SharePoint Connector v2.0, você fornece uma função com as políticas a seguir.

- Permissão para acessar o AWS Secrets Manager segredo que contém as credenciais de autenticação do SharePoint site. Para obter mais informações sobre o conteúdo do segredo, consulte Fontes de SharePoint dados da Microsoft.
- Permissão para usar a chave mestra AWS KMS do cliente (CMK) para descriptografar o nome de usuário e a senha secreta armazenados pelo. AWS Secrets Manager
- Permissão para usar as operações `BatchPutDocument` e `BatchDeleteDocument` para atualizar o índice.
- Permissão para acessar o Amazon S3 bucket que contém o certificado SSL usado para se comunicar com o SharePoint site.

Você também deve anexar uma política de confiança que Amazon Kendra permita assumir a função.

> **Note**
>
> Você pode conectar uma fonte de SharePoint dados da Microsoft Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
```

```
  "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": "*"
    }
  ]
}
```

Se você criptografou o Amazon S3 bucket que contém o certificado SSL usado para se comunicar com o SharePoint site, forneça uma política para dar Amazon Kendra acesso à chave.

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

```
            "Resource": [
                "arn:aws:kms:your-region:youraccount-id:key/key-id"
            ]
        }
    ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Microsoft SQL Server

Ao usar o Microsoft SQL Server, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua instância do Microsoft SQL Server.
- Permissão para chamar as APIs públicas necessárias para o conector do banco de dados do Microsoft SQL Server.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Microsoft SQL Server Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Microsoft Teams

Ao usar uma fonte de dados do Microsoft Teams, você Amazon Kendra fornece uma função que tem as permissões necessárias para se conectar ao site. Isso inclui:

- Permissão para obter e descriptografar o AWS Secrets Manager segredo que contém o ID do cliente e o segredo do cliente necessários para se conectar ao Microsoft Teams. Para obter mais informações sobre os conteúdos da senha, consulte fontes de dados do Microsoft Teams.

> **Note**
>
> Você pode conectar uma fonte de dados do Microsoft Teams Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

A IAM política a seguir fornece as permissões necessárias:

```
{
  "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"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Microsoft Yammer

Ao usar uma fonte de dados do Microsoft Yammer, você Amazon Kendra fornece uma função que tem as permissões necessárias para se conectar ao site. Isso inclui:

- Permissão para obter e descriptografar o AWS Secrets Manager segredo que contém o ID do aplicativo e a chave secreta necessários para se conectar ao site do Microsoft Yammer. Para obter mais informações sobre os conteúdos da senha, consulte fontes de dados do Microsoft Yammer.
- Permissão para usar as BatchDeleteDocumentAPIs BatchPutDocumente.

> **Note**
>
> Você pode conectar uma fonte de dados do Microsoft Yammer Amazon Kendra por meio Amazon VPC de. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

A IAM política a seguir fornece as permissões necessárias:

```
{
  "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"
  }]
}
```

Se você estiver armazenando a lista de usuários para indexar em um Amazon S3 bucket, também deverá fornecer permissão para usar a `GetObject` operação do S3. O modelo de política do IAM a seguir fornece as permissões necessárias.

```
{
  "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"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados MySQL

Ao usar um conector de fonte de dados do MySQL, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua instância da fonte de dados My SQL.
- Permissão para chamar as APIs públicas necessárias para o conector da fonte de dados do MySQL.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados MySQL por meio de Amazon Kendra . Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados Oracle

Ao usar um conector de fonte de dados do Oracle, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua instância de fonte de dados Oracle.
- Permissão para chamar as APIs públicas necessárias para o conector da fonte de dados do Oracle.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados Oracle Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do PostgreSQL

Ao usar um conector de fonte de dados do PostgreSQL, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua instância da fonte de dados PostgreSQL.
- Permissão para chamar as APIs públicas necessárias para o conector da fonte de dados do PostgreSQL.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados PostgreSQL por meio de. Amazon Kendra Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Quip

Ao usar o , você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Quip.
- Permissão para chamar as APIs públicas necessárias para o conector do .
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Quip Amazon Kendra por meio Amazon VPC de. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Salesforce

Ao usar um conector de fonte de dados do servidor do Salesforce, você fornece uma função com as políticas a seguir.

- Permissão para acessar o AWS Secrets Manager segredo que contém o nome de usuário e a senha do site do Salesforce. Para obter mais informações sobre os conteúdos da senha, consulte fontes de dados do Salesforce.
- Permissão para usar a chave mestra AWS KMS do cliente (CMK) para descriptografar o nome de usuário e a senha secreta armazenados pelo. Secrets Manager
- Permissão para usar as operações `BatchPutDocument` e `BatchDeleteDocument` para atualizar o índice.

> **Note**
>
> Você pode conectar uma fonte de dados do Salesforce por meio de Amazon Kendra . Amazon VPC Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

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

IAM funções para fontes ServiceNow de dados

Ao usar a ServiceNow como fonte de dados, você fornece uma função com as seguintes políticas:

- Permissão para acessar o Secrets Manager segredo que contém o nome de usuário e a senha do ServiceNow site. Para obter mais informações sobre os conteúdos da senha, consulte ServiceNow fontes de dados.
- Permissão para usar a chave mestra AWS KMS do cliente (CMK) para descriptografar o nome de usuário e a senha secreta armazenados pelo. Secrets Manager
- Permissão para usar as operações `BatchPutDocument` e `BatchDeleteDocument` para atualizar o índice.

> **Note**
>
> Você pode conectar uma fonte ServiceNow de dados Amazon Kendra por meio de Amazon VPC. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Slack

Ao usar o Slack, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar seu Slack.
- Permissão para chamar as APIs públicas necessárias para o conector do Slack.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Slack Amazon Kendra por meio Amazon VPC de. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para fontes de dados do Zendesk

Ao usar o Zendesk, você fornece uma função com as políticas a seguir.

- Permissão para acessar seu AWS Secrets Manager segredo para autenticar sua Zendesk Suite.
- Permissão para chamar as APIs públicas necessárias para o conector do Zendesk.
- Permissão para chamar as APIs `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping` e `ListGroupsOlderThanOrderingId`.

> **Note**
>
> Você pode conectar uma fonte de dados do Zendesk Amazon Kendra por meio Amazon VPC de. Se você estiver usando um Amazon VPC, precisará adicionar permissões adicionais.

```
{
  "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}}"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

Função de nuvem privada virtual (VPC) IAM

Se você usa uma nuvem privada virtual (VPC) para se conectar à sua fonte de dados, deverá fornecer as seguintes permissões adicionais.

Função 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": "*"
    }
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para perguntas frequentes (FAQs)

Ao usar a CreateFaqAPI para carregar perguntas e respostas em um índice, você deve fornecer Amazon Kendra uma IAM função com acesso ao Amazon S3 bucket que contém os arquivos de origem. Se os arquivos de origem estiverem criptografados, você deverá fornecer permissão para usar a chave mestra AWS KMS do cliente (CMK) para descriptografar os arquivos.

IAM funções para perguntas frequentes

Uma política de função necessária para Amazon Kendra permitir o acesso a um Amazon S3 bucket.

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

Uma política de função opcional Amazon Kendra para permitir o uso de uma chave mestra AWS KMS do cliente (CMK) para descriptografar arquivos em um bucket. Amazon S3

```
{
    "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"
                    ]
                }
            }
        }
    ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para sugestões de consulta

Ao usar um Amazon S3 arquivo como uma lista de bloqueio de sugestões de consulta, você fornece um papel que tem permissão para acessar o Amazon S3 arquivo e o Amazon S3 bucket. Se o arquivo de texto da lista de bloqueio (o Amazon S3 arquivo) no Amazon S3 bucket estiver criptografado, você deverá fornecer permissão para usar a chave mestra do AWS KMS cliente (CMK) para descriptografar os documentos.

IAM funções para sugestões de consulta

Uma política de função necessária Amazon Kendra para permitir o uso do Amazon S3 arquivo como sua lista de bloqueio de sugestões de consulta.

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

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

Uma política de função opcional Amazon Kendra para permitir o uso de uma chave mestra AWS KMS do cliente (CMK) para descriptografar documentos em um bucket. Amazon S3

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

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para mapeamento principal de usuários e grupos

Ao usar a PutPrincipalMappingAPI para mapear usuários para seus grupos para filtrar os resultados da pesquisa por contexto de usuário, você precisa fornecer uma lista de usuários ou subgrupos que pertencem a um grupo. Se sua lista tiver mais de 1.000 usuários ou subgrupos para um grupo, você precisará fornecer uma função que tenha permissão para acessar o Amazon S3 arquivo da sua lista e do Amazon S3 bucket. Se o arquivo de texto (o Amazon S3 arquivo) da lista no Amazon S3 bucket estiver criptografado, você deverá fornecer permissão para usar a chave mestra do AWS KMS cliente (CMK) para descriptografar os documentos.

IAM funções para mapeamento principal

Uma política de função necessária Amazon Kendra para permitir o uso do Amazon S3 arquivo como sua lista de usuários e subgrupos que pertencem a um grupo.

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

Uma política de função opcional Amazon Kendra para permitir o uso de uma chave mestra AWS KMS do cliente (CMK) para descriptografar documentos em um bucket. Amazon S3

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

```
            ]
        }
    ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

É recomendável que você inclua `aws:sourceAccount` e `aws:sourceArn` na política de confiança. Isso limita as permissões e verifica com segurança se `aws:sourceAccount` `aws:sourceArn` elas são as mesmas fornecidas na política de IAM função da `sts:AssumeRole` ação. Isso impede que entidades não autorizadas acessem suas IAM funções e suas permissões. Para obter mais informações, consulte o AWS Identity and Access Management guia sobre o problema confuso do deputado.

```
{
    "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 funções para AWS IAM Identity Center

Ao usar o UserGroupResolutionConfigurationobjeto para obter níveis de acesso de grupos e usuários de uma fonte de AWS IAM Identity Center identidade, você precisa fornecer uma função que tenha permissão de acesso IAM Identity Center.

IAM funções para AWS IAM Identity Center

Uma política de função necessária para Amazon Kendra permitir o acesso 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"
              ]
            }
          }
        }
    ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

IAM funções para Amazon Kendra experiências

Ao usar as UpdateExperienceAPIs CreateExperienceou para criar ou atualizar um aplicativo de pesquisa, você deve fornecer uma função que tenha permissão para acessar as operações necessárias e o IAM Identity Center.

IAM funções para experiência Amazon Kendra de pesquisa

Uma política de função necessária Amazon Kendra para permitir o acesso a `Query` operações, `QuerySuggestions` operações, `SubmitFeedback` operações e ao IAM Identity Center que armazena suas informações de usuários e 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"
          ]
        }
      }
    }
  ]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

É recomendável que você inclua `aws:sourceAccount` e `aws:sourceArn` na política de confiança. Isso limita as permissões e verifica com segurança se `aws:sourceAccount` `aws:sourceArn` elas são as mesmas fornecidas na política de IAM função da `sts:AssumeRole` ação. Isso impede que entidades não autorizadas acessem suas IAM funções e suas permissões. Para obter mais informações, consulte o AWS Identity and Access Management guia sobre o problema confuso do deputado.

```
{
    "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 funções para enriquecimento personalizado de documentos

Ao usar o CustomDocumentEnrichmentConfigurationobjeto para aplicar alterações avançadas nos metadados e no conteúdo do documento, você deve fornecer uma função que tenha as permissões necessárias para execução `PreExtractionHookConfiguration` e/ou. `PostExtractionHookConfiguration` Configure uma função do Lambda para `PreExtractionHookConfiguration` e/ou `PostExtractionHookConfiguration` aplicar alterações avançadas nos metadados e no conteúdo do documento durante o processo de ingestão. Se você optar por ativar a criptografia do lado do servidor para seu Amazon S3 bucket, deverá fornecer permissão para usar a chave mestra do AWS KMS cliente (CMK) para criptografar e descriptografar os objetos armazenados em seu bucket. Amazon S3

IAM funções para enriquecimento personalizado de documentos

Uma política de função necessária para Amazon Kendra permitir a execução `PreExtractionHookConfiguration` e `PostExtractionHookConfiguration` com criptografia para seu 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"
  }]
}
```

Uma política de função opcional para Amazon Kendra permitir a execução `PreExtractionHookConfiguration` e `PostExtractionHookConfiguration` sem criptografia para seu 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": [
      "lambda:InvokeFunction"
    ],
    "Resource": "arn:aws:lambda:your-region:your-account-id:function:lambda-function"
  }]
}
```

Uma política de confiança para Amazon Kendra permitir assumir uma função.

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

É recomendável que você inclua `aws:sourceAccount` e `aws:sourceArn` na política de confiança. Isso limita as permissões e verifica com segurança se `aws:sourceAccount` `aws:sourceArn`

elas são as mesmas fornecidas na política de IAM função da `sts:AssumeRole` ação. Isso impede que entidades não autorizadas acessem suas IAM funções e suas permissões. Para obter mais informações, consulte o AWS Identity and Access Management guia sobre o problema confuso do deputado.

```
{
    "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/*"
                }
            }
        }
    ]
}
```

Implantação de Amazon Kendra

Quando chega a hora de implantar a pesquisa do Amazon Kendra no site, fornecemos o código-fonte que pode ser usado com o React para obter uma vantagem inicial no aplicativo. O código-fonte é fornecido gratuitamente sob uma licença modificada do MIT. Você pode usá-lo como ele está ou alterá-lo de acordo com suas necessidades. O aplicativo React fornecido é um exemplo que pode ajudar você a começar. Ele não é um aplicativo pronto para produção.

Para implantar um aplicativo de pesquisa sem código e gerar uma URL de endpoint para a página de pesquisa com controle de acesso, consulte Amazon KendraExperience Builder.

O código de exemplo a seguir adiciona a pesquisa do Amazon Kendra a um aplicativo da Web React existente:

- https://kendrasamples.s3.amazonaws.com/kendrasamples-react-app.zip: arquivos de amostra que os desenvolvedores podem usar para criar uma experiência de pesquisa funcional no aplicativo da Web React existente.

Os exemplos são modelados de acordo com a página de pesquisa do console do Amazon Kendra. Eles têm os mesmos recursos para pesquisar e exibir os resultados da pesquisa. Você pode usar o exemplo completo ou escolher apenas um dos recursos para seu próprio uso.

Para ver os três componentes da página de pesquisa no console do Amazon Kendra, escolha o ícone do código (</>) no menu à direita. Passe o mouse sobre cada seção para ver uma breve descrição do componente e obter o URL da fonte do componente.

Tópicos

- Visão geral
- Pré-requisitos
- Configurar o exemplo
- Página de pesquisa principal
- Componente de pesquisa
- Componente de resultados
- Componente de facetas
- Componente de paginação

- Criando uma experiência de pesquisa sem código

Visão geral

Adicione o código de exemplo a um aplicativo da Web React existente para ativar a pesquisa. O código de exemplo inclui um arquivo Readme com etapas para configurar um novo ambiente de desenvolvimento do React. Os dados de exemplo no exemplo de código podem ser usados para demonstrar uma pesquisa. Os arquivos e os componentes de pesquisa no código de exemplo são estruturados da seguinte forma:

- Página de pesquisa principal (`Search.tsx`): essa é a página principal que contém todos os componentes. É nela que você integra o aplicativo com a API do Amazon Kendra.
- Barra de pesquisa: esse é o componente em que um usuário insere um termo de pesquisa e chama a função de pesquisa.
- Resultados: esse é o componente que exibe os resultados de Amazon Kendra. Ele tem três componentes: respostas sugeridas, resultados de perguntas frequentes e documentos recomendados.
- Facetas: esse é o componente que mostra as facetas nos resultados da pesquisa e permite que você escolha uma faceta para restringir a pesquisa.
- Paginação: esse é o componente que pagina a resposta do Amazon Kendra.

Pré-requisitos

Antes de começar, você precisará fazer o seguinte:

- Node.js e npm instalados. É necessária a versão 19 ou anterior do Node.js.
- Python 3 ou Python 2 baixados e instalados.
- SDK for Java ou AWS SDK for JavaScriptpara fazer chamadas de API para o Amazon Kendra.
- Uma aplicação da Web existente do React. O código de exemplo inclui um arquivo Readme com etapas para configurar um novo ambiente de desenvolvimento do React, incluindo o uso de estruturas/bibliotecas necessárias. Você também pode seguir as instruções de início rápido na documentação do React sobre a criação de um aplicativo da Web do React.
- As bibliotecas e as dependências necessárias configuradas no ambiente de desenvolvimento. O código de exemplo inclui um arquivo Readme que lista as bibliotecas e dependências de

pacotes necessárias. Isso `sass` é obrigatório, pois `node-sass` está obsoleto. Se você instalou anteriormente o `node-sass`, desinstale-o e instale o `sass`.

Configurar o exemplo

Um procedimento completo para adicionar a pesquisa do Amazon Kendra a um aplicativo React está no arquivo Readme incluído no exemplo de código.

Para começar a usar o kendrasamples-react-app.zip

1. Certifique-se de ter concluído o Pré-requisitos, incluindo o download e a instalação do Node.js e do npm.
2. Baixe o kendrasamples-react-app.zip e descompacte.
3. Abra o terminal e vá para `aws-kendra-example-react-app/src/services/`. Forneça suas credenciais da `local-dev-credentials.json`. Não adicione esse arquivo a nenhum repositório público.
4. Acesse `aws-kendra-example-react-app` e instale as dependências em `package.json`. Execute `npm install`.
5. Iniciar uma versão de demonstração do aplicativo no servidor local. Execute `npm start`. Você pode parar o servidor local digitando no teclado `Cmd/Ctrl + C`.
6. Você pode alterar a porta ou o host (por exemplo, endereço IP) acessando `package.json` e atualizando o host e a porta: `"start": "HOST=[host] PORT=[port] react-scripts start"`. Se você usa Windows: `"start": "set HOST=[host] && set PORT=[port] && react-scripts start"`.
7. Se você tiver um domínio de site registrado, poderá especificá-lo em `package.json` após o nome do aplicativo. Por exemplo, `"homepage": "https://mywebsite.com"`. Execute `npm install` novamente para atualizar novas dependências e, em seguida, execute `npm start`.
8. Para construir o aplicativo, execute `npm build`. Faça o upload do conteúdo do diretório de compilação para o seu provedor de host.

> **Warning**
>
> O aplicativo React não está pronto para produção. É um exemplo de implantação de um aplicativo para pesquisa do Amazon Kendra.

Página de pesquisa principal

A página de pesquisa principal (`Search.tsx`) contém todos os exemplos de componentes de pesquisa. Ele inclui o componente da barra de pesquisa para saída, os componentes de resultados para exibir a resposta da API de Consulta e um componente de paginação para paginar a resposta.

Componente de pesquisa

O componente de pesquisa fornece uma caixa de texto para inserir o texto da consulta. A função do `onSearch` é um hook que chama a função principal `Search.tsx` para fazer a chamada de API Amazon Kendra Consulta.

Componente de resultados

O componente de resultados mostra a resposta da API `Query`. Os resultados são mostrados em três áreas distintas.

- Respostas sugeridas: esses são os principais resultados retornados pela API `Query`. Ela contém até três respostas sugeridas. Na resposta, elas têm o tipo de resultado `ANSWER`.
- Respostas de perguntas frequentes: essas são os resultados das perguntas frequentes retornados pela resposta. As perguntas frequentes são adicionadas ao índice separadamente. Na resposta, elas têm o tipo de resultado `QUESTION_ANSWER`. Para obter mais informações, consulte Perguntas e respostas.
- Documentos recomendados: esses são documentos adicionais que Amazon Kendra retornam na resposta. Na resposta da API `Query`, elas têm o tipo de `DOCUMENT`.

Os componentes de resultados compartilham um conjunto de componentes para recursos como destaque, títulos, links e muito mais. Os componentes compartilhados devem estar presentes para que os componentes do resultado funcionem.

Componente de facetas

O componente de facetas lista as facetas disponíveis nos resultados da pesquisa. Cada faceta classifica a resposta em uma dimensão específica, como autor. Você pode refinar a pesquisa para uma faceta específica escolhendo uma na lista.

Depois de selecionar uma faceta, o componente chama `Query` com um filtro de atributo que restringe a pesquisa a documentos que correspondam à faceta.

Componente de paginação

O componente de paginação permite que você exiba os resultados da pesquisa da API `Query` em várias páginas. Ele chama a API `Query` com os parâmetros `PageSize` e `PageNumber` para obter uma página específica de resultados.

Criando uma experiência de pesquisa sem código

O aplicativo de pesquisa do Amazon Kendra pode ser implantado em alguns cliques sem a necessidade de nenhum código de front-end. Amazon Kendra O Experience Builder ajuda você a criar e implantar um aplicativo de pesquisa totalmente funcional em alguns cliques para começar a pesquisar imediatamente. Você pode personalizar a página de pesquisa e ajustar a pesquisa para adaptar a experiência às necessidades dos usuários. O Amazon Kendra gera um URL de endpoint exclusivo e totalmente hospedado da sua página de pesquisa para começar a pesquisar os documentos e as perguntas frequentes. Você pode criar rapidamente uma prova de conceito da experiência de pesquisa e compartilhá-la com outras pessoas.

Use o modelo de experiência de pesquisa disponível no construtor para personalizar sua pesquisa. Você pode convidar outras pessoas para colaborar na criação de sua experiência de pesquisa ou avaliar os resultados da pesquisa para fins de ajuste. Quando a experiência de pesquisa estiver pronta para que os usuários comecem a pesquisar, é só compartilhar o URL seguro do endpoint.

Como a pesquisa do Experience Builder funciona

O processo geral de criação de uma experiência de pesquisa é o seguinte:

1. Crie a experiência de pesquisa dando a ela um nome, uma descrição e escolhendo as fontes de dados que deseja usar para a experiência de pesquisa.
2. Configure a lista de usuários e grupos em AWS IAM Identity Center e, em seguida, atribua a eles direitos de acesso à sua experiência de pesquisa. Você se inclui como proprietário da experiência. Para obter mais informações, consulte the section called “Fornecer acesso à sua página de pesquisa”.
3. Abra o Experience Builder do Amazon Kendra para criar e ajustar a página de pesquisa. Compartilhe o URL do endpoint da sua experiência de pesquisa com outras pessoas a quem você atribui direitos de acesso de edição própria ou direitos de acesso de visualização e pesquisa.

Chame a API CreateExperience para criar e configurar sua experiência de pesquisa. Ao usar o console, selecione o índice e, em seguida, selecione Experiências no menu de navegação para configurar a experiência.

Projete e ajuste a experiência de pesquisa

Depois de criar e configurar a experiência de pesquisa, abra a experiência de pesquisa usando uma URL de endpoint para começar a personalizar a pesquisa como proprietário com direitos de acesso de editor. Digitalize a consulta na caixa de pesquisa e personaliza a pesquisa usando as opções de edição no painel lateral para ver como elas se aplicam à sua página. Quando estiver pronto para publicar, selecione Publicar. Você também pode alternar entre Alternar para visualização ao vivo, para ver a versão mais recente publicada da página de pesquisa, e Alternar para o modo de criação, para editar ou personalizar a página de pesquisa.

Veja a seguir formas de personalizar a experiência de pesquisa.

Filtro

Adicione pesquisa facetada ou filtre por atributos do documento. Isso inclui atributos personalizados. Você pode adicionar um filtro usando seus campos de metadados configurados. Por exemplo, para pesquisar por facetas por cada categoria de cidade, use um atributo de documento `_category` personalizado que contenha todas as categorias de cidade.

Resposta sugerida

Adicione respostas geradas por machine learning às consultas dos usuários. Por exemplo, “Quão difícil é esse curso?” . O Amazon Kendra pode recuperar o texto mais relevante em todos os documentos referentes à dificuldade de um curso e sugerir a resposta mais relevante.

Perguntas frequentes

Adicione um documento de perguntas frequentes para fornecer respostas às perguntas mais frequentes. Por exemplo, “Quantas horas faltam para concluir este curso?”. O Amazon Kendra pode usar o documento de perguntas frequentes que contém a resposta a essa pergunta e dar a resposta correta.

Classificar

Adicione a classificação dos resultados da pesquisa para que os usuários possam organizar os resultados por relevância, hora de criação, hora da última atualização e outros critérios de classificação.

Documentos

Configure como os documentos ou os resultados da pesquisa são exibidos na página de pesquisa. Você pode configurar quantos resultados são exibidos na página, incluir paginação, como números de página, ativar um botão de feedback do usuário e organizar como os campos de metadados do documento são exibidos em um resultado de pesquisa.

Idioma

Selecione um idioma para filtrar os resultados da pesquisa ou documentos no idioma selecionado.

Barra de pesquisa

Configure o tamanho e o espaço reservado para o texto da caixa de pesquisa, além de permitir sugestões de consulta.

Ajuste de relevância

Adicione impulsionamento aos campos de metadados do documento para dar mais peso a esses campos quando os usuários pesquisarem documentos. Você pode adicionar um peso que começa em 1 e aumenta gradualmente para 10. Você pode aumentar os tipos de campo de texto, data e numérico. Por exemplo, para dar a `_last_updated_at` e `_created_at` mais peso ou importância do que outros campos, atribua a esses campos um peso de 1 a 10, dependendo de sua importância. Você pode aplicar diferentes configurações de ajuste de relevância para cada aplicativo ou experiência de pesquisa.

Fornecer acesso à sua página de pesquisa

O acesso à sua experiência de pesquisa é feito pelo IAM Identity Center. Ao configurar a experiência de pesquisa, você concede a outras pessoas listadas no diretório do Identity Center acesso à sua página de pesquisa do Amazon Kendra. Eles recebem um e-mail que os orienta a fazer login usando as credenciais no IAM Identity Center para acessar a página de pesquisa. Você deve configurar o IAM Identity Center no nível da organização ou no nível do titular da conta no AWS Organizations. Para obter mais informações sobre o IAM Identity Center, consulte Introdução ao IAM Identity Center.

Você ativa as identidades do usuário no IAM Identity Center com a experiência de pesquisa e atribui permissões de acesso de Visualizador ou Proprietário usando a API ou o console.

- Visualizador: autorizado a fazer consultas, receber sugestões de respostas relevantes para a pesquisa e contribuir com comentários sobre o Amazon Kendra para continuar melhorando a pesquisa.
- Proprietário: autorizado a personalizar o design da página de pesquisa, ajustar a pesquisa e usar o aplicativo de pesquisa como Visualizador. Atualmente, não há suporte para desabilitar o acesso aos visualizadores no console.

Para atribuir acesso a outras pessoas à sua experiência de pesquisa, ative primeiro as identidades de usuário no IAM Identity Center com sua experiência do Amazon Kendra usando o objeto ExperienceConfiguration. Você especifica o nome do campo que contém os identificadores de usuários, como nome de usuário ou endereço de e-mail. Em seguida, conceda à sua lista de usuários acesso à experiência de pesquisa usando a API AssociateEntitiesToExperience e define as permissões como Visualizador ou Proprietário usando a API AssociatePersonasToEntities. Você especifica cada usuário ou grupo usando o objeto EntityConfiguration e se esse usuário ou grupo é um Visualizador ou Proprietário usando o objeto EntityPersonaConfiguraton.

Para atribuir a outras pessoas acesso à sua experiência de pesquisa usando o console, primeiro você precisa criar uma experiência e confirmar sua identidade e se você é o proprietário. Depois, você pode atribuir outros usuários ou grupos como visualizadores ou proprietários. No console, selecione seu índice e, em seguida, selecione Experiências no menu de navegação. Depois de criar sua experiência, você pode selecioná-la na lista. Acesse o Gerenciamento de acesso para atribuir usuários ou grupos como visualizadores ou proprietários.

Configurando uma experiência de pesquisa

Veja a seguir um exemplo de como configurar ou criar uma experiência de pesquisa.

Console

Para criar uma experiência de pesquisa do Amazon Kendra

1. No painel de navegação esquerdo, em Índices, selecione Experiências e, em seguida, selecione Criar experiência.
2. Na página Configurar experiência, insira um nome e uma descrição para a experiência, escolha as fontes de conteúdo e o perfil do IAM para sua experiência. Para obter mais informações sobre as funções do IAM, consulte as funções do IAM para as experiências do Amazon Kendra.

3. Na página Confirme sua identidade em um diretório do Identity Center, selecione o ID de usuário, como seu e-mail. Se você não tiver um diretório do Identity Center, digite seu nome completo e e-mail para criar um diretório do Identity Center. Isso inclui você como usuário da experiência e atribui automaticamente a você direitos de acesso de proprietário.
4. Na página Revisar para abrir o Experience Builder, revise os detalhes da configuração e selecione Criar experiência e abra o Experience Builder para editar a página de pesquisa.

CLI

Para criar uma experiência do 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 criar uma experiência do 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 criar um 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.");
    }
}
```

Ajustar a capacidade

Amazon Kendra fornece recursos para seu índice em unidades de capacidade. Cada unidade de capacidade fornece recursos adicionais para o índice. Existem unidades de capacidade separadas para o armazenamento de documentos e para consultas. Você só pode adicionar unidades de capacidade aos índices Amazon Kendra da Enterprise Edition. Você não pode adicionar capacidade a um índice do Developer Edition.

Uma unidade de capacidade de armazenamento de documentos fornece o seguinte armazenamento adicional para seu índice:

- 100 mil documentos ou 30 GB de armazenamento.

Uma unidade de capacidade de armazenamento de documentos fornece o seguinte armazenamento adicional para seu índice:

- 0,1 consulta por segundo ou aproximadamente 8 mil consultas por dia.

Cada índice vem com uma capacidade básica igual a 1 unidade de capacidade (30 GB de armazenamento e 0,1 consulta por segundo). Há um custo adicional para cada unidade de capacidade adicional. Para obter detalhes, consulte Definição de preço do Amazon Kendra.

Você pode adicionar até 100 unidades de capacidade extras aos seus recursos de armazenamento e consulta para um índice. Se precisar de mais unidades, entre em contato com o suporte.

Você pode ajustar as unidades de capacidade até 5 vezes por dia para atender às suas necessidades de uso. Você não pode reduzir a capacidade de armazenamento de documentos abaixo do número de documentos armazenados em seu índice. Por exemplo, se estiver armazenando 150 mil documentos, não poderá reduzir a capacidade de armazenamento abaixo de 1 unidade adicional.

Você pode visualizar os recursos que um índice está usando no console selecionando o nome do índice para abrir as configurações do índice e outras informações, ou você pode usar a DescribeIndexAPI.

Amazon Kendra também retorna exceções quando você excede a capacidade de um índice. Você recebe um `ServiceQuotaExceededException` quando o tamanho total extraído de todos os documentos excede o limite de um índice. Você recebe um `InvalidRequest` para

cada documento quando o número de documentos excede o limite de um índice. Você recebe um `ThrottlingException` quando o número de consultas por segundo excede o limite. Para obter mais informações sobre limites, consulte Cotas para o Amazon Kendra.

As consultas acumuladas durarão até 24 horas.

Visualizar a capacidade

Visualize os recursos que seu índice está usando com o Amazon Kendra console selecionando o nome do seu índice para acessar os detalhes. O console também apresenta gráficos de uso para você poder determinar quanta capacidade de armazenamento e consulta é usada pelo índice. Você pode usar essas informações para planejar quando adicionar mais capacidade.

Para visualizar o armazenamento de documentos e o uso de consultas (console)

1. Faça login no AWS Management Console e abra o Amazon Kendra console em https://console.aws.amazon.com/kendra/home.
2. Na lista de índices, escolha o índice que deseja acessar.
3. Role até a seção de configurações para ver o armazenamento total atual de documentos e a capacidade de consulta.

Para ver a capacidade usando a Amazon Kendra API, use o `CapacityUnits` parâmetro na DescribeIndexAPI.

Adicionar e remover capacidade

Se precisar de capacidade adicional para seu índice, você pode adicioná-la usando o console ou a Amazon Kendra API.

Como adicionar ou remover armazenamento ou capacidade de consulta (console)

1. Faça login no AWS Management Console e abra o Amazon Kendra console em https://console.aws.amazon.com/kendra/home.
2. Na lista de índices, escolha o índice que deseja acessar.
3. Selecione Editar ou selecione Editar no menu suspenso Ações.
4. Selecione Avançar para acessar a página de detalhes do aprovisionamento.

5. Adicione ou remova unidades de capacidade para armazenamento de documentos e/ou consultas.
6. Continue selecionando Avançar para acessar a página de revisão e, em seguida, selecione Atualizar para salvar as alterações.

Depois de atualizar a capacidade do índice, espere alguns minutos para que as alterações sejam aplicadas.

Para adicionar ou remover capacidade usando a Amazon Kendra API, use o `CapacityUnits` parâmetro na UpdateIndexAPI.

Amazon Kendra Capacidade de classificação inteligente

Uma unidade de capacidade oferece as solicitações adicionais de repontuação por segundo a seguir para um plano de execução de repontuação. Um plano de execução de repontuação é um recurso usado para provisionar a API Rescore.

- 0,01 solicitação por segundo.

Cada plano de execução de repontuação vem com uma capacidade básica igual a 1 unidade de capacidade (0,01 solicitação por segundo). Há um custo adicional para cada unidade de capacidade adicional. Para obter detalhes, consulte Definição de preço do Amazon Kendra.

Você pode adicionar até 1.000 unidades de capacidade extras para um plano de execução de repontuação. Se precisar de mais unidades, entre em contato com o suporte.

Capacidade para sugestões de consulta

Ao usar sugestões de consulta, há uma capacidade básica de consulta de 2,5 GetQuerySuggestionschamadas por segundo. A capacidade `GetQuerySuggestions` é cinco vezes a capacidade de consulta provisionada para um índice ou a capacidade básica de 2,5 chamadas por segundo, a que for maior. Por exemplo, a capacidade básica de um índice é de 0,1 consulta por segundo e a capacidade `GetQuerySuggestions` tem o valor básico de 2,5 chamadas por segundo. Se você adicionar mais 0,1 consulta por segundo para totalizar 0,2 consulta por segundo para um índice, a capacidade `GetQuerySuggestions` será de 2,5 chamadas por segundo (maior que 5 vezes 0,2 consulta por segundo).

Amazon Kendra capacidade de experiência

Capacidade para experiência de pesquisa

Amazon Kendra começa a limitar `Query` sua Amazon Kendra experiência com 15 solicitações por segundo e 40 solicitações por segundo para intermitência de consultas. `QuerySuggestions` `SubmitFeedback` Para um índice com mais de 150 unidades de capacidade de consulta, esses limites ainda são aplicáveis.

Por exemplo, as unidades de capacidade de consulta para seu índice são 150. Dessa maneira, a aplicação de experiência de pesquisa pode lidar com 15 solicitações por segundo. No entanto, se você escalasse para 200 unidades de capacidade de consulta, sua aplicação de experiência de pesquisa ainda processaria apenas 15 solicitações por segundo. Se você limitar o índice a 100 unidades de capacidade de consulta, a aplicação de experiência de pesquisa processará apenas 10 solicitações por segundo.

Expansão de consultas adaptável

Amazon Kendra tem uma capacidade básica provisionada de 1 unidade de capacidade de consulta. Você pode usar até 8 mil consultas por dia com um throughput mínimo de 0,1 consulta por segundo (por unidade de capacidade de consulta). As consultas acumuladas duram até 24 horas e podem acomodar picos de tráfego. O volume de expansão permitido varia porque depende da carga do cluster em um determinado momento. Provisione unidades de capacidade de consulta suficientes para lidar com os níveis de pico de carga.

Uma abordagem adaptativa para lidar com picos inesperados de tráfego além da taxa de transferência provisionada é Amazon Kendra o intermitente de consultas adaptável incorporado. A expansão de consultas adaptável está disponível no Amazon Kendra Enterprise Edition.

O intermitente adaptativo de consultas é um recurso incorporado que permite aplicar a capacidade de consulta não utilizada para lidar com tráfego inesperado. Amazon Kendra acumula suas consultas não utilizadas na taxa de consultas provisionadas por segundo, a cada segundo, até o número máximo de consultas que você provisionou para seu índice. Amazon Kendra Essas consultas acumuladas são usadas para tráfego inesperado acima da capacidade alocada. O desempenho ideal da expansão de consultas adaptável pode variar, dependendo de vários fatores, como o tamanho total do índice, a complexidade da consulta, o acúmulo de consultas não utilizadas e a carga geral do índice. Realize seus próprios testes de carga para medir com precisão a capacidade de expansão.

Conceitos básicos

Esta seção mostra como criar uma fonte de dados e adicionar seus documentos a um Amazon Kendra índice. As instruções são fornecidas para o AWS console, o AWS CLI, um programa Python usando o. AWS SDK for Python (Boto3) e um programa Java usando o. AWS SDK for Java

Tópicos

- Pré-requisitos
- Introdução ao Amazon Kendra console
- Conceitos básicos (AWS CLI)
- Conceitos básicos (AWS SDK for Python (Boto3))
- Conceitos básicos (AWS SDK for Java)
- Introdução a uma fonte de dados do Amazon S3 (console)
- Introdução a uma fonte de dados do banco de dados MySQL (console)
- Introdução a uma fonte de AWS IAM Identity Center identidade (console)

Pré-requisitos

As etapas a seguir são pré-requisitos para os exercícios de conceitos básicos. As etapas mostram como configurar sua conta, criar uma IAM função que dê Amazon Kendra permissão para fazer chamadas em seu nome e indexar documentos de um Amazon S3 bucket. Um bucket do S3 é usado como exemplo, mas você pode usar outras fontes de dados Amazon Kendra compatíveis. Escolha as Fontes de dados.

Inscreva-se para um Conta da AWS

Se você não tiver um Conta da AWS, conclua as etapas a seguir para criar um.

Para se inscrever em um Conta da AWS

1. Abra https://portal.aws.amazon.com/billing/signup.
2. Siga as instruções on-line.

   Parte do procedimento de inscrição envolve receber uma chamada telefônica e digitar um código de verificação no teclado do telefone.

Quando você se inscreve em um Conta da AWS, um Usuário raiz da conta da AWSé criado. O usuário-raiz tem acesso a todos os Serviços da AWS e recursos na conta. Como prática recomendada de segurança, atribua o acesso administrativo a um usuário e use somente o usuário-raiz para executar tarefas que exigem acesso de usuário-raiz.

AWS envia um e-mail de confirmação após a conclusão do processo de inscrição. A qualquer momento, é possível visualizar as atividades da conta atual e gerenciar sua conta acessando https://aws.amazon.com/ e selecionando Minha conta.

Criar um usuário com acesso administrativo

Depois de se inscrever em um Conta da AWS, proteja seu Usuário raiz da conta da AWS AWS IAM Identity Center, habilite e crie um usuário administrativo para que você não use o usuário root nas tarefas diárias.

Proteja seu Usuário raiz da conta da AWS

1. Faça login AWS Management Consolecomo proprietário da conta escolhendo Usuário raiz e inserindo seu endereço de Conta da AWS e-mail. Na próxima página, digite sua senha.

   Para obter ajuda ao fazer login usando o usuário-raiz, consulte Signing in as the root user (Fazer login como usuário-raiz) no Guia do usuário do Início de Sessão da AWS .
2. Habilite a autenticação multifator (MFA) para o usuário-raiz.

   Para obter instruções, consulte Habilitar um dispositivo de MFA virtual para seu usuário Conta da AWS raiz (console) no Guia do usuário do IAM.

Criar um usuário com acesso administrativo

1. Habilitar o IAM Identity Center.

   Para obter instruções, consulte Habilitar AWS IAM Identity Center no Guia do usuário do AWS IAM Identity Center .
2. No Centro de Identidade do IAM, conceda o acesso administrativo para um usuário.

   Para ver um tutorial sobre como usar o Diretório do Centro de Identidade do IAM como fonte de identidade, consulte Configurar o acesso do usuário com o padrão Diretório do Centro de Identidade do IAM no Guia AWS IAM Identity Center do usuário.

Iniciar sessão como o usuário com acesso administrativo

- Para fazer login com seu usuário do Centro de Identidade do IAM, use a URL de login que foi enviada ao seu endereço de e-mail quando você criou o usuário do Centro do Usuário do IAM.

  Para obter ajuda para fazer login usando um usuário do IAM Identity Center, consulte Como fazer login no portal de AWS acesso no Guia Início de Sessão da AWS do usuário.

Atribuir acesso a usuários adicionais

1. No Centro de Identidade do IAM, crie um conjunto de permissões que siga as práticas recomendadas de aplicação de permissões com privilégio mínimo.

   Para obter instruções, consulte Create a permission set no Guia do usuário do AWS IAM Identity Center .
2. Atribua usuários a um grupo e, em seguida, atribua o acesso de autenticação única ao grupo.

   Para obter instruções, consulte Add groups no Guia do usuário do AWS IAM Identity Center .

- Se você estiver usando um bucket do S3 contendo documentos para testar Amazon Kendra, crie um bucket do S3 na mesma região que você está usando. Amazon Kendra Para obter instruções, consulte Como criar um bucket do S3 no Guia do usuário do Amazon Simple Storage Service.

  Carregar os documentos no bucket do S3. Para obter mais informações, consulte Upload, download e gerenciamento de objetos no Manual do usuário do Amazon Simple Storage Service.

  Se estiver usando outra fonte de dados, use um site ativo e credenciais para se conectar à fonte de dados.

Se estiver usando o console, comece com Introdução ao Amazon Kendra console.

Amazon Kendra recursos: AWS CLI, SDK, console

Há certas permissões necessárias para usar a CLI, o SDK ou o console.

Amazon Kendra Para usar a CLI, o SDK ou o console, você deve ter permissões Amazon Kendra para criar e gerenciar recursos em seu nome. Dependendo do seu caso de uso, essas permissões

incluem acesso à própria Amazon Kendra API, AWS KMS keys se você quiser criptografar seus dados por meio de uma CMK personalizada, o diretório do Identity Center se quiser se integrar AWS IAM Identity Center ou criar uma experiência de pesquisa. Para obter uma lista completa de permissões para diferentes casos de uso, consulte as Funções do IAM.

Primeiro, anexe as permissões abaixo ao seu usuário do 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": "*"
    }
  ]
}
```

Segundo, se você usa a CLI ou o SDK, também deve criar uma IAM função e uma política para acessar. Amazon CloudWatch Logs Se você estiver usando o console, não precisará criar uma função e uma política da IAM para isso. Você cria isso como parte do procedimento do console.

Para criar uma IAM função e uma política para o AWS CLI SDK que permita Amazon Kendra acessar seu Amazon CloudWatch Logs.

1. Faça login AWS Management Console e abra o console do IAM em https://console.aws.amazon.com/iam/.
2. À esquerda, escolha Políticas e selecione Criar política.
3. Escolha JSON e substitua a política padrão conforme a seguir:

```
{
    "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. Escolha Revisar política.
5. Nomeie a política como "KendraPolicyForGettingStartedIndex" e clique em Criar política.
6. No painel de navegação, escolha Funções e Criar função.

7. Escolha Outra AWS conta e digite o ID da sua conta em ID da conta. Escolha Próximo: permissões.
8. Escolha a política criada acima e, em seguida, escolha Próximo: tags
9. Não adicione tag nenhuma. Escolha Próximo: revisar.
10. Nomeie a função como "KendraRoleForGettingStartedIndex" e clique em Criar função.
11. Encontre a função que você acabou de criar. Escolha o nome da função para abrir o Resumo. Escolha Relações de confiança e selecione Editar relação de confiança.
12. Substitua o relacionamento de confiança existente pelo seguinte:

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

13. Selecione Atualizar política de confiança.

Em terceiro lugar, se você usa um Amazon S3 para armazenar seus documentos ou está usando o S3 para testar Amazon Kendra, você também deve criar uma IAM função e uma política para acessar seu bucket. Se estiver usando outra fonte de dados, consulte Funções do IAM para fontes de dados.

Para criar uma IAM função e uma política que permitam Amazon Kendra acessar e indexar seu Amazon S3 bucket.

1. Faça login AWS Management Console e abra o console do IAM em https://console.aws.amazon.com/iam/.
2. À esquerda, escolha Políticas e selecione Criar política.
3. Escolha JSON e substitua a política padrão conforme a seguir:

```
{
```

```
    "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. Escolha Revisar política.
5. Nomeie a política KendraPolicyForGettingStartedDataSource "" e escolha Criar política.
6. No painel de navegação, escolha Funções e Criar função.
7. Escolha Outra AWS conta e digite o ID da sua conta em ID da conta. Escolha Próximo: permissões.
8. Escolha a política criada acima e, em seguida, escolha Próximo: tags
9. Não adicione tag nenhuma. Escolha Próximo: revisar.
10. Nomeie a função KendraRoleForGettingStartedDataSource "" e escolha Criar função.

11. Encontre a função que você acabou de criar. Escolha o nome da função para abrir o Resumo. Escolha Relações de confiança e selecione Editar relação de confiança.
12. Substitua o relacionamento de confiança existente pelo seguinte:

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

13. Selecione Atualizar política de confiança.

Dependendo de como você deseja usar a Amazon Kendra API, faça o seguinte.

- Conceitos básicos (AWS CLI)
- Conceitos básicos (AWS SDK for Java)
- Conceitos básicos (AWS SDK for Python (Boto3))

Introdução ao Amazon Kendra console

Os procedimentos a seguir mostram como criar e testar um Amazon Kendra índice usando o AWS console. Com base nos procedimentos, crie um índice e uma fonte de dados para um índice. Por fim, teste o índice fazendo uma solicitação de pesquisa.

Etapa 1: para criar um índice (console)

1. Faça login no AWS Management Console e abra o Amazon Kendra console em https://console.aws.amazon.com/kendra/.
2. Na seção Índices, escolha Criar índices.
3. Na página Especificar detalhes do índice, dê um nome e uma descrição ao índice.

4. Em Função do IAM , escolha Criar uma nova função e digite o nome da função. A IAM função terá o prefixo "AmazonKendra-”.
5. Deixe os outros campos nos padrões determinados. Escolha Próximo.
6. Na página Configurar controle de acesso do usuário, escolha Próximo.
7. Na página de Detalhes do provisionamento, escolha Developer Edition.
8. Escolha Criar para criar seu índice.
9. Aguarde a criação do seu índice. Amazon Kendra provisiona o hardware para seu índice. Essa operação pode levar algum tempo.

Etapa 2: para adicionar uma fonte de dados a um índice (console)

1. Visualize as fontes de dados disponíveis para se Amazon Kendra conectar e indexar seus documentos.
2. No painel de navegação, selecione Fontes de dados e, em seguida, selecione Adicionar fonte de dados para a fonte de dados escolhida.
3. Siga as etapas para configurar a fonte de dados.

Etapa 3: para pesquisar um índice (console)

1. No painel de navegação, escolha a opção para pesquisar no índice.
2. Insira um termo de pesquisa apropriado para o índice. Os principais resultados e os principais resultados do documento serão exibidos.

Conceitos básicos (AWS CLI)

O procedimento a seguir mostra como criar um Amazon Kendra índice usando AWS CLI o. O procedimento cria uma fonte de dados, um índice e executa uma consulta no índice.

Para criar um Amazon Kendra índice (CLI)

1. Faça Pré-requisitos.
2. Insira o seguinte comando para criar um 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. Amazon Kendra Aguarde a criação do índice. Verifique o andamento usando o seguinte comando: Quando o status como `ACTIVE`, siga para a próxima etapa.

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

4. No prompt de comando, insira o comando a seguir para criar uma fonte de dados.

```
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"}}'
```

Se você se conectar à fonte de dados usando um esquema de modelo, configure o esquema de modelo.

```
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. A criação da fonte de dados demorará Amazon Kendra um pouco. Insira o comando a seguir para acompanhar o progresso. Quando o status for `ACTIVE`, siga para a próxima etapa.

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

6. Insira o comando a seguir para sincronizar a fonte de dados.

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

7. Amazon Kendra indexará sua fonte de dados. O tempo necessário depende do número de documentos. Você pode verificar o status da tarefa usando o seguinte: Quando o status for `ACTIVE`, siga para a próxima etapa.

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

8. Digite o comando a seguir para fazer uma consulta.

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

   Os resultados da pesquisa são exibidos no formato JSON.

Conceitos básicos (AWS SDK for Python (Boto3))

O programa a seguir é um exemplo de uso Amazon Kendra em um programa Python. O programa realiza as seguintes ações:

1. Cria um novo índice usando a CreateIndexoperação.
2. Aguardar a conclusão da criação do índice. Ele usa a DescribeIndexoperação para monitorar o status do índice.
3. Quando o índice está ativo, ele cria uma fonte de dados usando a CreateDataSourceoperação.
4. Aguardar a conclusão da criação da fonte de dados. Ele usa a DescribeDataSourceoperação para monitorar o status da fonte de dados.
5. Quando a fonte de dados está ativa, ela sincroniza o índice com o conteúdo da fonte de dados usando a StartDataSourceSyncJoboperação.

```
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.")
```

Conceitos básicos (AWS SDK for Java)

O programa a seguir é um exemplo de uso Amazon Kendra em um programa Java. O programa realiza as seguintes ações:

1. Cria um novo índice usando a CreateIndexoperação.
2. Aguardar a conclusão da criação do índice. Ele usa a DescribeIndexoperação para monitorar o status do índice.

3. Quando o índice está ativo, ele cria uma fonte de dados usando a CreateDataSourceoperação.
4. Aguardar a conclusão da criação da fonte de dados. Ele usa a DescribeDataSourceoperação para monitorar o status da fonte de dados.
5. Quando a fonte de dados está ativa, ela sincroniza o índice com o conteúdo da fonte de dados usando a StartDataSourceSyncJoboperação.

```
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");
    }
}
```

Introdução a uma fonte de dados do Amazon S3 (console)

Use o console do Amazon Kendra para começar a usar um bucket do Amazon S3como armazenamento de dados. Ao usar o console, especifique todas as informações de conexão necessárias para indexar o conteúdo do bucket. Para obter mais informações, consulte Amazon S3.

Use o procedimento a seguir para criar uma fonte de dados básica do bucket do S3 usando a configuração padrão. O procedimento pressupõe que você criou um índice seguindo as etapas na etapa 1 do Introdução ao Amazon Kendra console.

Para criar uma fonte de dados do bucket do S3 usando o console do Amazon Kendra

1. Faça login no AWS Management Console e abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/home.
2. Na lista de índices, escolha o índice que deseja adicionar à fonte de dados.
3. Escolha Adicionar fonte de dados.
4. Na lista de conectores de fonte de dados, escolha Amazon S3.
5. Na página Definir atributos, dê um nome à sua fonte de dados e, opcionalmente, uma descrição. Deixe o campo Tags em branco. Escolha Próximo para continuar.
6. No campo Inserir a localização da fonte de dados, insira o nome do bucket do S3 que contém seus documentos. Você pode inserir o nome diretamente ou procurar o nome escolhendo Procurar. O bucket deve estar na mesma região que o índice.
7. Em Função IAM, Escolha Criar uma nova função e digite o nome da função. Para obter mais informações, Funções IAM para fontes de dados do Amazon S3.
8. Na seção Definir cronograma de execução de sincronização, escolha Executar sob demanda.
9. Escolha Próximo para continuar.
10. Na página Revisar e criar, revise os detalhes da fonte de dados do S3. Se quiser fazer alterações, escolha o botão Editar ao lado do item que deseja alterar. Quando estiver satisfeito com suas escolhas, escolha Criar para criar a fonte de dados do S3.

Depois de escolher Criar, o Amazon Kendra começa a criar a fonte de dados. A criação da VPC leva alguns minutos para criar a fonte de dados.. Quando concluído, o status da fonte de dados muda de Criando para Ativo.

Depois de criar a fonte de dados, sincronize o índice Amazon Kendra com a fonte de dados. Escolha Sincronizar agora para iniciar o processo de sincronização. A sincronização da fonte de dados pode levar de alguns minutos a várias horas, dependendo do número e do tamanho dos documentos.

Introdução a uma fonte de dados do banco de dados MySQL (console)

Use o console Amazon Kendra para usar um banco de dados MySQL como fonte de dados. Ao usar o console, especifique todas as informações de conexão necessárias para indexar o conteúdo do banco de dados MySQL. Para obter mais informações, consulte Uso de uma origem dos dados de banco de dados.

Primeiro, crie um banco de dados MySQL e, em seguida, crie uma fonte de dados para o banco de dados.

Use o procedimento a seguir para criar um banco de dados MySQL básico. O procedimento pressupõe que você criou um índice seguindo a etapa 1 do Introdução ao Amazon Kendra console.

Para criar um banco de dados MySQL

1. Faça login no AWS Management Console e abra o console do Amazon RDS em https://console.aws.amazon.com/rds/.
2. No painel de navegação, escolha Grupos de sub-redes e Criar grupo de sub-rede DB.
3. Dê um nome ao grupo e selecione sua nuvem privada virtual (VPC). Para obter mais informações sobre como configurar uma VPC, consulte Como configurar o Amazon Kendra para usar uma VPC.
4. Adicione as sub-redes privadas da VPC. As sub-redes privadas são aquelas que não estão conectadas ao NAT. Escolha Criar.
5. No painel de navegação, selecione Bancos de dados e, em seguida, selecione Criar banco de dados.
6. Use os parâmetros a seguir para criar o banco de dados. Deixe os outros campos nos padrões determinados.

    - Opções de mecanismo: MySQL
    - Modelos: nível gratuito
    - Configurações de credenciais: insira e confirme uma senha
    - Em Conectividade, escolha Configuração de conectividade adicional. Faça as seguintes escolhas.
        - Grupo de sub-redes: escolha o grupo de sub-redes que você criou na etapa 4.

    - Grupo de segurança da VPC: escolha o grupo que contém as regras de entrada e saída criadas na sua VPC. Por exemplo, **`DataSourceSecurityGroup`**. Para obter mais informações sobre como configurar uma VPC, consulte Como configurar o Amazon Kendra para usar uma VPC.
    - Em Configuração adicional, defina o Nome do banco de dados inicial como **`content`**.
7. Escolha Criar banco de dados.
8. Na lista de bancos de dados, selecione o novo banco de dados. Anote o endpoint do banco de dados.
9. Depois de criar o banco de dados, crie uma tabela para armazenar os documentos. A criação de uma tabela está fora do escopo dessas instruções. Ao usar ou para criar tarefas, observe o seguinte:

    - Nome do banco de dados: **`content`**
    - Nome da tabela: **`documents`**
    - Colunas: **`ID`**, **`Title`**, **`Body`** e **`LastUpdate`**. Você pode incluir colunas adicionais, se quiser.

Agora que criou o banco de dados MySQL, crie uma fonte de dados para o banco de dados.

Para criar uma fonte de dados MySQL

1. Faça login no AWS Management Console e abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/home.
2. No painel de navegação, escolha Índices e depois escolha o seu índice.
3. Escolha Adicionar fontes de dados e, em seguida, escolha Amazon RDS.
4. Digite um nome e uma descrição para a fonte de dados e, em seguida, escolha Próximo.
5. Escolha MySQL.
6. Em Acesso à conexão, insira as seguintes informações:

    - Endpoint: o endpoint do banco de dados criado anteriormente.
    - Porta: o número da porta do banco de dados. A porta padrão do MySQL é 3.306.
    - Tipo de autenticação: escolha Novo.
    - Novo nome de contêiner secreto: um nome para o contêiner do Secrets Manager para as credenciais do banco de dados.
    - Nome de usuário: o nome de um usuário com acesso administrativo ao banco de dados.

- Senha: a senha do usuário e, em seguida, escolha Salvar autenticação.
- Nome do banco de dados: **`content`**.
- Nome da tabela: **`documents`**.
- Perfil do IAM: escolha Criar uma nova função e digite o nome da função.

7. Na Configuração da coluna: insira o seguinte:
   - Nome da coluna do ID do documento: **`ID`**
   - Nome da coluna do título do documento: **`Title`**
   - Nome da coluna de dados do documento: **`Body`**
8. Em Detecção de alteração da coluna, insira o seguinte:
   - Colunas de detecção de alterações: **`LastUpdate`**
9. Em Configurar VPC e grupo de segurança, forneça o seguinte:
   - Em Nuvem privada virtual [VPC], selecione uma VPC.
   - Em sub-redes: escolha as sub-redes privadas criadas para a sua VPC.
   - Em Grupo de segurança da VPC, escolha o grupo que contém as regras de entrada e saída criadas na sua VPC para os bancos de dados MySQL. Por exemplo, **`DataSourceSecurityGroup`**.
10. Na seção Definir cronograma de execução de sincronização, escolha Executar sob demanda e, em seguida, Próximo.
11. Em Mapeamento do campo da fonte de dados, escolha Próximo.
12. Revise a configuração da fonte de dados para se certificar de que está correta. Quando estiver convencido de que tudo está correto, escolha Criar.

Introdução a uma fonte de AWS IAM Identity Center identidade (console)

Uma fonte de AWS IAM Identity Center identidade contém informações sobre seus usuários e grupos. Isso é útil para configurar a filtragem de contexto do usuário, que Amazon Kendra filtra os resultados da pesquisa para diferentes usuários com base no acesso do usuário ou do grupo aos documentos.

Para criar uma fonte de identidade do IAM Identity Center, ative o IAM Identity Center e crie uma organização no AWS Organizations. Ao ativar o IAM Identity Center e criar uma organização pela primeira vez, ele automaticamente usa como padrão o diretório do Identity Center como fonte de identidade. Você pode mudar para o Active Directory (gerenciado ou autogerenciado pela Amazon) ou um provedor de identidade externo como sua fonte de identidade. Você deve seguir a orientação correta para isso, consulte Alteração da fonte de identidade do IAM Identity Center. Você pode ter somente uma fonte de identidade por organização.

Para que usuários e grupos tenham diferentes níveis de acesso aos documentos, inclua os usuários e os grupos na sua lista de controle de acesso ao ingerir documentos no índice. Isso permite que seus usuários e grupos Amazon Kendra pesquisem documentos de acordo com seu nível de acesso. Ao emitir uma consulta, o ID do usuário precisa corresponder exatamente ao nome do usuário no IAM Identity Center.

Você também deve conceder as permissões necessárias para usar o IAM Identity Center com Amazon Kendra. Para obter mais informações, consulte Funções para o IAM Identity Center do IAM.

Para configurar uma fonte de identidade do IAM Identity Center

1. Abra o console do Centro de Identidade do IAM.
2. Escolha Ativar o IAM Identity Center e, em seguida, escolha Criar AWS organização.

   O diretório do Identity Center é criado por padrão e um e-mail é enviado a você para verificar o endereço de e-mail associado à organização.
3. Para adicionar um grupo à sua AWS organização, no painel de navegação, escolha Grupos.
4. Na página Grupos, escolha Criar grupo e insira um nome e uma descrição para o grupo na caixa de diálogo.. Escolha Criar.
5. Para adicionar um usuário à sua organização, no painel de navegação, escolha Usuários.
6. Na página Usuários, selecione Adicionar usuário. Em Detalhes do usuário, especifique todos os campos obrigatórios. Em Senha, escolha Enviar um e-mail para o usuário. Escolha Próximo.
7. Para adicionar um usuário a um grupo, escolha Grupos e selecione um grupo.
8. Na página Detalhes do grupo, em Membros do grupo, escolha Adicionar usuários.
9. Na página Adicionar usuários ao grupo, localize os usuários que você deseja adicionar como membros do grupo. É possível selecionar vários usuários para adicionar ao grupo.
10. Para sincronizar sua lista de usuários e grupos com o IAM Identity Center, altere a fonte de identidade para Active Directory ou provedor de identidade externo.

O diretório do Identity Center é a fonte de identidade padrão e exige adicionar manualmente os usuários e os grupos usando essa fonte se você não tiver sua lista gerenciada por um provedor. Você deve seguir a orientação correta para isso, consulte Alteração da fonte de identidade do IAM Identity Center.

> **Note**
>
> Se estiver usando o Active Directory ou um provedor de identidade externo como fonte de identidade, mapeie os endereços de e-mail dos usuários para os nomes de usuário do IAM Identity Center ao especificar o protocolo System for Cross-domain Identity Management (SCIM) . Para obter mais informações, consulte o Guia do IAM Identity Center no SCIM para ativar o IAM Identity Center.

Depois de configurar a fonte de identidade do IAM Identity Center, você pode ativá-la no console ao criar ou editar o índice. Acesse Controle de acesso do usuário nas configurações do índice e edite as configurações para permitir a busca de informações do grupo de usuários no IAM Identity Center.

Você também pode ativar o IAM Identity Center usando o UserGroupResolutionConfigurationobjeto. Você fornece `UserGroupResolutionMode` o anúncio `AWS_SSO` e cria uma IAM função que dá permissão para chamar`sso:ListDirectoryAssociations`,`sso-directory:SearchUsers`,`sso-directory:ListGroupsForUser`,`sso-directory:DescribeGroups`.

> **Warning**
>
> Amazon Kendra atualmente não suporta o uso `UserGroupResolutionConfiguration` com uma conta de membro da AWS organização para sua fonte de identidade do IAM Identity Center. É necessário criar o índice na conta de gerenciamento da organização para usar o `UserGroupResolutionConfiguration`.

Veja a seguir uma visão geral de como configurar uma fonte de dados com `UserGroupResolutionConfiguration` e controle de acesso do usuário para filtrar os resultados da pesquisa no contexto do usuário. Isso pressupõe que você já tenha criado um índice e uma IAM função para índices. Você cria um índice e fornece a IAM função usando a CreateIndexAPI.

Configurando uma fonte de dados com **`UserGroupResolutionConfiguration`** e filtragem de contexto de usuário

1. Crie uma função do IAM que dê permissão para acessar sua fonte de identidade do IAM Identity Center.
2. Configure `UserGroupResolutionConfiguration`definindo o modo `AWS_SSO` e ligue UpdateIndexpara atualizar seu índice para usar o IAM Identity Center.
3. Se você quiser usar o controle de acesso do usuário baseado em tokens para filtrar os resultados da pesquisa no contexto do usuário, UserContextPolicydefina como `USER_TOKEN` quando você ligar. `UpdateIndex` Caso contrário, Amazon Kendra rastreia a lista de controle de acesso de cada um dos seus documentos para a maioria dos conectores de fonte de dados. Você também pode filtrar os resultados da pesquisa no contexto do usuário na API Consulta fornecendo informações do usuário e do grupo no `UserContext`. Você também pode mapear usuários para seus grupos usando PutPrincipalMappingpara que você só precise fornecer o ID do usuário ao emitir a consulta.
4. Crie uma função do IAM que conceda permissão para acessar a fonte de dados.
5. Configurar a fonte de dados Fornece as informações de conexão necessárias para se conectar a fonte de dados.
6. Crie uma fonte de dados usando a CreateDataSourceAPI. Forneça o objeto do `DataSourceConfiguration`, que inclui `TemplateConfiguration`, a ID do seu índice, a função do IAM da fonte de dados, o tipo da fonte de dados e dê um nome à fonte de dados. Você também pode atualizar a fonte de dados.

Para alterar uma fonte de identidade do IAM Identity Center

> **Warning**
>
> Alterar sua fonte de identidade nas Configurações do IAM Identity Center pode afetar a preservação das informações do usuário e do grupo. Para fazer isso com segurança, é recomendável que você revise as Considerações para alterar a fonte de identidade. Quando você altera a fonte de identidade, uma nova ID da fonte de identidade é gerada. Verifique se você está usando a ID correta antes de ativar o `AWS_SSO` modo UserGroupResolutionConfiguration.

Para alterar uma fonte de identidade do IAM Identity Center

1. Abra o console do Centro de Identidade do IAM.
2. Escolha Configurações.
3. Na página Configurações, em Fonte de identidade, escolha Alterar.
4. Na página Alterar fonte de identidade, selecione sua fonte de identidade preferida e escolha Próximo.

Criar um índice

Você pode criar um índice usando o console ou chamando a CreateIndexAPI. Você pode usar o AWS Command Line Interface (AWS CLI) ou o SDK com a API. Depois de criar o índice, você pode adicionar documentos diretamente a ele ou de uma fonte de dados.

Para criar um índice, você deve fornecer o Amazon Resource Name (ARN) de uma função AWS Identity and Access Management (IAM) para os índices acessarem. CloudWatch Para obter mais informações, consulte Funções do IAM para índices.

As guias a seguir fornecem um procedimento para criar um índice usando o AWS Management Console, e exemplos de código para usar os AWS CLI SDKs de Python e Java.

Console

Para criar um índice

1. Faça login no AWS Management Console e abra o Amazon Kendra console em https://console.aws.amazon.com/kendra/.
2. Na seção Índices, escolha Criar índices.
3. Na página Especificar detalhes do índice, dê um nome e uma descrição.
4. Na IAM função, forneça uma IAM função. Para encontrar uma função, escolha uma das funções em sua conta que contenham a palavra “kendra” ou insira o nome de outra função. Para obter mais informações sobre as permissões que a função exige, consulte funções do IAM para índices.
5. Escolha Próximo.
6. Na página Configurar controle de acesso do usuário, escolha Próximo. Você pode atualizar o índice para usar tokens para controle de acesso depois de criar um índice. Para obter mais informações, consulte Controlando o acesso aos documentos.
7. Na página Detalhes de provisionamento, escolha Criar.
8. Pode levar algum tempo para que o índice seja criado. Verifique a lista de índices para acompanhar o progresso da criação do seu índice. Quando o status do índice é `ACTIVE`, o índice está pronto para uso.

AWS CLI

Para criar um índice

1. Insira o seguinte comando para criar um índice. `role-arn`Deve ser o Amazon Resource Name (ARN) de uma IAM função que possa executar Amazon Kendra ações. Para obter mais informações, consulte Funções do IAM.

   O comando a seguir é formatado para Linux e macOS. Para Windows, substitua o caractere de continuação de linha do Unix (\) pelo circunflexo (^).

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

2. Pode levar algum tempo para que o índice seja criado. Para verificar o estado do índice, use o ID do índice retornado `create-index` com o comando a seguir. Quando o status do índice é `ACTIVE`, o índice está pronto para uso.

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

Python

Para criar um índice

- Forneça valores para as seguintes variáveis no exemplo de código a seguir:
  - `description`: uma descrição da índice que você está criando. Isso é opcional.
  - `index_name`: o nome da índice que você está criando.
  - `role_arn`— O Amazon Resource Name (ARN) de uma função que pode executar Amazon Kendra APIs. Para obter mais informações, consulte Funções do 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 criar um índice

- Forneça valores para as seguintes variáveis no exemplo de código a seguir:
  - `description`: uma descrição da índice que você está criando. Isso é opcional.
  - `index_name`: o nome da índice que você está criando.
  - `role_arn`— O Amazon Resource Name (ARN) de uma função que pode executar Amazon Kendra APIs. Para obter mais informações, consulte Funções do 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.");
    }
}
```

Depois de criar o índice, você adiciona documentos a ele. Você pode adicioná-los diretamente ou criar uma fonte de dados que atualize o índice regularmente.

Tópicos

- Como adicionar documentos diretamente a um índice com o upload em lote.
- Adicionar perguntas frequentes a um índice
- Criação de campos de documentos personalizados
- Controle do acesso de usuários a documentos por token

Como adicionar documentos diretamente a um índice com o upload em lote.

Você pode adicionar documentos diretamente a um índice usando a BatchPutDocumentAPI. Você não pode excluir documentos diretamente usando o console. Se você usa o console, você se conecta a uma fonte de dados para adicionar documentos ao seu índice. Os documentos podem ser adicionados de um bucket do S3 ou fornecidos como dados binários. Para obter uma lista dos tipos de documentos suportados pelo, Amazon Kendra consulte Tipos de documentos.

A adição de documentos a um índice usando `BatchPutDocument` é uma operação assíncrona. Depois de chamar a `BatchPutDocument` API, você usa a BatchGetDocumentStatusAPI para monitorar o progresso da indexação de seus documentos. Quando você chama a API `BatchGetDocumentStatus` com uma lista de IDs de documentos, ela retorna o status do documento. Quando o status do documento é `INDEXED` ou `FAILED`, o processamento do documento está concluído. Quando o status é `FAILED`, a API `BatchGetDocumentStatus` retorna o motivo pelo qual o documento não pôde ser indexado.

Se você quiser alterar os metadados ou os atributos e o conteúdo do documento durante o processo de absorção do documento, consulte Enriquecimento personalizado de documentos no Amazon Kendra. Se você quiser usar uma fonte de dados personalizada, cada documento enviado usando a API `BatchPutDocument` exige uma ID da fonte de dados e uma ID de execução como atributos ou campos. Para obter mais informações, consulte Atributos obrigatórios para fontes de dados personalizadas.

> **Note**
>
> Cada ID de documento deve ser exclusiva por índice. Você não pode criar uma fonte de dados para indexar os documentos com os IDs exclusivos e depois usar a API `BatchPutDocument` para indexar os mesmos documentos ou vice-versa. Você pode criar uma fonte de dados e depois usar a API `BatchPutDocument` para indexar os mesmos documentos ou vice-versa. Usar as APIs `BatchPutDocument` e `BatchDeleteDocument` em combinação com um conector de fonte de dados do Amazon Kendra para o mesmo conjunto de documentos pode causar inconsistências entre os dados. Em vez disso, recomendamos usar o conector de fonte de dados personalizado do Amazon Kendra.

Os seguintes documentos do guia do desenvolvedor mostram como adicionar documentos diretamente a um índice.

Tópicos

- Adicionar documentos com a BatchPutDocument API
- Adicionar documentos de um bucket do S3

Adicionar documentos com a BatchPutDocument API

O exemplo a seguir adiciona uma bolha de texto a um índice BatchPutDocumentchamando. Você pode usar a `BatchPutDocument` API para adicionar documentos diretamente ao seu índice. Para obter uma lista dos tipos de documentos suportados pelo, Amazon Kendra consulte Tipos de documentos.

Para ver um exemplo de criação de um índice usando os SDKs AWS CLI e, consulte Criação de um índice. Para configurar a CLI e os SDKs, consulte Configurando Amazon Kendra.

> **Note**
>
> Os arquivos adicionados ao índice devem estar em um fluxo de bytes codificado UTF-8.

Nos exemplos a seguir, o texto codificado em UTF-8 é adicionado ao índice.

CLI

No AWS Command Line Interface, use o comando a seguir. O comando a seguir é formatado para Linux e macOS. Para Windows, substitua o caractere de continuação de linha do Unix (\) pelo circunflexo (^).

```
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));
    }
}
```

Adicionar documentos de um bucket do S3

Você pode adicionar documentos diretamente ao seu índice a partir de um Amazon S3 bucket usando a BatchPutDocumentAPI. Você pode inserir até 10 documentos na mesma chamada. Ao usar um bucket do S3, você deve fornecer uma IAM função com permissão para acessar o bucket que contém seus documentos. Você especifica a função com o parâmetro `RoleArn`.

Usar a BatchPutDocumentAPI para adicionar documentos de um Amazon S3 bucket é uma operação única. Para manter um índice sincronizado com o conteúdo de um bucket, crie uma fonte de Amazon S3 dados. Para obter mais informações, consulte fonte de dados do Amazon S3.

Para ver um exemplo de criação de um índice usando os SDKs AWS CLI e, consulte Criação de um índice. Para configurar a CLI e os SDKs, consulte Configurando Amazon Kendra. Para obter informações sobre como criar um bucket do S3, consulte a documentação do Amazon Simple Storage Service.

No exemplo a seguir, dois documentos do Microsoft Word são adicionados ao índice usando a 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));
    }
}
```

Adicionar perguntas frequentes a um índice

Você pode adicionar perguntas frequentes (FAQs) diretamente ao seu índice usando o console ou a CreateFaqAPI. A adição de perguntas frequentes a um índice é uma operação assíncrona. Você coloca os dados das perguntas frequentes em um arquivo que você armazena em um Amazon Simple Storage Service bucket. Você pode usar arquivos CSV ou JSON como entrada para as perguntas frequentes:

- CSV básico: um arquivo CSV em que cada linha contém uma pergunta, uma resposta e um URI de origem opcional.
- CSV personalizado: um arquivo CSV que contém perguntas, respostas e cabeçalhos para campos/atributos personalizados que você pode usar para facetar, exibir ou classificar respostas de perguntas frequentes. Você também pode definir campos de controle de acesso para limitar a resposta das perguntas frequentes a determinados usuários e grupos que têm permissão para ver a resposta das perguntas frequentes.
- JSON: um arquivo JSON que contém perguntas, respostas e cabeçalhos para campos/atributos personalizados que você pode usar para facetar, exibir ou classificar respostas de perguntas frequentes. Você também pode definir campos de controle de acesso para limitar a resposta das perguntas frequentes a determinados usuários e grupos que têm permissão para ver a resposta das perguntas frequentes.

Por exemplo, o seguinte é um arquivo CSV básico que fornece respostas a perguntas sobre clínicas gratuitas em Spokane, Washington, EUA, e Mountain View, Missouri e EUA.

```
How many free clinics are in Spokane WA?, 13
```

```
How many free clinics are there in Mountain View Missouri?, 7
```

> **Note**
>
> O arquivo de perguntas frequentes deve ser um arquivo com codificação UTF-8.

Tópicos

- Criação de campos de índice para um arquivo de perguntas frequentes
- Arquivo CSV básico
- Arquivo CSV personalizado
- Arquivo JSON
- Usando seu arquivo de perguntas frequentes
- Arquivos de perguntas frequentes em idiomas diferentes do inglês

Criação de campos de índice para um arquivo de perguntas frequentes

Ao usar um arquivo CSV ou JSON personalizado para entrada, você pode declarar campos personalizados para suas perguntas de perguntas frequentes. Por exemplo, você pode criar um campo personalizado que atribua cada pergunta a um departamento comercial. Quando as perguntas frequentes são retornadas em uma resposta, você pode usar o departamento como uma faceta para restringir a pesquisa apenas a “RH” ou “Finanças”, por exemplo.

Um campo personalizado deve ser mapeado para um campo de índice. No console, você usa a página de definição de facetas para criar um campo de índice. Ao usar a API, você deve primeiro criar um campo de índice usando a UpdateIndexAPI.

O tipo de campo/atributo no arquivo de perguntas frequentes deve corresponder ao tipo do campo de índice associado. Por exemplo, o campo “Departamento” é um campo de tipo `STRING_LIST`. Portanto, você deve fornecer valores para o campo do departamento como uma lista de strings no arquivo de perguntas frequentes. Você pode verificar o tipo de campos de índice usando a página de definição de facetas no console ou usando a DescribeIndexAPI.

Ao criar um campo de índice mapeado para um atributo personalizado, você pode marcá-lo como exibível, facetável ou classificável. Não é possível fazer um atributo personalizado pesquisável.

Além dos atributos personalizados, você também pode usar os campos Amazon Kendra reservados ou comuns em um arquivo CSV ou JSON personalizado. Para obter mais informações, consulte Campos ou atributos personalizados.

Arquivo CSV básico

Use um arquivo CSV básico quando quiser usar uma estrutura simples para suas perguntas frequentes. Em um arquivo CSV básico, cada linha tem dois ou três campos: uma pergunta, uma resposta e um URI de origem opcional que aponta para um documento com mais informações.

O conteúdo do arquivo deve seguir o formato comum RFC 4180 e o tipo MIME para arquivos de valores separados por vírgula (CSV).

A seguir está um arquivo de perguntas frequentes no 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
```

Arquivo CSV personalizado

Use um arquivo CSV personalizado quando quiser adicionar campos/atributos personalizados às suas perguntas de perguntas frequentes. Para um arquivo CSV personalizado, você usa uma linha de cabeçalho no arquivo CSV para definir os atributos adicionais.

O arquivo CSV deve conter os dois campos obrigatórios a seguir:

- `_question`: perguntas frequentes
- `_answer`: a resposta para as perguntas frequentes

Seu arquivo pode conter campos Amazon Kendra reservados e campos personalizados. Este é um exemplo de um arquivo .

```
_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
```

O conteúdo do arquivo deve seguir o formato comum RFC 4180 e o tipo MIME para arquivos de valores separados por vírgula (CSV).

A seguir, listamos os tipos de campos personalizados:

- Data: valores de data e hora codificados ISO 8601.

  Por exemplo, 2012-03-25T12:30:10+01:00 é o formato de data e hora do ISO 8601 para 25 de março de 2012 às 12h30 (mais 10 segundos) no horário da Europa Central.
- Longos: números, como `1234`.
- String: valores de string. Se a string contiver vírgulas, coloque o valor inteiro entre aspas duplas (“) (por exemplo, `"custom attribute, and more"`).
- Lista de strings: uma lista de valores de string. Liste os valores em uma lista separada por virgulas que está entre aspas (“) (por exemplo, `"item1, item2, item3"`). Se a lista contiver somente uma única entrada, você poderá omitir as aspas (por exemplo, `item1`).

Um arquivo CSV personalizado pode conter campos de controle de acesso do usuário. Você pode usar esses campos para limitar o acesso às perguntas frequentes a determinados usuários e grupos. Para filtrar o contexto do usuário, o usuário deve fornecer informações do usuário e do grupo na consulta. Caso contrário, todas as perguntas frequentes relevantes serão retornadas. Para obter mais informações, consulte Filtrar o contexto do usuário.

A seguir, são listados os filtros de contexto do usuário para perguntas frequentes:

- `_acl_user_allow`: os usuários na lista de permissões podem ver as perguntas frequentes na resposta da consulta. A pergunta não é devolvida a outros usuários.
- `_acl_user_deny`: os usuários na lista de bloqueios não podem ver as perguntas frequentes na resposta da consulta. A pergunta é devolvida a todos os outros usuários quando é relevante para a consulta.
- `_acl_group_allow`: os usuários que são membros de um grupo permitido podem ver as perguntas frequentes na resposta da consulta. As perguntas frequentes não são devolvidas aos usuários que são membros de outro grupo.
- `_acl_group_deny`: os usuários que são membros de um grupo permitido podem ver as perguntas frequentes na resposta da consulta. A pergunta é devolvida a todos os outros usuários quando é relevante para a consulta.

Forneça os valores para as listas de permissão e negação em listas separadas por vírgulas entre aspas (por exemplo, `"user1,user2,user3"`). Você pode incluir um usuário ou um grupo em uma lista de permissões ou negações, mas não em ambas, quando o mesmo usuário é permitido individualmente, mas também em um grupo negado. Se você incluir um usuário ou grupo em ambos, receberá um erro.

Este é um exemplo de um arquivo .

```
_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
```

Arquivo JSON

Você pode usar um arquivo JSON para fornecer perguntas, respostas e campos para seu índice. Você pode adicionar qualquer um dos campos Amazon Kendra reservados ou personalizados às perguntas frequentes.

O seguinte é o esquema para o arquivo 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
    ]
```

```
}
```

O exemplo de arquivo JSON a seguir mostra dois documentos de perguntas frequentes. Um dos documentos contém apenas as perguntas e respostas necessárias. O outro documento também inclui informações adicionais de campo e contexto do usuário ou controle de acesso.

```
{
    "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 seguir, listamos os tipos de campos personalizados:

- Data: um valor de string JSON com valores de data e hora codificados em ISO 8601. Por exemplo, 2012-03-25T12:30:10+01:00 é o formato de data e hora do ISO 8601 para 25 de março de 2012 às 12h30 (mais 10 segundos) no horário da Europa Central.

- Longo: um valor numérico JSON, como `1234`.
- String: um valor de string JSON (por exemplo, `"custom attribute"`).
- Lista de strins: uma matriz JSON de valores de string (por exemplo, `["item1,item2,item3"]`).

Um arquivo JSON pode conter campos de controle de acesso do usuário. Você pode usar esses campos para limitar o acesso às perguntas frequentes a determinados usuários e grupos. Para filtrar o contexto do usuário, o usuário deve fornecer informações do usuário e do grupo na consulta. Caso contrário, todas as perguntas frequentes relevantes serão retornadas. Para obter mais informações, consulte Filtrar o contexto do usuário.

Você pode incluir um usuário ou um grupo em uma lista de permissões ou negações, mas não em ambas, quando o mesmo usuário é permitido individualmente, mas também em um grupo negado. Se você incluir um usuário ou grupo em ambos, receberá um erro.

Veja a seguir um exemplo de como incluir controle de acesso do usuário em uma pergunta frequente JSON.

```
"AccessControlList": [
                {
                    "Name": "group or user name",
                    "Type": "GROUP | USER",
                    "Access": "ALLOW | DENY"
                },
                additional user context
            ]
```

Usando seu arquivo de perguntas frequentes

Depois de armazenar seu arquivo de entrada de perguntas frequentes em um bucket do S3, você use console ou a API `CreateFaq`para colocar as perguntas e respostas no índice. Se você quiser atualizar uma pergunta frequente, exclua ela e crie-a novamente. Use a API `DeleteFaq` para excluir uma FAQ.

Você deve fornecer uma IAM função que tenha acesso ao bucket do S3 que contém seus arquivos de origem. Você especifica a função no console ou no parâmetro `RoleArn`. Veja a seguir um exemplo de como adicionar um arquivo de pergunta frequente a um 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));

    }
}
```

Arquivos de perguntas frequentes em idiomas diferentes do inglês

Você pode indexar uma FAQ em um idioma compatível. Amazon Kendra indexa as perguntas frequentes em inglês por padrão se você não especificar um idioma. Você especifica o código do idioma ao chamar a CreateFaqoperação ou pode incluir o código do idioma de uma FAQ nos metadados da FAQ como um campo. Se uma pergunta frequente não tiver um código de idioma em seus metadados especificado em um campo de metadados, a pergunta frequente será indexada usando o código de idioma especificado quando você chama a operação `CreateFAQ`. Para indexar um documento de perguntas frequentes em um idioma compatível no console, acesse Perguntas frequentes e selecione Adicionar perguntas frequentes. Você escolhe um idioma no menu suspenso Idioma.

Criação de campos de documentos personalizados

Você pode criar atributos ou campos personalizados para seus documentos no seu índice Amazon Kendra. Por exemplo, você pode criar um campo ou atributo personalizado chamado “Departamento” com os valores de “RH”, “Vendas” e “Fabricação”. Se você mapear esses campos ou atributos personalizados para o seu índice Amazon Kendra, você pode usá-los para filtrar os resultados da pesquisa e incluir documentos pelo atributo de departamento “RH”, por exemplo.

Antes de usar um campo ou atributo personalizado, você deve primeiramente criar o campo no índice. Use o console para editar os mapeamentos de campo da fonte de dados para adicionar um campo personalizado ou use a UpdateIndexAPI para criar o campo de índice. Você não pode alterar o tipo de campo depois de criar o campo.

Para a maioria das fontes de dados, você mapeia campos na fonte de dados externa para os campos correspondentes em Amazon Kendra. Para obter mais informações, consulte Mapear campos de fonte de dados. Para origens dos dados do S3, você pode criar campos ou atributos personalizados usando um arquivo de metadados JSON.

Crie até 500 campos ou atributos personalizados.

Você também pode usar campos Amazon Kendra reservados ou comuns. Para obter mais informações, consulte Campos ou atributos personalizados.

Tópicos

- Atualização de campos de documentos personalizados

Atualização de campos de documentos personalizados

Com a API `UpdateIndex`, você adiciona campos ou atributos personalizados usando o parâmetro `DocumentMetadataConfigurationUpdates`.

O exemplo de JSON a seguir usa `DocumentMetadataConfigurationUpdates` para adicionar um campo chamado “Departamento” ao índice.

```
"DocumentmetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE"
   }
]
```

As seções a seguir incluem exemplos para adicionar atributos ou campos personalizados usando BatchPutDocumente para uma fonte de dados do Amazon S3.

Tópicos

- Adicionar atributos ou campos personalizados com a BatchPutDocument API

- Adicionar atributos ou campos personalizados a uma fonte de dados do Amazon S3

Adicionar atributos ou campos personalizados com a BatchPutDocument API

Ao usar a BatchPutDocumentAPI para adicionar um documento ao seu índice, você especifica campos ou atributos personalizados como parte do`Attributes`. Você pode adicionar vários campos ou atributos ao chamar a API. Crie até 500 campos ou atributos personalizados. O exemplo a seguir é um campo ou atributo personalizado que adiciona “Departamento” a um documento.

```
"Attributes":
    {
        "Department": "HR",
        "_category": "Vacation policy"
    }
```

Adicionar atributos ou campos personalizados a uma fonte de dados do Amazon S3

Ao usar um bucket do S3 como fonte de dados para o índice, você adiciona metadados aos documentos com arquivos de metadados complementares. Você coloca os arquivos JSON de metadados em uma estrutura de diretórios paralela aos documentos. Para mais informações, consulte metadados de documento do S3.

Você especifica campos ou atributos personalizados na estrutura `Attributes` JSON. Crie até 500 campos ou atributos personalizados. Por exemplo, o exemplo a seguir usa `Attributes` para definir três campos ou atributos personalizados e um campo reservado.

```
"Attributes": {
        "brand": "Amazon Basics",
        "price": 1595,
        "_category": "sports",
        "subcategories": ["outdoors", "electronics"]
    }
```

As etapas a seguir orientam você a adicionar atributos personalizados a uma fonte de dados do Amazon S3.

Tópicos

- Etapa 1: criar um índice do Amazon Kendra
- Etapa 2: atualizar o índice para adicionar campos de documentos personalizados

- Etapa 3: Crie uma fonte de dados do Amazon S3 e mapeie os campos da fonte de dados para atributos personalizados

Etapa 1: criar um índice do Amazon Kendra

Siga as etapas Criar um índice para criar seu índice Amazon Kendra.

Etapa 2: atualizar o índice para adicionar campos de documentos personalizados

Depois de criar um índice, você adiciona campos a ele. O procedimento a seguir mostra como adicionar campos a um índice usando o console e a CLI.

Console

Para criar campos de índice

1. Verifique se você criou um índice.
2. Em seguida, no menu de navegação à esquerda, em Gerenciamento de dados, escolha Definição de faceta.
3. No Guia de configurações do campo Índice, em Campos de índice, escolha Adicionar campo para adicionar campos personalizados.
4. Na caixa de diálogo Adicionar campo de índice, faça o seguinte:
   - Nome do campo — Adicione um nome de campo.
   - Tipo de dados — Selecione o tipo de dados, seja String, String list ou Data.
   - Tipos de uso — Selecione os tipos de uso, sejam eles facetáveis, pesquisáveis, exibíveis e classificáveis.

     Em seguida, selecione Adicionar.

   Repita a última etapa para qualquer outro campo que você queira 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
        }
    }
...
]"
```

Etapa 3: Crie uma fonte de dados do Amazon S3 e mapeie os campos da fonte de dados para atributos personalizados

Para criar uma fonte de dados do Amazon S3 e mapear campos para ela, siga as instruções em. Amazon S3

Se você estiver usando a API, use o `fieldMappings` atributo abaixo `configuration` ao usar a CreateDataSourceAPI.

Para obter uma visão geral de como os campos da fonte de dados são mapeados, consulteMapeando campos de fontes de dados.

Controle do acesso de usuários a documentos por token

Você pode controlar quais usuários ou grupos podem acessar determinados documentos no índice ou ver determinados documentos noa resultados de pesquisa. Isso é chamado de filtragem de contexto do usuário. É uma espécie de pesquisa personalizada com o benefício de controlar o

acesso aos documentos. Por exemplo, nem todas as equipes que pesquisam informações no portal da empresa devem acessar documentos ultrassecretos da empresa, nem esses documentos são relevantes para todos os usuários. Somente usuários específicos ou grupos de equipes com acesso a documentos ultrassecretos devem ver esses documentos nos resultados de pesquisa.

Amazon Kendra oferece suporte ao controle de acesso de usuário baseado em token usando os seguintes tipos de token:

- Open ID
- JWT com senha compartilhada
- JWT com chave pública
- JSON

O Amazon Kendra oferece pesquisa corporativa altamente segura para os aplicativos de pesquisa. Os resultados da pesquisa refletem o modelo de segurança da organização. Os clientes são responsáveis por autenticar e autorizar os usuários a obter acesso ao aplicativo de pesquisa. No momento da pesquisa, o serviço do Amazon Kendra filtra os resultados da pesquisa com base no ID do usuário fornecido pelo aplicativo de pesquisa do cliente e nas listas de controle de acesso a documentos (ACLs) coletadas pelos conectores do Amazon Kendra durante o tempo de rastreamento/indexação. Os resultados da pesquisa retornam URLs que apontam para os repositórios de documentos originais, além de pequenos trechos. O acesso ao documento completo ainda é imposto pelo repositório original.

Tópicos

- Usando o OpenID
- Como usar um JSON Web Token (JWT) com uma senha compartilhada
- Usando um JSON Web Token (JWT) com uma chave pública
- Usar o JSON

Usando o OpenID

Para configurar um Amazon Kendra índice para usar um token OpenID para controle de acesso, você precisa da URL JWKS (JSON Web Key Set) do provedor OpenID. Na maioria dos casos, a URL do JWKS está no seguinte formato (se elas estiverem seguindo a descoberta do OpenID) `https://domain-name/.well_known/jwks.json`:

Os exemplos a seguir mostram como usar um token do OpenID para controle de acesso do usuário ao criar um índice.

Console

1. Escolha Criar índice para criar um novo índice.
2. Na página Especificar detalhes do índice, dê um nome e uma descrição ao índice.
3. Para a função do IAM , selecione uma função ou Criar uma nova função para e especifique um nome de função para criar uma nova função. A função do IAM terá o prefixo "AmazonKendra-”.
4. Deixe os outros campos nos padrões determinados. Escolha Próximo.
5. Na página Configurar controle de acesso do usuário, em Configurações de controle de acesso, escolha Sim para usar tokens para controle de acesso.
6. Em Configuração de token, selecione OpenID como o Tipo de token.
7. Especifique a URL da chave de assinatura. A URL deve apontar para um conjunto de chaves da web JSON.
8. Opcional Em Configuração avançada:
    a. Especifique um Nome de usuário para usar na verificação da ACL.
    b. Especifique um ou mais Grupos para serem usados na verificação da ACL.
    c. Especifique o Emissor que validará o emissor do token.
    d. Especifique o(s) ID(s) do cliente. Você deve especificar uma expressão regular que corresponda ao público no JWT.
9. Na página de Detalhes do provisionamento, escolha Developer Edition.
10. Escolha Criar para criar seu índice.
11. Aguarde até que seu índice seja criado. Amazon Kendra provisiona o hardware para seu índice. Essa operação pode levar algum tempo.

CLI

Para criar um índice AWS CLI usando um arquivo de entrada JSON, primeiro crie um arquivo JSON com os parâmetros desejados:

```
{
    "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"
}
```

Você pode substituir os nomes de campo padrão do usuário e do grupo. O valor padrão para `UserNameAttributeField` é "usuário". O valor padrão para `GroupAttributeField` é "grupos".

Em seguida, chame o `create-index` usando o arquivo de entrada. Por exemplo, se o nome do arquivo JSON for `create-index-openid.json`, você poderá usar o seguinte:

```
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'
)
```

Como usar um JSON Web Token (JWT) com uma senha compartilhada

Os exemplos a seguir mostram como usar o JSON Web Token (JWT) com um token de segredo compartilhado para controle de acesso do usuário ao criar um índice.

Console

1. Escolha Criar índice para começar a criar um novo índice.
2. Na página Especificar detalhes do índice, dê um nome e uma descrição ao índice.
3. Para o perfil do IAM, selecione uma função ou selecione Criar uma nova função para e especifique um nome de função para criar uma nova função. A IAM função terá o prefixo "AmazonKendra-”.
4. Deixe os outros campos nos padrões determinados. Escolha Próximo.
5. Na página Configurar controle de acesso do usuário, em Configurações de controle de acesso, escolha Sim para usar tokens para controle de acesso.
6. Em Configuração de token, selecione JWT com senha compartilhada como o Tipo de token.
7. Em Parâmetros para entrar com senha compartilhada, escolha o Tipo de senha. Você pode usar uma senha compartilhada do AWS Secrets Manager existente ou criar uma nova senha compartilhada.

   Para criar uma nova senha compartilhada, escolha Novo e siga estas etapas:

   a. Em Novo AWS Secrets Manager segredo, especifique um nome secreto. O prefixo `AmazonKendra-` será adicionado ao salvar a chave pública.
   b. Especifique um ID de chave. O ID de chave é uma dica que indica qual foi a chave usada para proteger a JSON Web Signature (JWS) do token.
   c. Escolha o Algoritmo de assinatura para o token. Esse é o algoritmo criptográfico usado para proteger o token de ID. Para obter mais informações sobre o RSA, consulte a Criptografia RSA.
   d. Especifique uma senha compartilhada inserindo uma senha codificada em URL base64. Você também pode selecionar Gerar senha para que uma senha seja gerada para você. Você deve garantir que ela seja uma senha codificada em URL base64.

e. (Opcional) Especifique quando a senha compartilhada for válida. Você pode especificar a data e a hora a partir da qual uma senha é válida, válido até ou os dois. A senha será válida no intervalo especificado.

f. Selecione Salvar senha para salvar a nova senha.

8. (Opcional) Em Configuração avançada:

a. Especifique um Nome de usuário para usar na verificação da ACL.

b. Especifique um ou mais Grupos para serem usados na verificação da ACL.

c. Especifique o Emissor que validará o emissor do token.

d. Especifique o(s) ID(s) da reclamação. Especifique uma expressão regular que corresponda ao público no JWT.

9. Na página de Detalhes do provisionamento, escolha Developer Edition.

10. Escolha Criar para criar seu índice.

11. Aguarde até que seu índice seja criado. Amazon Kendra provisiona o hardware para seu índice. Essa operação pode levar algum tempo.

CLI

Você pode usar o token JWT com um segredo compartilhado dentro do AWS Secrets Manager. A senha deve ser um senha codificada com uma URL base64. Você precisa do Secrets Manager ARN e sua Amazon Kendra função deve ter acesso ao `GetSecretValue` Secrets Manager recurso. Se você estiver criptografando o Secrets Manager recurso com AWS KMS, a função também deverá ter acesso à ação de descriptografia.

Para criar um índice AWS CLI usando um arquivo de entrada JSON, primeiro crie um arquivo JSON com os parâmetros desejados:

```
{
    "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"
}
```

Você pode substituir os nomes de campo padrão do usuário e do grupo. O valor padrão para `UserNameAttributeField` é "usuário". O valor padrão para `GroupAttributeField` é "grupos".

Em seguida, chame o `create-index` usando o arquivo de entrada. Por exemplo, se o nome do arquivo JSON for `create-index-openid.json`, use o seguinte:

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

O segredo deve ter o seguinte formato em 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 obter mais informações sobre o JWT, consulte jwt.io.

Python

Você pode usar o token JWT com um segredo compartilhado dentro do AWS Secrets Manager. A senha deve ser um senha codificada com uma URL base64. Você precisa do Secrets Manager ARN e sua Amazon Kendra função deve ter acesso ao `GetSecretValue` Secrets Manager

recurso. Se você estiver criptografando o Secrets Manager recurso com AWS KMS, a função também deverá ter acesso à ação de descriptografia.

```
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'
)
```

Usando um JSON Web Token (JWT) com uma chave pública

Os exemplos a seguir mostram como usar o JSON Web Token (JWT) com uma chave pública para controle de acesso do usuário ao criar um índice. Para obter mais informações sobre o JWT, consulte jwt.io.

Console

1. Escolha Criar índice para começar a criar um novo índice.
2. Na página Especificar detalhes do índice, dê um nome e uma descrição ao índice.
3. Para o perfil do IAM, selecione uma função ou selecione Criar uma nova função para e especifique um nome de função para criar uma nova função. A IAM função terá o prefixo "AmazonKendra-”.
4. Deixe os outros campos nos padrões determinados. Escolha Próximo.
5. Na página Configurar controle de acesso do usuário, em Configurações de controle de acesso, escolha Sim para usar tokens para controle de acesso.

6. Em Configuração de token, selecione JWT com chave pública como o Tipo de token.
7. Em Parâmetros para assinar a chave pública, escolha o Tipo de senha. Você pode usar uma senha existente do AWS Secrets Manager ou criar outra.

   Para criar uma nova senha, escolha Novo e siga estas etapas:

   a. Em Novo AWS Secrets Manager segredo, especifique um nome secreto. O prefixo `AmazonKendra-` será adicionado ao salvar a chave pública.
   b. Especifique um ID de chave. O ID de chave é uma dica que indica qual foi a chave usada para proteger a JSON Web Signature (JWS) do token.
   c. Escolha o Algoritmo de assinatura para o token. Esse é o algoritmo criptográfico usado para proteger o token de ID. Para obter mais informações sobre o RSA, consulte a Criptografia RSA.
   d. Em Atributos do certificado, especifique uma Cadeia de certificados opcional. A cadeia de certificados é composta por uma lista de certificados. Ele começa com o certificado do servidor e termina com o certificado raiz.
   e. Opcional Especifique a impressão digital ou do polegar. Ele deve ser uma confirmação de certificado, uma verificação além de todos os dados do certificado e sua assinatura.
   f. Especifique o Expoente. Esse é o valor expoente da chave pública RSA. Ele é representado como um valor codificado como Base64urlUInt.
   g. Especifique os Módulos. Esse é o valor expoente da chave pública RSA. Ele é representado como um valor codificado como Base64urlUInt.
   h. Selecione Salvar chave para salvar a nova chave.
8. Opcional Na Configuração avançada:

   a. Especifique um Nome de usuário para usar na verificação da ACL.
   b. Especifique um ou mais Grupos para serem usados na verificação da ACL.
   c. Especifique o Emissor que validará o emissor do token.
   d. Especifique o(s) ID(s) do cliente. Você deve especificar uma expressão regular que corresponda ao público no JWT.
9. Na página de Detalhes do provisionamento, escolha Developer Edition.
10. Escolha Criar para criar seu índice.
11. Aguarde até que seu índice seja criado. Amazon Kendra provisiona o hardware para seu índice. Essa operação pode levar algum tempo.

CLI

Use o JWT com uma chave pública dentro de um AWS Secrets Manager. Você precisa do Secrets Manager ARN e sua Amazon Kendra função deve ter acesso ao `GetSecretValue` Secrets Manager recurso. Se você estiver criptografando o Secrets Manager recurso com AWS KMS, a função também deverá ter acesso à ação de descriptografia.

Para criar um índice AWS CLI usando um arquivo de entrada JSON, primeiro crie um arquivo JSON com os parâmetros desejados:

```
{
    "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"
}
```

Você pode substituir os nomes de campo padrão do usuário e do grupo. O valor padrão para `UserNameAttributeField` é "usuário". O valor padrão para `GroupAttributeField` é "grupos".

Em seguida, chame o `create-index` usando o arquivo de entrada. Por exemplo, se o nome do arquivo JSON for `create-index-openid.json`, use o seguinte:

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

O segredo deve ter o seguinte formato em 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 obter mais informações sobre o 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'
)
```

Usar o JSON

Os exemplos a seguir mostram como usar o JSON para controle de acesso do usuário ao criar um índice.

> **Warning**
>
> O token JSON é uma carga não validada. Ele só deve ser usado quando as solicitações de Amazon Kendra vêm de um servidor confiável e nunca de um navegador.

Console

1. Escolha Criar índice para começar a criar um novo índice.
2. Na página Especificar detalhes do índice, dê um nome e uma descrição ao índice.
3. Para a função do IAM , selecione uma função ou Criar uma nova função para e especifique um nome de função para criar uma nova função. A IAM função terá o prefixo "AmazonKendra-”.
4. Deixe os outros campos nos padrões determinados. Escolha Próximo.
5. Na página Configurar controle de acesso do usuário, em Configurações de controle de acesso, escolha Sim para usar tokens o para controle de acesso.
6. Em Configuração de token, selecione JSON como o Tipo de token.
7. Especifique um Nome de usuário a ser usado na verificação da ACL.
8. Especifique um ou mais Grupos a serem usados na verificação da ACL.
9. Escolha Próximo.
10. Na página de Detalhes do provisionamento, escolha Developer Edition.
11. Escolha Criar para criar seu índice.
12. Aguarde até que seu índice seja criado. Amazon Kendra provisiona o hardware para seu índice. Essa operação pode levar algum tempo.

CLI

Para criar um índice AWS CLI usando um arquivo de entrada JSON, primeiro crie um arquivo JSON com os parâmetros desejados:

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",
    "UserTokenConfigurations": [
        {
            "JsonTokenTypeConfiguration": {
                "UserNameAttributeField": "user",
                "GroupAttributeField": "group"
            }
        }
    ],
    "UserContextPolicy": "USER_TOKEN"
}
```

Em seguida, chame o `create-index` usando o arquivo de entrada. Por exemplo, se o nome do arquivo JSON for `create-index-openid.json`, use o seguinte:

```
aws kendra create-index --cli-input-json file://create-index-openid.json
```

Se você não estiver usando o Open ID para AWS IAM Identity Center, poderá nos enviar o token no formato JSON. Se você fizer isso, deverá especificar qual campo no token JSON contém o nome do usuário e os grupos. Os valores do campo de grupos devem ser uma matriz de sequências de caracteres JSON. Por exemplo, se estiver usando SAML, o token será semelhante ao seguinte:

```
{
     "username" : "user1",
     "groups": [
        "group1",
        "group2"
     ]
}
```

O `TokenConfiguration` especificaria o nome do usuário e os nomes dos campos do grupo:

```
{
    "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'
)
```

Criar um conector de fonte de dados

Você pode criar um conector de fonte de dados Amazon Kendra para se conectar e indexar seus documentos. Amazon Kendra pode se conectar à Microsoft SharePoint, ao Google Drive e a muitos outros provedores. Ao criar um conector de fonte de dados, você fornece Amazon Kendra as informações de configuração necessárias para se conectar ao seu repositório de origem. Ao contrário de adicionar documentos diretamente em um índice, digitalize periodicamente a fonte de dados para atualizar o índice.

Por exemplo, digamos que você tenha um repositório de documentos fiscais armazenados em um Amazon S3 bucket. De tempos em tempos, os documentos existentes são alterados e novos documentos são adicionados ao repositório. Se você adicionar o repositório Amazon Kendra como fonte de dados, poderá manter seu índice atualizado configurando sincronizações periódicas entre a fonte de dados e o índice.

Você pode optar por atualizar um índice manualmente usando o console ou a StartDataSourceSyncJobAPI. Caso contrário, configure uma agenda para atualizar um índice e sincronizá-lo com a fonte de dados.

Um índice pode ter mais de uma fonte de dados. Cada fonte de dados pode ter seu próprio cronograma de atualização. Por exemplo, atualize o índice dos documentos de trabalho diariamente, ou até mesmo de hora em hora, enquanto atualiza os documentos arquivados manualmente sempre que o arquivo for alterado.

Se você quiser alterar os metadados ou os atributos e o conteúdo do documento durante o processo de absorção do documento, consulte Enriquecimento personalizado de documentos no Amazon Kendra.

> **Note**
>
> Cada ID de documento deve ser exclusiva por índice. Você não pode criar uma fonte de dados para indexar os documentos com os IDs exclusivos e depois usar a API `BatchPutDocument` para indexar os mesmos documentos ou vice-versa. Você pode criar uma fonte de dados e depois usar a API `BatchPutDocument` para indexar os mesmos documentos ou vice-versa. Usar as `BatchDeleteDocument` APIs `BatchPutDocument` e em combinação com um conector de fonte de Amazon Kendra dados para o mesmo conjunto

de documentos pode causar inconsistências com seus dados. Em vez disso, recomendamos usar o conector de fonte de dados personalizado do Amazon Kendra.

> **Note**
>
> Os arquivos adicionados ao índice devem estar em um fluxo de bytes codificado UTF-8. Para obter mais informações sobre documentos em Amazon Kendra, consulte Documentos.

Definindo um cronograma de atualização

Configure a fonte de dados para ser atualizada periodicamente com o console ou usando o parâmetro `Schedule` ao criar ou atualizar uma fonte de dados. O conteúdo do parâmetro é uma string que contém uma string de agendamento em formato `cron` ou uma string vazia para indicar que o índice é atualizado sob demanda. Para o formato de uma expressão cron, consulte Programar expressões para regras no Guia do Amazon CloudWatch Events usuário. Amazon Kendra suporta somente expressões cron. Ele não suporta expressões rate.

Configurações de idioma

Você pode indexar todos os documentos em uma fonte de dados em um idioma compatível. Você especifica o código do idioma para todos os seus documentos em sua fonte de dados ao ligar CreateDataSource. Se um documento não tiver um código de idioma especificado em um campo de metadados, o documento será indexado usando o código de idioma especificado para todos os documentos no nível da fonte de dados. Se você não especificar um idioma, o Amazon Kendra indexa documentos em uma fonte de dados em inglês por padrão. Para obter mais informações sobre os idiomas suportados, incluindo os códigos, consulte Adicionar documentos em outros idiomas além do inglês.

Você pode indexar todos os documentos em uma fonte de dados em um idioma compatível. Acesse Fontes de dados e edite a fonte de dados ou Adicione a fonte de dados se estiver adicionando uma nova fonte de dados. Na página Especificar detalhes da fonte de dados, escolha um idioma no menu suspenso Idioma. Selecione a opção Atualizar ou continue inserindo as informações de configuração para se conectar à sua fonte de dados.

Conectores de fontes de dados

Esta seção mostra como se conectar Amazon Kendra a bancos de dados e repositórios de fontes de dados compatíveis usando Amazon Kendra as Amazon Kendra APIs AWS Management Console e as.

Tópicos

- Esquemas de modelos de fonte de dados
- Adobe Experience Manager
- Alfresco
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (Windows)
- Amazon FSx (EM UM NetApp TOQUE)
- Amazon RDS/Aurora
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon S3
- Amazon Kendra Rastreador da Web
- Amazon WorkDocs
- Box
- Confluence
- Conectores de fontes de dados 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 modelos de fonte de dados

A seguir estão os esquemas de modelo para fontes de dados em que os modelos são compatíveis.

Tópicos

- Esquema de modelo do Adobe Experience Manager
- Amazon FSx Esquema de modelo (Windows)
- Amazon FSx Esquema de modelo (NetApp ONTAP)
- Esquema de modelo do Alfresco
- Aurora Esquema de modelo (MySQL)
- Aurora Esquema de modelo (PostgreSQL)
- Amazon RDS Esquema de modelo (Microsoft SQL Server)
- Amazon RDS Esquema de modelo (MySQL)
- Amazon RDS Esquema de modelo (Oracle)
- Amazon RDS Esquema de modelo (PostgreSQL)
- Amazon S3 esquema de modelo
- Amazon Kendra Esquema do modelo do Web Crawler

- Esquema do modelo do Confluence
- Esquema de modelos do Dropbox
- Esquema de modelos do Drupal
- GitHub esquema de modelo
- Esquema de modelos do Gmail
- Esquema do modelo do Google Drive
- Esquema de modelo do IBM DB2
- Esquema de modelo do Microsoft Exchange
- Esquema OneDrive de modelos da Microsoft
- Esquema SharePoint de modelos da Microsoft
- Esquema de modelo do Microsoft SQL Server
- Esquema de modelo do Microsoft Teams
- Esquema de modelo do Microsoft Yammer
- Esquema de modelo do MySQL
- Esquema de modelos do Oracle Database
- Esquema de modelo do (PostgreSQL
- Esquema de modelo do Salesforce
- ServiceNow esquema de modelo
- Esquema de modelos do Slack
- Esquema do modelo do Zendesk

Esquema de modelo do Adobe Experience Manager

Inclua um JSON que contém o esquema da fonte de dados como parte do objeto do TemplateConfiguration. Forneça a URL do host do Adobe Experience Manager, o tipo de autenticação e se você usa o Adobe Experience Manager (AEM) como um serviço de nuvem ou o AEM on-premises como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `AEM`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, especifique `TEMPLATE` como `Type` ao chamar CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Para ter mais informações, consulte Esquema JSON do Adobe Experience Manager.

A tabela a seguir descreve os parâmetros do esquema JSON do AEM.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
aemUrl	O URL do host do Adobe Experience Manager. Por exemplo, ao usar o AEM on-premises, inclua o nome do host e a porta:. https://hostname:port Ou, ao usar o AEM como um serviço de nuvem, use a URL do autor: https://author-xxxxxx-xxxxxxx.adobeaemcloud.com.
authType	O tipo de autenticação que você usa: `Basic` ou `OAuth2`.
deploymentType	O tipo de Adobe Experience Manager que você usa: `CLOUD` ou`ON_PREMISE` .
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• page • asset	Uma lista de objetos que mapeiam os atributos ou nomes de campo de suas Adobe Experience Manager páginas e ativos para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
timeZoneId	Se você usa o AEM On-Premise e o fuso horário do seu servidor é diferente do fuso

Configuração	Descrição
	horário do conector ou índice do Amazon Kendra AEM, você pode especificar o fuso horário do servidor para alinhar com o conector ou índice do AEM. O fuso horário padrão para o AEM On-Premis e é o fuso horário do conector ou índice do Amazon Kendra AEM. O fuso horário padrão para o AEM como serviço de nuvem é o Greenwich Mean Time.
• pageRootPaths • assetRootPaths	Uma lista de caminhos raiz para páginas e ativos. Por exemplo, o caminho raiz de uma página pode ser /content/sub e o caminho raiz de um ativo pode ser /content/sub/asset1.
crawlAssets	`true` para rastrear ativos.
crawlPages	`true` para rastrear páginas.
• pagePathInclusionPadrões • pageNameInclusionPadrões • assetPathInclusionPadrões • assetTypeInclusionPadrões • assetNameInclusionPadrões	Uma lista de padrões de expressões regulares para incluir determinadas páginas e ativos em sua fonte de dados do Adobe Experience Manager. As páginas e os ativos que correspon dem aos padrões são incluídos no índice. As páginas e os ativos que não correspondem aos padrões são excluídos do índice. Se a página ou ativo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o conteúdo não será incluído no índice.

Configuração	Descrição
• pagePathExclusionPadrões • pageNameExclusionPadrões • assetPathExclusionPadrões • assetTypeInclusionPadrões • assetNameInclusionPadrões	Uma lista de padrões de expressões regulares para incluir determinadas páginas e ativos em sua fonte de dados do Adobe Experience Manager. As páginas e os ativos que correspondem aos padrões são excluídos do índice. As páginas e os ativos que não correspondem aos padrões são incluídos no índice. Se a página ou ativo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o conteúdo não será incluído no índice.
pageComponents	Uma lista de nomes para os componentes de página específicos que você deseja indexar.
contentFragmentVariations	Uma lista de nomes para as variações salvas específicas dos fragmentos de conteúdo do Adobe Experience Manager que você deseja indexar.
tipo	O tipo da fonte de dados. Especifique `AEM` como seu tipo de fonte de dados.
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
secretArn	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Adobe Experience Manager. Para obter informações sobre esses pares de valores-c have, consulte Instruções de conexão para o Adobe Experience Manager.

Configuração	Descrição
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo (Windows)

Inclua um JSON que contém o esquema da fonte de dados como parte do objeto do TemplateConfiguration. Você fornece a ID do sistema de arquivos como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Você também deve especificar o tipo de fonte de dados comoFSX, um segredo para suas credenciais de autenticação e outras configurações necessárias. Em seguida, especifique `TEMPLATE` como `Type` ao chamar CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Amazon FSx Esquema JSON (Windows).

A tabela a seguir descreve os parâmetros do esquema JSON Amazon FSx (Windows).

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
fileSystemId	O identificador do sistema Amazon FSx de arquivos. Você pode encontrar o ID do sistema de arquivos no painel Sistemas de arquivos no Amazon FSx console.
fileSystemType	O tipo Amazon FSx de sistema de arquivos. Para usar Windows File Server como seu tipo de sistema de arquivos, especifique`WINDOWS`.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
Todos	Uma lista de objetos que mapeiam atributos ou nomes de campo de seus arquivos em sua fonte de Amazon FSx dados para Amazon

Configuração	Descrição
	Kendra indexar nomes de campo. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
isCrawlAcl	`true`para rastrear as informações da lista de controle de acesso (ACL) de seus documento s, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
inclusionPatterns	Uma lista de padrões de expressão regular para incluir determinados arquivos em sua fonte Amazon FSx de dados. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
exclusionPatterns	Uma lista de padrões de expressão regular para excluir determinados arquivos na sua fonte Amazon FSx de dados. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
tipo	O tipo da fonte de dados. Para fontes de dados do sistema de arquivos do Windows, especifiq ue`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 modelo (NetApp ONTAP)

Inclua um JSON que contém o esquema da fonte de dados como parte do objeto do TemplateConfiguration. Você fornece a ID do sistema de arquivos e a máquina virtual de armazenamento (SVM) como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Você também deve especificar o tipo de fonte de dados como`FSXONTAP`, um segredo para suas credenciais de autenticação e outras configurações necessárias. Em seguida, especifique `TEMPLATE` como `Type` ao chamar CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Amazon FSx Esquema NetApp JSON (ONTAP).

A tabela a seguir descreve os parâmetros do esquema JSON Amazon FSx (NetApp ONTAP).

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
fileSystemId	O identificador do sistema Amazon FSx de arquivos. Você pode encontrar o ID do sistema

Configuração	Descrição
	de arquivos no painel Sistemas de arquivos no Amazon FSx console. Para obter informações sobre como criar um sistema de arquivos no Amazon FSx console para o NetApp ONTAP, consulte o Guia de introdução do NetApp ONTAP no Guia do FSx for ONTAP usuário.
fileSystemType	O tipo Amazon FSx de sistema de arquivos. Para usar NetApp ONTAP como seu tipo de sistema de arquivos, especifique0NTAP.
SVMid	O identificador da máquina virtual de armazenamento (SVM) usada com seu sistema de Amazon FSx arquivos paraNetAp p ONTAP. Você pode encontrar sua ID SVM acessando o painel Sistemas de arquivos no Amazon FSx console, selecionando a ID do sistema de arquivos e, em seguida, seleciona ndo Máquinas virtuais de armazenamento. Para obter informações sobre como criar um sistema de arquivos no Amazon FSx console paraNetApp ONTAP, consulte o Guia de introdução do NetApp ONTAP no Guia do FSx for ONTAP usuário.
Tipo de protocolo	Se você usa o protocolo Common Internet File System (CIFS) para Windows ou o protocolo Network File System (NFS) para Linux.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.

Configuração	Descrição
file	Uma lista de objetos que mapeiam atributos ou nomes de campo de seus arquivos em sua fonte de Amazon FSx dados para Amazon Kendra indexar nomes de campo. Para obter mais informações, consulte Mapear campos de fonte de dados. Os nomes dos campos da fonte de dados devem existir nos metadados personalizados dos seus arquivos.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
crawlAcl	`true`para rastrear as informações da lista de controle de acesso (ACL) de seus documento s, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
inclusionPatterns	Uma lista de padrões de expressão regular para incluir determinados arquivos em sua fonte Amazon FSx de dados. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
exclusionPatterns	Uma lista de padrões de expressão regular para excluir determinados arquivos na sua fonte Amazon FSx de dados. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
tipo	O tipo da fonte de dados. Para fontes de dados do sistema de NetApp ONTAP arquivos, especifique`FSXONTAP`.
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição

secretArn	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu sistema de arquivos. Amazon FSx O segredo deve conter uma estrutura JSON com as seguintes chaves:
	`{ "username": " user@corp.example.com ", "password": " password" }`
	Se você usa o protocolo NFS para seu sistema de Amazon FSx arquivos, o segredo é armazenado em uma estrutura JSON com as seguintes chaves:
	`{ "leftId": "left ID", "rightId": " right ID", "preSharedKey": " pre-shared key " }`

Amazon FSx Esquema NetApp JSON (ONTAP)

```
{
  "$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 modelo do Alfresco

Inclua um JSON que contém o esquema da fonte de dados como parte do objeto do TemplateConfiguration. Forneça o ID do site, o URL do repositório, o URL da interface do usuário e o tipo de autenticação do Alfresco, se você usa a nuvem ou on-premises, e o tipo de conteúdo que deseja rastrear. Forneça isso como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `ALFRESCO`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, especifique `TEMPLATE` como `Type` ao chamar CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Alfresco.

A tabela a seguir descreve os parâmetros do esquema JSON do Alfresco.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
siteId	O identificador do site Alfresco.
repoUrl	O URL do seu repositório do Alfresco. Você pode obter o URL do repositório com o administrador do Alfresco. Por exemplo, se você usa o Cloud (PaaS) do Alfresco, o URL do repositório pode ser https://company.alfrescocloud.com. Ou, se você usa o Alfresco on-premises, o URL do repositório pode ser https://company-alfresco-instance.company-domain.suffix:port.
webAppUrl	O URL da sua interface de usuário do Alfresco. Você pode obter o URL da interface do usuário do Alfresco com o administrador do Alfresco. Por exemplo, o URL da interface do usuário pode ser https://example.com.

Configuração	Descrição
repositoryAdditionalProperties	Propriedades adicionais para se conectar ao endpoint do repositório/fonte de dados.
authType	O tipo de autenticação que você usa: `OAuth2` ou `Basic`.
tipo implantação	O tipo de Alfresco que você usa: `PAAS` ou`ON-PREM`.
crawlType	O tipo de conteúdo que você deseja rastrear, seja `ASPECT` (conteúdo marcado com “Aspectos” no Alfresco), `SITE_ID` (conteúdo em um site específico do Alfresco) ou `ALL_SITES` (conteúdo em todos os sites do Alfresco).
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• document • comentário	Uma lista de objetos que mapeiam os atributos ou nomes de campo de seus documentos e comentários do Alfresco para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
aspectName	O nome de um “Aspecto” específico que você deseja indexar.
aspectProperties	Uma lista de propriedades de conteúdo de “Aspecto” específicas que você deseja indexar.

Configuração	Descrição
enableFineGrainedControle	`true` para rastrear os “Aspectos”.
isCrawlComment	`true`para rastrear comentários.
• inclusionFileNamePadrões • inclusionFileTypePadrões • inclusionFilePathPadrões	Uma lista de padrões de expressões regulares para incluir determinadas páginas e ativos em sua fonte de dados do Alfresco. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
• exclusionFileNamePadrões • exclusionFileTypePadrões • exclusionFilePathPadrões	Uma lista de padrões de expressões regulares para excluir determinadas páginas e ativos em sua fonte de dados do Alfresco. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
tipo	O tipo da fonte de dados. Especifique `ALFRESCO` como seu tipo de fonte de dados.

Configuração	Descrição

secretArn	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu. Alfresco O segredo deve conter uma estrutura JSON com as seguintes chaves:
	
	Se estiver usando a autenticação básica:
	
	`{ "username": " user name", "password": " password" }`
	
	Se estiver usando a autenticação OAuth 2.0:
	
	`{ "clientId": " client ID", "clientSecret": " client secret", "tokenUrl": " token URL" }`

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisad os publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappin gAPI para carregar informações de acesso de usuários e grupos.
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo (MySQL)

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `mysql`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Aurora Esquema JSON (MySQL).

A tabela a seguir descreve os parâmetros do esquema JSON Aurora (MySQL).

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql` `db2`, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes

Configuração	Descrição
	de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.

Configuração	Descrição
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

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 modelo (PostgreSQL)

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `postgresql`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Aurora (PostgreSQL) Esquema JSON.

A tabela a seguir descreve os parâmetros do esquema Aurora JSON (PostgreSQL).

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql`db2, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos

Configuração	Descrição
	específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.

Configuração	Descrição
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Aurora (PostgreSQL) 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 modelo (Microsoft SQL Server)

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `sqlserver`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Amazon RDS Esquema JSON (Microsoft SQL Server).

A tabela a seguir descreve os parâmetros do esquema JSON Amazon RDS (Microsoft SQL Server).

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql`db2, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos

Configuração	Descrição
	específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.

Configuração	Descrição
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

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 modelo (MySQL)

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `mysql`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Amazon RDS Esquema JSON (MySQL).

A tabela a seguir descreve os parâmetros do esquema JSON Amazon RDS (MySQL).

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql`db2, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos

Configuração	Descrição
	específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.

Configuração	Descrição
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

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 modelo (Oracle)

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `oracle`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Amazon RDS Esquema JSON (Oracle).

A tabela a seguir descreve os parâmetros do esquema JSON Amazon RDS (Oracle).

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql`db2, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos

Configuração	Descrição
	específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.

Configuração	Descrição
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Amazon RDS Esquema JSON (Oracle)

```
{
  "$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 modelo (PostgreSQL)

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `postgresql`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Amazon RDS (PostgreSQL) Esquema JSON.

A tabela a seguir descreve os parâmetros do esquema Amazon RDS JSON (PostgreSQL).

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql`db2, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos

Configuração	Descrição
	específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.

Configuração	Descrição
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Amazon RDS (PostgreSQL) 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 S3 esquema de modelo

Inclua um JSON que contém o esquema da fonte de dados como parte da configuração do modelo. Forneça o nome do bucket S3 como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `S3` e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON S3.

A tabela a seguir descreve os parâmetros do esquema Amazon S3 JSON.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
BucketName	O nome do seu Amazon S3 balde.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
• inclusionPatterns • exclusionPatterns	Uma lista de padrões de expressão regular para incluir ou excluir arquivos específicos na

Configuração	Descrição
• inclusionPrefixes • exclusionPrefixes	sua fonte Amazon S3 de dados. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
aclConfigurationFileCaminho	O caminho do arquivo que controla o acesso aos documentos em um índice do Amazon Kendra .
metadataFilesPrefix	O local em seu bucket para arquivos de metadados.
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
tipo	O tipo da fonte de dados. Especifique `S3` como seu tipo de fonte de dados.

Configuração	Descrição
versão	A versão do modelo que é compatível.

Esquema JSON 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 do modelo do Web Crawler

Inclua um JSON que contém o esquema da fonte de dados como parte do objeto do TemplateConfiguration.

Forneça os URLs iniciais ou de ponto de partida ou os URLs do mapa do site, como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Em vez de listar manualmente todos os seus URLs, você pode fornecer o caminho para o Amazon S3 bucket que armazena um arquivo de texto para sua lista de URLs iniciais ou arquivos XML de sitemap, que você pode agrupar em um arquivo ZIP no S3.

Você também especifica o tipo de fonte de dados como`WEBCRAWLERV2`, as credenciais de autenticação do site e o tipo de autenticação, se seus sites exigirem autenticação, e outras configurações necessárias.

Em seguida, especifique `TEMPLATE` como `Type` ao chamar CreateDataSource.

> **Important**
>
> A criação do conector Web Crawler v2.0 não é suportada pelo. AWS CloudFormation Use o conector Web Crawler v1.0 se precisar de suporte. AWS CloudFormation

Ao selecionar sites para indexar, você precisa aderir à Política de uso aceitável da Amazon e a todos os outros termos da Amazon. Lembre-se de que você só deve usar o Amazon Kendra Web Crawler para indexar suas próprias páginas da Web ou páginas da Web que você tenha autorização para indexar. Para saber como impedir que o Web Crawler do Amazon Kendra indexe seus sites, consulte Configurando o arquivo do *`robots.txt`* para o Web Crawler do Amazon Kendra.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Amazon Kendra Esquema JSON do Web Crawler.

A tabela a seguir descreve os parâmetros do esquema JSON do Amazon Kendra Web Crawler.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
siteMapUrls	A lista de URLs de mapa dos sites nos quais você deseja fazer o crawling. Você pode listar até três URLs de mapa de site.

Configuração	Descrição
s3 SeedUrl	O caminho do S3 para o arquivo de texto que armazena a lista de URLs semente ou de partida. Por exemplo, s3://bucket-name/directory/. Cada URL no arquivo de texto deve ser formatado em uma linha separada. Você pode listar até 100 URLs semente em um arquivo.
s3 SiteMapUrl	O caminho do S3 para os arquivos XML do mapa do site. Por exemplo, s3://bucket-name/directory/. Você pode listar até três arquivos XML do mapa do site. Você pode agrupar vários arquivos de sitemap em um arquivo ZIP e armazená-lo em seu Amazon S3 bucket.
seedUrlConnections	A lista de URLs semente ou de partida dos sites nos quais você deseja fazer o crawling. Você pode listar até 100 URLs semente.
seedUrl	O URL semente ou de partida.
authentication	O tipo de autenticação dos sites exigem a mesma autenticação, caso contrário, especifique `NoAuthentication` .
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.

Configuração	Descrição
• webPage • attachment	Uma lista de objetos que mapeiam os atributos ou nomes de campo de suas páginas da Web e arquivos de páginas da Web para Amazon Kendra indexar nomes de campos. Por exemplo, a tag de título da página da web em HTML pode ser mapeada para o campo de índice `_document_title` . Para obter mais informações, consulte Mapear campos de fonte de dados.
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
rateLimit	O número máximo de URLs que o crawling percorre por host de site por minuto.

Configuração	Descrição
maxFileSize	O tamanho máximo (em MB) de uma página da Web ou anexo para crawling.
crawlDepth	O número de níveis do URL semente para crawling. Por exemplo, a página de URL semente tem profundidade 1 e todos os hiperlinks nessa página que também são rastreados têm profundidade 2.
maxLinksPerURL	O número máximo de URLs em uma página da Web a serem incluídos no crawling de um site. Esse número é por página da Web. À medida que as páginas de um site passam pelo crawling, todos os URLs aos quais as páginas se vinculam também são incluídos nele. Os URLs em uma página da Web passam pelo crawling por ordem de exibição.
crawlSubDomain	`true`: fazer crawling dos domínios do site com subdomínios. Por exemplo, se o URL semente for "abc.example.com“, então “a.abc.example.com“ e “b.abc.example.com“ também serão rastreados. Se você não definir `crawlSubDomain` ou `crawlAllDomain` selecionar`true`, Amazon Kendra rastreará apenas os domínios dos sites que você deseja rastrear.
crawlAllDomain	`true`: fazer crawling dos domínios do site com subdomínios e outros domínios aos quais as páginas da Web estão vinculadas. Se você não definir `crawlSubDomain` ou `crawlAllDomain` selecionar`true`, Amazon Kendra rastreará apenas os domínios dos sites que você deseja rastrear.

Configuração	Descrição
honorRobots	`true` para respeitar as diretivas robots.tx t dos sites nos quais você deseja fazer o crawling. Essas diretivas controlam como o Amazon Kendra Web Crawler rastreia os sites, se Amazon Kendra pode rastrear somente conteúdo específico ou não rastrear nenhum conteúdo.
crawlAttachments	`true` para rastrear arquivos aos quais as páginas da web estão vinculadas.
• URL de inclusão CrawlPatterns • URL de inclusão IndexPatterns	Uma lista de padrões de expressão regular que inclui o crawling de determinados URLs e a indexação de quaisquer hiperlinks nessas páginas da Web com URL. Os URLs que correspondem aos padrões são incluídos no índice. Os URLs que não correspondem aos padrões são excluídos do índice. Se um URL corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o URL/páginas da Web não serão incluídos no índice.
• URL de exclusão CrawlPatterns • URL de exclusão IndexPatterns	Uma lista de padrões de expressão regular que inclui o crawling de determinados URLs e a indexação de quaisquer hiperlinks nessas páginas da Web com URL. Os URLs que correspondem aos padrões são excluídos do índice. Os URLs que não correspondem aos padrões são incluídos no índice. Se um URL corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o URL/páginas da Web não serão incluídos no índice.

Configuração	Descrição
inclusionFileIndexPadrões	Uma lista de padrões de expressões regulares para incluirdeterminados arquivos de páginas da Web. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo correspon der tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
exclusionFileIndexPadrões	Uma lista de padrões de expressões regulares para excluirdeterminados arquivos de páginas da Web. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo correspon der tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
proxy	Informações de configuração necessárias para se conectar aos seus sites internos por meio de um proxy da Web.
host	O nome do host do servidor proxy ao qual você deseja se conectar por sites internos. Por exemplo, o nome do host de https://a.example.com/page1.html é "a.example.com”.
porta	O número da porta do servidor proxy ao qual você deseja se conectar por sites internos. Por exemplo, 443 é a porta padrão para HTTPS.

Configuração	Descrição
secretArn (proxy)	Se forem necessárias credenciais de proxy da web para se conectar a um host de site, você poderá criar um AWS Secrets Manager segredo que armazene as credenciais. Forneça o nome do recurso da Amazon (ARN) da senha.
tipo	O tipo da fonte de dados. Especifique `WEBCRAWLERV2` como seu tipo de fonte de dados.

Configuração	Descrição
secretArn	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo usado se seus sites precisarem de autenticação para acessá-los. Você armazena as credenciais de autenticação do site na senha que contém pares de valores-chave JSON. Se você usa o básico ou NTML/Kerberos, digite o nome de usuário e a senha. As chaves JSON na senha devem ser `userName` e `password`. O protocolo de autenticação NTLM inclui hash de senha e o protocolo de autenticação Kerberos inclui criptografia de senha. Se você usar SAML ou autenticação de formulário, insira o nome de usuário e a senha, XPath para o campo de nome de usuário (e botão de nome de usuário se estiver usando SAML), XPaths para o campo e botão de senha e a URL da página de login. As chaves JSON na senha devem ser `userName`, `password`, `userNameFieldXpath` , `userNameButtonXpath` , `passwordFieldXpath` , `passwordButtonXpath` e `loginPageUrl` . Você pode encontrar os XPaths (XML Path Language) dos elementos usando as ferramentas de desenvolvedor do navegador. Os XPaths geralmente seguem este formato:`//tagname[@Attribute='Value']` . Amazon Kendra também verifica se as informações do endpoint (URLs iniciais) incluídas no segredo são as mesmas informaçõ

Configuração	Descrição
	es do endpoint especificadas nos detalhes de configuração do endpoint da fonte de dados.
versão	Atualmente, apenas a versão do modelo tem suporte.

Amazon Kendra Esquema JSON do 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 do modelo do Confluence

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Forneça o URL do host do Confluence, o método de host e o tipo de autenticação como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `CONFLUENCEV2`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Confluence.

A tabela a seguir descreve os parâmetros do esquema JSON do Confluence.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
hostUrl	O URL da sua instância do Confluence. Por exemplo, *https://example.confluence.com* .
tipo	O método de host para sua instância do Confluence: `SAAS` ou `ON_PREM`.
authType	O método de autenticação para sua instância do Confluence: `Basic`, `OAuth2` ou `Personal-token` .
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• space • page • blog • comentário • attachment	Uma lista de objetos que mapeiam os atributos ou nomes de campo de seus espaços, páginas, blogs, comentários e anexos do Confluenc e para indexar Amazon Kendra nomes de campos. Para obter mais informações, consulte Mapping data source fields (Mapear campos de fonte de dados). Os nomes dos campos da fonte de dados do Confluence devem existir nos metadados personalizados do Confluence.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
isCrawlAcl	`true`para rastrear as informações da lista de controle de acesso (ACL) de seus documento s, se você tiver uma ACL e quiser usá-la para

Configuração	Descrição
	controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
fieldForUserIdentificação	Especifique `email` se você deseja usar o e-mail do usuário como ID do usuário. `email`é usado por padrão e atualmente é o único tipo de ID de usuário compatível.
• inclusionSpaceKeyFiltro • exclusionSpaceKeyFiltro • pageTitleRegEX • blogTitleRegEX • commentTitleRegEX • attachmentTitleRegEX • inclusionFileTypePadrões • exclusionFileTypePadrões • inclusionUrlPatterns • exclusionUrlPatterns	Uma lista de padrões de expressões regulares para incluir e/ou excluir determinadas páginas e ativos em sua fonte de dados do Confluenc e. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
proxyHost	O nome do host do proxy da web que você usa, sem o `https://` protocolo `http://` ou.
proxyPort	O número da porta usada pelo protocolo de transporte de URL do host. Esse valor deve estar entre 0 e 65.535.

Configuração	Descrição
• isCrawlPersonalEspaço • isCrawlArchivedEspaço • isCrawlArchivedPágina • isCrawlPage • isCrawlBlog • isCrawlPageComente • isCrawlPageAnexo • isCrawlBlogComente • isCrawlBlogAnexo	`true`para rastrear arquivos em seus espaços pessoais, páginas, blogs, comentários de página, anexos de página, comentários de blog e anexos de blog do Confluence.
maxFileSizeInMegaBytes	Especifique o limite de tamanho do arquivo em MBs que Amazon Kendra pode ser rastreado. Amazon Kendra rastreia somente os arquivos dentro do limite de tamanho definido. O tamanho padrão do arquivo é 50 MB. O tamanho máximo do arquivo deve ser maior que 0MB e menor ou igual a 50MB.
tipo	O tipo da fonte de dados. Especifique `CONFLUENCEV2` como seu tipo de fonte de dados.

Configuração	Descrição
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisad os publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappin gAPI para carregar informações de acesso de usuários e grupos.
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Confluence. Para obter informações sobre esses pares de valores-chave, consulte Instruções de conexão para o Confluence.
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelos do Dropbox

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Você fornece a chave do aplicativo, a senha do aplicativo e o token de acesso do Dropbox como parte da senha que armazena suas credenciais de autenticação. Especifique também o tipo de fonte de dadosDROPBOX, o tipo de token de acesso que você deseja usar (temporário ou permanente) e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Dropbox.

A tabela a seguir descreve os parâmetros do esquema JSON do Dropbox.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados. Essa fonte de dados não especifica um endpoint em `repositoryEndpoint Metadata` . Em vez disso, as informaçõ es de conexão são incluídas em um AWS Secrets Manager segredo que você fornece `secretArn` a.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• file • paper • papert • shortcut	Uma lista de objetos que mapeiam os atributos ou nomes de campo de seus arquivos do Dropbox, do Dropbox Paper e dos atalhos para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz

Configuração	Descrição
	ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos.

Configuração	Descrição
secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Dropbox. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "appKey": "Dropbox app key", "appSecret": " Dropbox app secret", "accesstoken": " temporary access token or refresh access token" }`
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
isCrawlAcl	`true`para rastrear as informações da lista de controle de acesso (ACL) de seus documento s, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

Configuração	Descrição
• inclusionFileNamePadrões • inclusionFileTypePadrões	Uma lista de padrões de expressões regulares para incluir determinados nomes e tipos de arquivos na fonte de dados do Dropbox. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
• exclusionFileNamePadrões • exclusionFileTypePadrões	Uma lista de padrões de expressões regulares para excluir determinados nomes e tipos de arquivos na fonte de dados do Dropbox. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
• crawlFile • crawlPaper • crawlPapert • crawlShortcut	`true`para rastrear arquivos em seu Dropbox, documentos do Dropbox Paper, modelos do Dropbox Paper e atalhos de páginas da web armazenados em seu Dropbox.
tipo	O tipo da fonte de dados. Especifique `DROPBOX` como seu tipo de fonte de dados.

Configuração	Descrição
tokenType	Especifique o tipo de token de acesso: token de acesso permanente ou temporário. É recomendável criar um token de acesso de atualização que nunca expire no Dropbox, em vez de confiar em um token de acesso único que expira após quatro horas. Crie um aplicativo e um token de acesso de atualizaç ão no console do desenvolvedor do Dropbox e forneça o token de acesso na senha.
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelos do Drupal

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Forneça a URL do host do Drupal e o tipo de autenticação como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como DRUPAL, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Drupal.

A tabela a seguir descreve os parâmetros do esquema JSON do Drupal.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
hostUrl	O URL do host do site do Drupal. Por exemplo, *https:///* <hostname><drupalsitename>.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados.
• content • comentário • attachment	Uma lista de objetos que mapeia atributos ou nomes de campos dos arquivos do Drupal. Para obter mais informações, consulte Mapear campos de fonte de dados. Os nomes dos campos da fonte de dados do Drupal devem existir nos metadados personalizados do Drupal .
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
• inclusionFileNamePadrões • articleTitleInclusionPadrões • pageTitleInclusionPadrões • customContentTitleInclusionPatterns • basicBlockTitleInclusionPatterns • customBlockTitleInclusionPatterns	Uma lista de padrões de expressões regulares para incluir determinados arquivos em sua fonte de dados do Drupal. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
• exclusionFileNamePadrões • articleTitleExclusionPadrões • pageTitleExclusionPadrões • customContentTitleExclusionPatterns • basicBlockTitleExclusionPatterns • customBlockTitleExclusionPatterns	Uma lista de padrões de expressões regulares para excluir determinadas páginas e ativos em sua fonte de dados do Drupal. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
contentDefinitions • contentType • Definição de campo • isCrawlComments • isCrawlFiles • isCrawlArticle • isCrawlBasicPágina • isCrawlBasicBloquear • isCrawlCustomContentTypesList	Especifique os tipos de conteúdo a serem rastreados e se os comentários e anexos devem ser rastreados para os tipos de conteúdo selecionados.
tipo	O tipo da fonte de dados. Especifique `DRUPAL` como seu tipo de fonte de dados.
authType	O tipo de autenticação que você usa: `BASIC-AUTH` ou `OAUTH2`.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisad os publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappin gAPI para carregar informações de acesso de usuários e grupos.

secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Drupal. O segredo deve conter uma estrutura JSON com as seguintes chaves:
	
	Se estiver usando a autenticação básica:
	
	`{ "username": "user name", "passwords": "password" }`
	
	Se estiver usando a autenticação OAuth 2.0:
	
	`{ "username": "user name", "password": "password" , "clientId": "client id", "clientSecret": "client secret" }`

Configuração	Descrição
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Você fornece a URL do GitHub host, o nome da organização e se usa a GitHub nuvem ou o GitHub local como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `GITHUB`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte GitHub Esquema JSON.

A tabela a seguir descreve os parâmetros do esquema GitHub JSON.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
tipo	Especifique o tipo como `SAAS` ou`ON_PREMISE` .
hostUrl	O URL do GitHub host. Por exemplo, se você usa GitHub SaaS/Enterprise Cloud:. https://api.github.com Ou, se você usa um servidor GitHub local/corporativo:. https://on-prem-host-url/api/v3/
organizationName	Você pode encontrar o nome da sua organização ao fazer login no GitHub desktop e acessar Suas organizações no menu suspenso da foto do perfil.

Configuração	Descrição
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• Repositório GH • GHCommit • ghIssueDocument • ghIssueComment • ghIssueAttachment • Documento GHPR • Comentário GHPR • Anexo GHPR	Uma lista de objetos que mapeiam os atributos ou nomes de campo do seu GitHub conteúdo para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
isCrawlAcl	`true`para rastrear as informações da lista de controle de acesso (ACL) de seus documento s, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especific a quais documentos os usuários e grupos podem acessar e pesquisar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

Configuração	Descrição
fieldForUserIdentificação	Especifique o tipo de ID de usuário que você deseja usar para o rastreamento da ACL. Especifique `email` se você deseja usar o e-mail do usuário para a ID do usuário ou `username` se deseja usar o nome do usuário para a ID do usuário. Se você não especificar uma opção, ela `email` será usada por padrão.
Filtro de repositório	Uma lista dos nomes dos repositórios e ramificações específicos que você deseja indexar.
Repositório de rastreamento	`true`para rastrear repositórios.
crawlRepositoryDocuments	`true`para rastrear documentos do repositório.
Problema de rastreamento	`true`para rastrear problemas.
crawlIssueComment	`true`para rastrear os comentários do problema.
crawlIssueCommentAnexo	`true`para rastrear anexos de comentários de problemas.
crawlPullRequest	`true`para rastrear pull requests.
crawlPullRequestComente	`true`para rastrear os comentários do pull request.
crawlPullRequestCommentAttachment	`true`para rastrear anexos de comentários do pull request.

Configuração	Descrição
• inclusionFolderNamePadrões • inclusionFileTypePadrões • inclusionFileNamePadrões	Uma lista de padrões de expressão regular para incluir determinado conteúdo em sua fonte GitHub de dados. O conteúdo que corresponde aos padrões é incluído no índice. O conteúdo que não corresponde aos padrões é excluído do índice. Se algum conteúdo corresponder a um padrão de inclusão e exclusão, o padrão de exclusão terá precedência e o conteúdo não será incluído no índice.
• exclusionFolderNamePadrões • exclusionFileTypePadrões • exclusionFileNamePadrões	Uma lista de padrões de expressão regular para excluir determinado conteúdo em sua fonte GitHub de dados. O conteúdo que corresponde aos padrões é excluído do índice. O conteúdo que não corresponde aos padrões é incluído no índice. Se algum conteúdo corresponder a um padrão de inclusão e exclusão, o padrão de exclusão terá precedênc ia e o conteúdo não será incluído no índice.
tipo	O tipo da fonte de dados. Especifique `GITHUB` como seu tipo de fonte de dados.
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisad os publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappin gAPI para carregar informações de acesso de usuários e grupos.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu. GitHub O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "personalToken": " token" }`
versão	A versão desse modelo que é compatível atualmente.

GitHub Esquema JSON

A seguir está o 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 modelos do Gmail

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `GMAIL`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Gmail.

A tabela a seguir descreve os parâmetros do esquema JSON do Gmail.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados. Essa fonte de dados não especifica um endpoint em `repositoryEndpoint Metadata` . Em vez disso, as informaçõ es de conexão são incluídas em um AWS Secrets Manager segredo que você fornece `secretArn` a.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
• message • attachments	Uma lista de objetos que mapeiam os atributos ou nomes de campo de suas mensagens e anexos do Gmail para Amazon Kendra indexar nomes de campos. Para obter mais informaçõ es, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
• inclusionLabelNamePadrões • exclusionLabelNamePadrões • inclusionAttachmentTypePadrões • exclusionAttachmentTypePadrões • inclusionAttachmentNamePadrões • exclusionAttachmentNamePadrões	Uma lista de padrões de expressões regulares para incluir ou excluir mensagens com nomes de assuntos específicos na sua fonte de dados do Gmail. Os arquivos que correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão

Configuração	Descrição
• inclusionSubjectFilter • exclusionSubjectFilter • isSubjectAnd • inclusionFromFilter • exclusionFromFilter • inclusionToFilter • exclusionToFilter • inclusionCcFilter • exclusionCcFilter • inclusionBccFilter • exclusionBccFilter	terá precedência e o arquivo não será incluído no índice.
beforeDateFilter	Especifique mensagens e anexos a serem incluídos antes de uma determinada data.
afterDateFilter	Especifique mensagens e anexos a serem incluídos antes de uma determinada data.
isCrawlAttachment	Um valor booleano para escolher se você deseja rastrear anexos. As mensagens são rastreadas automaticamente.
tipo	O tipo da fonte de dados. Especifique `GMAIL` como seu tipo de fonte de dados.
shouldCrawlDraftMensagens	Um valor booleano para escolher se você deseja rastrear rascunhos de mensagens.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. **Important** Como não há uma API para atualizar mensagens do Gmail excluídas permanentemente, qualquer conteúdo novo, modificado ou excluído é sincronizado: • Não removerá mensagens que foram excluídas permanentemente do Gmail do seu índice Amazon Kendra • Não sincronizará alterações nas etiquetas de e-mail do Gmail Para sincronizar as alterações no rótulo da fonte de dados do Gmail e

Configuração	Descrição
	as mensagens de e-mail excluídas permanentemente com seu Amazon Kendra índice, você deve executar rastreamentos completos periodica mente.
secretARN	O nome do recurso da Amazon (ARN) de uma senha do Secrets Manager que contém os pares de chave/valor necessários para se conectar ao Gmail. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "adminAccountEmailId": " service account email", "clientEmailId": " user account email", "privateKey": " private key" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 do modelo do Google Drive

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `GOOGLEDRIVE2`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Google Drive.

A tabela a seguir descreve os parâmetros do esquema JSON do Google Drive.

Configuração	Descrição
connectionConfiguration	Informações de configuração sobre a fonte de dados
repositoryEndpointMetadata	Informações do endpoint da fonte de dados. Essa fonte de dados não especifica um endpoint. Você escolhe o tipo de autenticação: `serviceAccount` e `OAuth2`. As informaçõ es de conexão estão incluídas em um AWS

Configuração	Descrição
	Secrets Manager segredo que você fornece `secretArn` a.
authType	Escolha entre `serviceAccount` e `OAuth2` com base no caso de uso.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• file • comentário	Uma lista de objetos que mapeia atributos ou nomes de campos do Google Drive para o Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
• maxFileSizeInMegaBytes	Especifique um limite de tamanho de arquivo em MBs que Amazon Kendra deve ser rastreado.
• iscrawlComment	`true`para rastrear comentários na sua fonte de dados do Google Drive.
• isCrawlMyDriveAndSharedWithMe	`true`para rastrear MyDrive e compartilhar unidades do Shared With Me na sua fonte de dados do Google Drive.
• isCrawlSharedConduz	`true`para rastrear unidades compartilhadas na sua fonte de dados do Google Drive.

Configuração	Descrição
isCrawlAcl	`true`para rastrear as informações da lista de controle de acesso (ACL) de seus documento s, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especific a quais documentos os usuários e grupos podem acessar e pesquisar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
• excludeUserAccounts • excludeSharedDrives • excludeMimeTypes • exclusionFileTypePadrões • exclusionFileNamePadrões • exclusionFilePathFiltro	Uma lista de padrões de expressões regulares para excluir determinados arquivos em sua fonte de dados do Google Drive. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
• includeUserAccounts • includeSharedDrives • includeMimeTypes • inclusionFileTypePadrões • inclusionFileNamePadrões • inclusionFilePathFiltro	Uma lista de padrões de expressões regulares para incluir determinados arquivos em sua fonte de dados do Google Drive. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
tipo	O tipo da fonte de dados. Especifique `GOOGLEDRIVEV2` como seu tipo de fonte de dados.
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição

secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Google Drive. O segredo deve conter uma estrutura JSON com as seguintes chaves:
	
	Se estiver usando a autenticação da conta de serviço do Google:
	
	`{ "clientEmail": " user account email", "adminAccountEmail": " service account email", "privateKey": " private key" }`
	
	Se estiver usando a autenticação OAuth 2.0:
	
	`{ "clientID": " OAuth client ID", "clientSecret": " client secret", "refreshToken": " refresh token" }`

versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo do IBM DB2

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados db2, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do IBM DB2.

A tabela a seguir descreve os parâmetros do esquema JSON do IBM DB2.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql` `db2`, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.

Configuração	Descrição
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.

Configuração	Descrição
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo do Microsoft Exchange

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Forneça o URL do host como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `MSEXCHANGE`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Microsoft Exchange.

A tabela a seguir descreve os parâmetros do esquema JSON do Microsoft Exchange.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
tenantId	O ID do locatário do Microsoft 365. Encontre o ID de locatário nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.

Configuração	Descrição
• email • attachment • calendar • contacts • notes	Uma lista de objetos que mapeiam os atributos ou nomes de campo da sua fonte de dados do Microsoft Exchange para campos de Amazon Kendra índice. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções de configuração adicionais para conteúdo em sua fonte de dados
inclusionPatterns	Uma lista de padrões de expressões regulares para incluir determinados arquivos em sua fonte de dados do Microsoft Exchange. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
exclusionPatterns	Uma lista de padrões de expressões regulares para excluir determinados arquivos em sua fonte de dados do Microsoft Exchange. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
• inclusionUsersList • inclusionUsersFileNome • inclusionDomainUsers	Uma lista de padrões de expressões regulares para incluir determinados arquivos em sua fonte de dados do Microsoft Exchange. Os URLs que correspondem aos padrões são incluídos no índice. Os usuários que não correspondem aos padrões são excluídos do índice. Se um usuário corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o usuário não será incluído no índice.
• exclusionUsersList • exclusionUsersFileNome • exclusionDomainUsers	Uma lista de padrões de expressões regulares para excluir determinados arquivos em sua fonte de dados do Microsoft Exchange. Os usuários que não correspondem aos padrões são excluídos do índice. Os usuários que não correspondem aos padrões são incluídos no índice. Se um usuário corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o usuário não será incluído no índice.
s3bucketName	O nome do seu bucket do S3, se quiser usar.
• crawlCalendar • crawlNotes • crawlContacts • crawlFolderAcl	`true`para rastrear esses tipos de conteúdo e informações de controle de acesso à sua fonte de dados do Microsoft Exchange.
startCalendarDateHora	Você pode configurar uma data e hora de início específica para o conteúdo do calendário.
endCalendarDateHora	Você pode configurar uma data e hora de início específica para o conteúdo do calendário.

Configuração	Descrição
subject	Você pode configurar uma linha de assunto específica para o conteúdo do e-mail.
emailFrom	Você pode configurar um e-mail específico para o conteúdo do e-mail “De” ou do remetente.
emailTo	Você pode configurar um e-mail específico para o conteúdo do e-mail “De” ou do remetente.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
tipo	O tipo da fonte de dados. Especifique `MSEXCHANGE` como seu tipo de fonte de dados.

Configuração	Descrição
secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Microsoft Exchange. Isso inclui o ID de cliente e a senha de cliente que são gerados ao criar um aplicativo OAuth no portal do Azure.
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelos da Microsoft

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Forneça o ID do locatário como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `ONEDRIVEV2`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema Microsoft OneDrive JSON.

A tabela a seguir descreve os parâmetros do esquema Microsoft OneDrive JSON.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
tenantId	O ID do locatário do Microsoft 365. Encontre o ID de locatário nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
file	Uma lista de objetos que mapeiam os atributos ou nomes de campo de seus OneDrive arquivos da Microsoft para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
• userNameFilter • userFilterPath • inclusionFileTypePadrões • exclusionFileTypePadrões • inclusionFileNamePadrões • exclusionFileNamePadrões • inclusionFilePathPadrões • exclusionFilePathPadrões • inclusionOneNoteSectionNamePatterns • exclusionOneNoteSectionNamePatterns	Você pode optar por indexar arquivos, OneNote seções e OneNote páginas específicos e filtrar por nome de usuário.

Configuração	Descrição
• inclusionOneNotePageNamePatterns • exclusionOneNotepageNamePatterns	
isUserNameEm S3	`true` para fornecer uma lista de nomes de usuário em um arquivo armazenado em um Amazon S3.
tipo	O tipo da fonte de dados. Especifique `ONEDRIVEV2` como seu tipo de fonte de dados.
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisad os publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappin gAPI para carregar informações de acesso de usuários e grupos.
tipo	O tipo da fonte de dados. Especifique `ONEDRIVEV2` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar à sua Microsoft. OneDrive O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "clientId": " client ID", "clientSecret": " client secret" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema Microsoft OneDrive 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
      }
     },
     "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 modelos da Microsoft

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Você fornece o URL/URLs do SharePoint site, o domínio e também um ID do locatário, se necessário, como parte da configuração da conexão ou dos detalhes do

endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `SHAREPOINTV2`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como o Tipo ao ligar CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte SharePoint Esquema JSON.

A tabela a seguir descreve os parâmetros do esquema Microsoft SharePoint JSON.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
tenantId	O ID do inquilino da sua SharePoint conta.
Domínio	O domínio da sua SharePoint conta.
siteUrls	Os URLs do host da sua SharePoint conta.
repositoryAdditionalProperties	Propriedades adicionais para se conectar ao endpoint do repositório/fonte de dados.
s3bucketName	O nome do Amazon S3 bucket que armazena seu certificado X.509 autoassinado do Azure AD.
s3certificateName	O nome do certificado X.509 autoassinado do Azure AD armazenado em seu bucket. Amazon S3
authType	O tipo de autenticação que você usa, seja `OAuth2OAuth2Certificate` ,`OAuth2App` ,`Basic`,`OAuth2_RefreshToken` ,`NTLM`, ou`Kerberos`.
versão	A SharePoint versão que você usa, seja `Server` ou`Online`.

Configuração	Descrição
onPremVersion	A versão do SharePoint servidor que você usa`2013`, se `20162019`, ou`SubscriptionEdition` .
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• evento • page • file • link • attachment • comentário	Uma lista de objetos que mapeiam os atributos ou nomes de campo do seu SharePoint conteúdo para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
• eventTitleFilterRegEx • pageTitleFilterRegEx • linkTitleFilterRegEx • inclusionFilePath • exclusionFilePath • inclusionFileTypePadrões • exclusionFileTypePadrões • inclusionFileNamePadrões • exclusionFileNamePadrões • inclusionOneNoteSectionNamePatterns • exclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns • exclusionOneNotePageNamePatterns	Uma lista de padrões de expressão regular para incluir/excluir determinado conteúdo em sua fonte de SharePoint dados. Os itens de conteúdo que correspondem aos padrões de inclusão são incluídos no índice. Os itens de conteúdo que não correspondem aos padrões de inclusão são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
• crawlFiles • crawlPages • crawlEvents • crawlComments • crawlLinks • crawlAttachments	`true`para rastrear esses tipos de conteúdo.
crawlAcl	`true`para rastrear as informações da lista de controle de acesso (ACL) de seus documento s, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especific a quais documentos os usuários e grupos podem acessar e pesquisar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
fieldForUserIdentificação	Especifique `email` se você deseja usar o e-mail do usuário para a ID do usuário ou `userPrincipalName` se deseja usar um nome de usuário para a ID do usuário. Se você não especificar uma opção, ela `email` será usada por padrão.
aclConfiguration	`ACLWithLDAPEmailFmt` Especifique `ACLWithManualEmailFmt` ou`ACLWithUs ernameFmtM` .
emailDomain	O domínio do e-mail. Por exemplo, *`"amazon.c om"`* .

Configuração	Descrição
• isCrawlLocalGroupMapping • isCrawlAdGroupMapping	`true`para rastrear informações de mapeament o de grupos.
proxyHost	O nome do host do proxy da web que você usa, sem o protocolo http://ou https://.
proxyPort	O número da porta usada pelo protocolo de transporte de URL do host. Esse valor deve estar entre 0 e 65.535.
tipo	Especifique `SHAREPOINTV2` como seu tipo de fonte de dados.
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisad os publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappin gAPI para carregar informações de acesso de usuários e grupos.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu. SharePoint Para obter informações sobre esses pares de valores-chave, consulte as instruções de conexão para o SharePoint Online e SharePoint o Server.

Configuração	Descrição
versão	Atualmente, apenas a versão do modelo tem suporte.

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 modelo do Microsoft SQL Server

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `sqlserver`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Microsoft SQL Server.

A tabela a seguir descreve os parâmetros do esquema JSON do Microsoft SQL Server.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql` `db2`, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.

Configuração	Descrição
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.

Configuração	Descrição
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo do Microsoft Teams

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Forneça o ID do locatário como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `MSTEAMS`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Microsoft Teams.

A tabela a seguir descreve os parâmetros do esquema JSON do Microsoft Teams.

Configuração	Descrição
connectionConfiguration	Informações de configuração para um endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
tenantId	O ID do locatário do Microsoft 365. Encontre o ID de locatário nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• chatMessage • chatAttachment	Uma lista de objetos que mapeiam os atributos ou nomes de campo do seu conteúdo do

Configuração	Descrição
• channelPost • channelWiki • channelAttachment • meetingChat • meetingFile • meetingNote • calendarMeeting	Microsoft Teams para Amazon Kendra indexar nomes de campo. Para obter mais informaçõ es, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
paymentModel	Especifica o tipo de modelo de pagamento a ser usado com fonte de dados do Microsoft Teams. Os modelos de pagamento do modelo A são restritos aos modelos de licenciam ento e pagamento que exigem conformidade de segurança. Os modelos de pagamento do modelo A são restritos aos modelos de licenciamento e pagamento que exigem conformidade de segurança.
• inclusionTeamNameFiltro • inclusionChannelNameFiltro • inclusionFileNamePadrões • inclusionFileTypePadrões • inclusionUserEmailFiltro • inclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns	Uma lista de padrões de expressões regulares para incluir determinado conteúdo em sua fonte de dados do Microsoft Teams. O conteúdo que corresponde aos padrões é incluído no índice. O conteúdo que não corresponde aos padrões é excluído do índice. Se o conteúdo correspon der tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o conteúdo não será incluído no índice.

Configuração	Descrição
• exclusionTeamNameFiltro • exclusionChannelNameFiltro • exclusionFileNamePadrões • exclusionFileTypePadrões • exclusionUserEmailFiltro • exclusionOneNoteSectionNamePatterns • exclusionOneNotePageNamePatterns	Uma lista de padrões de expressões regulares para excluir determinados conteúdos em sua fonte de dados do Microsoft Teams. O conteúdo que corresponde aos padrões é excluído do índice. O conteúdo que não corresponde aos padrões é incluído no índice. Se o conteúdo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o conteúdo não será incluído no índice.
• isCrawlChatMensagem • isCrawlChatAnexo • isCrawlChannelPublicar • isCrawlChannelAnexo • isCrawlChannelWiki • isCrawlCalendarReunião • isCrawlMeetingBate-papo • isCrawlMeetingArquivo • isCrawlMeetingNota	`true`para rastrear esses tipos de conteúdo em sua fonte de dados do Microsoft Teams.
startCalendarDateHora	Você pode configurar uma data e hora de início específica para o conteúdo do calendário.
endCalendarDateHora	Você pode configurar uma data e hora de início específica para o conteúdo do calendário.
tipo	O tipo da fonte de dados. Especifique `MSTEAMS` como seu tipo de fonte de dados.

Configuração	Descrição
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
secretArn	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Microsoft Teams. Isso inclui o ID de cliente e a senha de cliente que são gerados ao criar um aplicativo OAuth no portal do Azure.

Configuração	Descrição
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo do Microsoft Yammer

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `YAMMER`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como o Tipo ao ligar CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor.

A tabela a seguir descreve os parâmetros do esquema JSON do Microsoft Yammer.

Configuração	Descrição
connectionConfiguration	Informações de configuração sobre a fonte de dados
repositoryEndpointMetadata	Informações do endpoint da fonte de dados. Essa fonte de dados não especifica um endpoint em `repositoryEndpoint Metadata` . Em vez disso, as informaçõ es de conexão são incluídas em um AWS Secrets Manager segredo que você fornece `secretArn` a.

Configuração	Descrição
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• community • usuário • message • attachment	Uma lista de objetos que mapeia atributos de fonte de dados ou nomes de campos do Microsoft Yammer para nomes de campos de índice do Amazon Kendra. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
inclusionPatterns	Uma lista de padrões de expressões regulares para incluir determinados arquivos em sua fonte de dados do Microsoft Yammer. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
exclusionPatterns	Uma lista de padrões de expressões regulares para excluir determinados arquivos em sua fonte de dados do Microsoft Yammer. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
sinceDate	Opte por configurar um parâmetro `psinceDate` ara que o conector do Microsoft Yammer rastreie o conteúdo com base em um `sinceDate` específico.
communityNameFilter	Você pode optar por indexar um conteúdo específico da comunidade. | • isCrawlMessage • isCrawlAttachment • isCrawlPrivateMensagem	`true`para rastrear mensagens, anexos de mensagens e mensagens privadas.
tipo	Especifique `YAMMER` como seu tipo de fonte de dados.
secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Microsoft Yammer. Isto inclui o seu nome de usuário e a chave do Microsoft Yammer, assim como o ID e a senha do cliente que são gerados ao criar uma aplicação OAuth no portal do Azure.

Configuração	Descrição
useChangeLog	`true`para usar o log de alterações do Microsoft Yammer para determinar quais documentos precisam ser atualizados no índice.
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisad os publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappin gAPI para carregar informações de acesso de usuários e grupos.

Esquema JSON do 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 modelo do MySQL

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `mysql`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema MySQL JSON.

A tabela a seguir descreve os parâmetros do esquema JSON do MySQL.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql` `db2`, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.

Configuração	Descrição
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.

Configuração	Descrição
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema MySQL 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"
  ]
}
```

Esquema de modelos do Oracle Database

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `oracle`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Oracle Database.

A tabela a seguir descreve os parâmetros do esquema JSON do banco de dados Oracle.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql`db2, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos

Configuração	Descrição
	específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.

Configuração	Descrição
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo do (PostgreSQL

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Especifique o tipo de fonte de dados como `JDBC`, o tipo de banco de dados `postgresql`, como uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do PostgreSQL.

A tabela a seguir descreve os parâmetros do esquema JSON do PostgreSQL.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações de configuração necessárias para conectar sua fonte de dados. • dbtype—O tipo de banco de dados Java que você usa, seja,, `mysql`db2, `postgresql` ou. `oracle` `sqlserver` • dbhost: o nome do host do banco de dados. • DBPort: a porta do banco de dados. • DBInstance: a instância do banco de dados.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos

Configuração	Descrição
	específicos de mapeamentos de conteúdo e campo. Especifique o tipo de fonte de dados e o ARN da senha.
document	Uma lista de objetos que mapeiam os atributos ou nomes de campo do conteúdo do seu banco de dados para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados. Use para incluir ou excluir um conteúdo específico em sua fonte de dados do banco de dados.
primaryKey	Forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
titleColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
bodyColumn	Forneça o nome da coluna do título do documento na tabela do banco de dados.
sqlQuery	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
timestampColumn	Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informaçõ es de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.

Configuração	Descrição
timestampFormat	Insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.
timezone	Insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
changeDetectingColumns	Insira os nomes das colunas que Amazon Kendra serão usadas para detectar alteraçõe s no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas
allowedUsersColumns	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
allowedGroupsColumn	Insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
sourceURIColumn	Insira o nome da coluna que contém os URLs de origem a serem indexados.
isSslEnabled	Insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB. Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
tipo	O tipo da fonte de dados. Especifique `JDBC` como seu tipo de fonte de dados.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição
secretArn	O nome do recurso da Amazon (ARN) de uma senha do Secret Manager que contém o nome do usuário e a senha para se conectar ao banco de dados. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "user name": "database user name", "password": " password" }`
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo do Salesforce

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Forneça o URL do host do Salesforce como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `SALESFORCEV2`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Salesforce.

A tabela a seguir descreve os parâmetros do esquema JSON do Salesforce.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
hostUrl	O URL da instância do Salesforce a ser indexado.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• account • contact	Uma lista de objetos que mapeiam os atributos ou nomes de campo de suas entidades do

Configuração	Descrição
• campaign • case • product • lead • contract • partner • profile • idea • pricebook • task • solution • attachment • user • document • knowledgeArticles • group • opportunity • chatter • customEntity	Salesforce para Amazon Kendra indexar nomes de campo. Para obter mais informaçõ es, consulte Mapear campos de fonte de dados.

Configuração	Descrição
secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Salesforce. O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "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	Opções adicionais de configuração para o conteúdo em sua fonte de dados.

Configuração	Descrição
• accountFilter • contactFilter • caseFilter • campaignFilter • contractFilter • groupFilter • leadFilter • productFilter • opportunityFilter • partnerFilter • pricebookFilter • ideaFilter • profileFilter • taskFilter • solutionFilter • userFilter • chatterFilter • documentFilter • knowledgeArticleFilter • customEntities	Uma coleção de sequências de caracteres que especifica quais entidades filtrar.

Configuração	Descrição
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	Uma lista de padrões de expressões regulares para incluir determinadas páginas e ativos em sua fonte de dados do Salesforce. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
exclusionPatterns • exclusionDocumentFileTypePatterns • exclusionDocumentFileNamePatterns • exclusionAccountFileTypePatterns • exclusionCampaignFileTypePatterns • exclusionCampaignFileNamePatterns • exclusionCaseFileTypePatterns • exclusionCaseFileNamePatterns • exclusionContactFileTypePatterns • exclusionContractFileNamePatterns • exclusionLeadFileTypePatterns • exclusionLeadFileNamePatterns • exclusionOpportunityFileTypePatterns • exclusionOpportunityFileNamePatterns • exclusionSolutionFileTypePatterns • exclusionSolutionFileNamePatterns • exclusionTaskFileTypePatterns • exclusionTaskFileNamePatterns • exclusionGroupFileTypePatterns • exclusionGroupFileNamePatterns • exclusionChatterFileTypePatterns • exclusionChatterFileNamePatterns • exclusionCustomEntityFileTypePatterns • exclusionCustomEntityFileNamePatterns	Uma lista de padrões de expressões regulares para excluir determinadas páginas e ativos em sua fonte de dados do Salesforce. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
• isCrawlAccount • isCrawlContact • isCrawlCase • isCrawlCampaign • isCrawlProduct • isCrawlLead • isCrawlContract • isCrawlPartner • isCrawlProfile • isCrawlIdea • isCrawlPricebook • isCrawlDocument • crawlSharedDocument • isCrawlGroup • isCrawlOpportunity • isCrawlChatter • isCrawlUser • isCrawlSolution • isCrawlTask • isCrawlAccountAnexos • isCrawlContactAnexos • isCrawlCaseAnexos • isCrawlCampaignAnexos • isCrawlLeadAnexos • isCrawlContractAnexos • isCrawlGroupAnexos • isCrawlOpportunityAnexos • isCrawlChatterAnexos • isCrawlSolutionAnexos	`true`para rastrear esses tipos de arquivos em sua conta do Salesforce.

Configuração	Descrição

• isCrawlTaskAnexos
• isCrawlCustomEntityAttachments
• isCrawlKnowledgeArtigos
  • isCrawlDraft
  • isCrawlPublish
  • isCrawlArchived	

tipo	O tipo da fonte de dados. Especifique `SALESFORCEV2` como seu tipo de fonte de dados.
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
versão	Atualmente, apenas a versão do modelo tem suporte.

Esquema JSON do 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 modelo

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Você fornece o URL do ServiceNow host, o tipo de autenticação e a versão da instância como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `SERVICENOWV2`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte ServiceNow Esquema JSON.

A tabela a seguir descreve os parâmetros do esquema ServiceNow JSON.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
hostUrl	O URL do ServiceNow host. Por exemplo, *`your-domain.service-now.com`* .
authType	O tipo de autenticação que você usa: `basicAuth` ou `OAuth2`.
servicenowInstanceVersion	A ServiceNow versão que você usa. Você pode escolher entre `Tokyo` `Sandiego` `Rome`,, `Others` e.

Configuração	Descrição
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• knowledgeArticle • attachment • serviceCatalog • incident	Uma lista de objetos que mapeiam os atributos ou nomes de campo de seus artigos de ServiceNow conhecimento, anexos, catálogo de serviços e incidente s para Amazon Kendra indexar nomes de campos. Para obter mais informações, consulte Mapear campos de fonte de dados. Os nomes dos campos da fonte de ServiceNow dados devem existir nos seus metadados ServiceNow personalizados.
additional properties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
maxFileSizeInMegaBytes	Especifique o limite de tamanho do arquivo em MBs que o Amazon Kendra rastreará. O Amazon Kendra rastreará somente os arquivos dentro do limite de tamanho que você definir. O tamanho padrão do arquivo é 50 MB. O tamanho máximo do arquivo deve ser maior que 0MB e menor ou igual a 50MB.

Configuração	Descrição
• knowledgeArticleFilter • incidentQueryFilter • serviceCatalogQueryFiltro • knowledgeArticleTitleRegExp • serviceCatalogTitleRegExp • incidentTitleRegExp • inclusionFileTypePadrões • exclusionFileTypePadrões • inclusionFileNamePadrões • exclusionFileNamePadrões • incidentStateType	Uma lista de padrões de expressão regular para incluir e/ou excluir determinados arquivos em sua fonte ServiceNow de dados. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
• isCrawlKnowledgeArtigo • isCrawlKnowledgeArticleAttachment • includePublicArticlesSomente • isCrawlServiceCatálogo • isCrawlServiceCatalogAttachment • isCrawlActiveServiceCatalog • isCrawlInactiveServiceCatalog • isCrawlIncident • isCrawlIncidentAnexo • isCrawlActiveIncidente • isCrawlInactiveIncidente • Aplicar ACL ForKnowledgeArticle • Aplicar ACL ForServiceCatalog • Aplicar ACL ForIncident	`true`para rastrear artigos de ServiceNow conhecimento, catálogos de serviços, incidentes e anexos.
tipo	O tipo da fonte de dados. Especifique `SERVICENOWV2` como seu tipo de fonte de dados.

Configuração	Descrição
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidad e para sincronizar informações de identidade/princip al sobre usuários e grupos com acesso a determina dos documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreado r de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informaçõ es de acesso de usuários e grupos.
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

Configuração	Descrição

secretARN	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-c have necessários para se conectar ao seu. ServiceNo w O segredo deve conter uma estrutura JSON com as seguintes chaves:
	`{ "username": " user name", "password": " password" }`
	Se você usar uma autenticação OAuth 2.0, a senha deverá conter uma estrutura JSON com as seguintes chaves:
	`{ "username": " user name", "password": " password", "clientId": " client id", "clientSecret": " client secret" }`

versão	Atualmente, apenas a versão do modelo tem suporte.

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 modelos do Slack

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Forneça o URL do host como parte da configuração da conexão ou

dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `SLACK`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Slack.

A tabela a seguir descreve os parâmetros do esquema JSON do Slack.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
ID da equipe	O ID da equipe do Slack que você copiou do URL da página principal do Slack.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
Todos	Uma lista de objetos que mapeiam os atributos ou nomes de campo do seu Slack conteúdo para Amazon Kendra indexar nomes de campos.
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
inclusionPatterns	Uma lista de padrões de expressão regular para incluir conteúdo específico em sua fonte Slack de dados. O conteúdo que corresponde aos padrões é incluído no índice. O conteúdo que não corresponde aos padrões é excluído do índice. Se algum conteúdo corresponder a um padrão de inclusão e exclusão, o padrão de

Configuração	Descrição
	exclusão terá precedência e o conteúdo não será incluído no índice.
exclusionPatterns	Uma lista de padrões de expressão regular para excluir conteúdo específico na sua fonte Slack de dados. O conteúdo que corresponde aos padrões é excluído do índice. O conteúdo que não corresponde aos padrões é incluído no índice. Se algum conteúdo corresponder a um padrão de inclusão e exclusão, o padrão de exclusão terá precedência e o conteúdo não será incluído no índice.
crawlBotMessages	`true`para rastrear mensagens de bots.
Excluir arquivado	`true`para excluir o rastreamento de mensagens arquivadas.
Tipo de conversa	O tipo de conversa que você deseja indexar se `PUBLIC_CHANNEL PRIVATE_CHANNEL` , `GROUP_MESSAGE DIRECT_MESSAGE` e.
Filtro de canais	O tipo de canal que você deseja indexar, seja `private_channel` ou`public_channel` .
sinceDate	Você pode optar por configurar um `sinceDate` parâmetro para que o Slack conector rastreie o conteúdo com base em um parâmetro específico. `sinceDate`
Olhe para trás	Você pode optar por configurar um `lookBack` parâmetro para que o Slack conector rastreie o conteúdo atualizado ou excluído até um determinado número de horas antes da última sincronização do conector.

Configuração	Descrição
syncMode	Especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Escolha uma das seguintes opções: • `FORCED_FULL_CRAWL` para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice. • `FULL_CRAWL` para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincroniz ada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização. • `CHANGE_LOG` para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
tipo	O tipo da fonte de dados. Especifique `SLACK` como seu tipo de fonte de dados.

Configuração	Descrição
enableIdentityCrawler	`true`usar o rastreador Amazon Kendra de identidade para sincronizar informações de identidade/principal sobre usuários e grupos com acesso a determinados documentos. Se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisad os publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappin gAPI para carregar informações de acesso de usuários e grupos.
secretArn	O Amazon Resource Name (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu. Slack O segredo deve conter uma estrutura JSON com as seguintes chaves: `{ "slackToken": " token" }`
versão	A versão desse modelo atualmente compatível.

Esquema JSON do 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 do modelo do Zendesk

Você inclui um JSON que contém o esquema da fonte de dados como parte do TemplateConfigurationobjeto. Forneça o URL do host como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Além disso, especifique o tipo de fonte de dados como `ZENDESK`, uma senha para suas credenciais de autenticação e outras configurações necessárias. Em seguida, você especifica `TEMPLATE` como `Type` quando você liga CreateDataSource.

Você pode usar o modelo fornecido neste guia do desenvolvedor. Consulte Esquema JSON do Zendesk.

A tabela a seguir descreve os parâmetros do esquema JSON do Zendesk.

Configuração	Descrição
connectionConfiguration	Informações de configuração para o endpoint da fonte de dados.
repositoryEndpointMetadata	Informações do endpoint da fonte de dados.
hostURL	O URL do host do Zendesk. Por exemplo, https://yoursubdomain.zendesk.com.
repositoryConfigurations	Informações de configuração de conteúdo da fonte de dados. Por exemplo, configurar tipos específicos de mapeamentos de conteúdo e campo.
• ticket • ticketComment • ticketCommentAttachment • article • articleComment • articleAttachment • communityTopic • communityPostComment	Uma lista de objetos que mapeia atributos de fonte de dados ou nomes de campos do Zendesk para nomes de campos de índice do Amazon Kendra. Para obter mais informações, consulte Mapear campos de fonte de dados.
secretARN	O nome de recurso da Amazon (ARN) de um AWS Secrets Manager segredo que contém os pares de valores-chave necessários para se conectar ao seu Zendesk. A senha deve conter uma estrutura JSON com as seguintes chaves: URL do host, ID do cliente, senha do cliente, nome de usuário e senha.

Configuração	Descrição
additionalProperties	Opções adicionais de configuração para o conteúdo em sua fonte de dados.
organizationNameFilter	Você pode optar por indexar os tíquetes em uma organização específica.
sinceDate	Você pode optar por configurar um parâmetro `sinceDate` para que o conector do Zendesk rastreie o conteúdo com base em um `sinceDate` específico.
inclusionPatterns	Uma lista de padrões de expressões regulares para incluir determinadas páginas e ativos em sua fonte de dados do Zendesk. Os arquivos que correspondem aos padrões são incluídos no índice. Os arquivos que não correspondem aos padrões são excluídos do índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.
exclusionPatterns	Uma lista de padrões de expressões regulares para excluir determinados arquivos em sua fonte de dados do Zendesk. Os arquivos que correspondem aos padrões são excluídos do índice. Os arquivos que não correspondem aos padrões são incluídos no índice. Se um arquivo corresponder tanto a um padrão de inclusão como a um de exclusão, o padrão de exclusão terá precedência e o arquivo não será incluído no índice.

Configuração	Descrição
• isCrawlTicket • isCrawlTicketComente • isCrawlTicketCommentAttachment • isCrawlArticle • isCrawlArticleComente • isCrawlArticleAnexo • isCrawlCommunityTópico • isCrawlCommunityPublicar • isCrawlCommunityPostComment	Insira "`true`" para rastrear esses tipos de conteúdo.
tipo	Especifique `ZENDESK` como seu tipo de fonte de dados.
useChangeLog	Insira "`true`" para usar o registro de alterações do Zendesk para determinar quais documentos precisam ser atualizados no índice. Dependendo do tamanho do log de alterações, talvez seja mais rápido digitalizar os documentos no Zendesk. Se estiver sincronizando a fonte de dados do Zendesk com o índice pela primeira vez, todos os documentos serão digitalizados.

Esquema JSON do 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

O Adobe Experience Manager é um sistema de gerenciamento de conteúdo usado para criar conteúdo de site ou aplicativo móvel. Você pode usar Amazon Kendra para se conectar Adobe Experience Manager e indexar suas páginas e ativos de conteúdo.

Amazon Kendra suporta Adobe Experience Manager (AEM) como instância de autor do Cloud Service e instância de autoria e Adobe Experience Manager publicação no local.

Você pode se conectar Amazon Kendra à sua fonte de Adobe Experience Manager dados usando o Amazon Kendra console ou a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Adobe Experience Manager, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão

Atributos compatíveis

o conector de fonte de dados do Adobe Experience Manager oferece suporte aos seguintes recursos:

- Mapeamentos de campos

- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- OAuth 2.0 e autenticação básica
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de Adobe Experience Manager dados, faça essas alterações em suas Adobe Experience Manager AWS contas.

No Adobe Experience Manager, verifique se você:

- Acesso a uma conta com privilégios administrativos ou a um usuário administrador.
- Copiou o URL do host do Adobe Experience Manager.

> **Note**
>
> (Local/servidor) Amazon Kendra verifica se as informações do endpoint incluídas são iguais às informações do endpoint especificadas nos AWS Secrets Manager detalhes de configuração da fonte de dados. Isso ajuda a proteger contra o problema de assistência confusa, que é um problema de segurança em que um usuário não tem permissão para realizar uma ação, mas usa o Amazon Kendra como proxy para acessar a senha configurada e realizar a ação. Se você alterar posteriormente as informações do endpoint, crie uma nova senha para sincronizar essas informações.

- Você anotou as credenciais básicas de autenticação do nome de usuário e senha do administrador.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Opcional: configurou as credenciais do OAuth 2.0 no Adobe Experience Manager (AEM) como um serviço de nuvem ou AEM On-Premise. Se você usa o AEM on-premises, as credenciais incluem o ID do cliente, a senha do cliente e chave privada. Se você usa o AEM como um serviço de nuvem, as credenciais incluem o ID do cliente, a senha do cliente, a chave privada, o ID da organização, o ID da conta técnica e o host (IMS) do Adobe Identity Management System. Para obter mais informações sobre como gerar essas credenciais para o AEM como serviço de nuvem, consulte a documentação do Adobe Experience Manager. Para o AEM on-premises, a implementação do servidor do Adobe Granite OAuth 2.0 (com.adobe.granite.oauth.server) oferece suporte às funcionalidades do servidor OAuth 2.0 no AEM.
- Verifique se cada documento é exclusivo no Adobe Experience Manager e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Adobe Experience Manager em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Adobe Experience Manager Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de Adobe Experience Manager dados, você deve fornecer os detalhes necessários da sua fonte de Adobe Experience Manager dados para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Adobe Experience Manager para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra a Adobe Experience Manager

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector do Adobe Experience Manager e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Adobe Experience Manager com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir acesso e segurança, insira as informações a seguir:

a. Fonte: escolha o AEM on-premises ou o AEM como um serviço em nuvem.

Insira o URL do host do Adobe Experience Manager. Por exemplo, ao usar o AEM on-premises, inclua o nome do host e a porta:. https://hostname:port Ou, ao usar o AEM como um serviço de nuvem, use a URL do autor: https://author-xxxxxx-xxxxxxx.adobeaemcloud.com.

b. Local do certificado SSL: insira o caminho para o certificado SSL armazenado em um bucket do Amazon S3 . Você usa isso para se conectar ao AEM On-Premise com uma conexão SSL segura.

c. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

d. Autenticação: escolha a Autenticação básica ou Autenticação OAuth 2.0. Em seguida, escolha um AWS Secrets Manager segredo existente ou crie um novo segredo para armazenar suas Adobe Experience Manager credenciais. Se você optar por criar um novo segredo, uma janela AWS Secrets Manager secreta será aberta.

Se escolher a Autenticação básica, insira um nome para a senha, o nome de usuário e a senha do site do Adobe Experience Manager. O usuário deve ter permissão de administrador ou ser um usuário administrador.

Se escolheu a Autenticação OAuth 2.0 e usa o AEM on-premises, insira um nome para a senha, o ID do cliente, a senha do cliente e a chave privada. Se você usa o AEM como um serviço de nuvem, insira um nome para a senha, o ID do cliente, a senha do cliente, a chave privada, o ID da organização, o ID da conta técnica e o host (IMS) do Adobe Identity Management System.

Salve e adicione seu segredo.

e. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

f. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

g. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

h. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Escopo de sincronização: defina limites de crawling determinados tipos de conteúdo, componentes da página e caminhos raiz e filtre o conteúdo usando padrões de expressão regex.

i. Tipos de conteúdo: escolha se deseja rastrear somente páginas, ativos ou os dois.

ii. (Opcional) Configurações adicionais: defina as seguintes configurações opcionais:

- Componentes da página: os nomes específicos dos componentes da página. O componente de página é um componente de página extensível projetado para funcionar com o editor de modelos do Adobe Experience Manager e permite que os componentes de cabeçalho/rodapé e estrutura da página sejam montados com o editor de modelos.

- Variações de fragmentos de conteúdo: os nomes específicos das variações de fragmentos de conteúdo. Os fragmentos de conteúdo permitem que você projete, crie, organize e publique conteúdo independente de página no Adobe Experience Manager. Eles permitem que você prepare conteúdo pronto para uso em diferentes locais/canais.
- Caminhos raiz: os caminhos raiz para um conteúdo específico.
- Padrões Regex: os padrões de expressão regular para incluir ou excluir determinadas páginas e ativos.

b. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

c. ID do fuso horário: se você usa o AEM on-premises e o fuso horário do servidor é diferente do fuso horário do conector ou índice do AEM do Amazon Kendra , especifique o fuso horário do servidor para alinhar com o conector ou índice do AEM. O fuso horário padrão para o AEM on-premises é o fuso horário do conector ou índice do AEM do Amazon Kendra . O fuso horário padrão para o AEM como serviço de nuvem é o Greenwich Mean Time.

d. Cronograma de execução de sincronização, por frequência — escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Selecione entre os campos da fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice. Para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   b. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Adobe Experience Manager

Você deve especificar um JSON do esquema da fonte de dados usando a API TemplateConfiguration. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `AEM` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- URL do host do AEM: especifique a URL do host do Adobe Experience Manager. Por exemplo, ao usar o AEM on-premises, inclua o nome do host e a porta:. https://hostname:port Ou, ao usar o AEM como um serviço de nuvem, use a URL do autor: https://author-xxxxxx-xxxxxxx.adobeaemcloud.com.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo

da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Tipo de autenticação: especifique qual tipo de autenticação deseja usar: `Basic` ou `OAuth2`.
- Tipo de AEM: especifique o tipo de Adobe Experience Manager usado: `CLOUD` ou `ON_PREMISE`.
- Nome do recurso da Amazon (ARN) da senha: se quiser usar a autenticação básica para o AEM on-Premises ou Cloud, forneça uma senha que armazena as credenciais de autenticação do nome de usuário e da senha. Você fornece o Amazon Resource Name (ARN) de um AWS Secrets Manager segredo. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "aemUrl": "Adobe Experience Manager On-Premise host URL",
    "username": "user name with admin permissions",
    "password": "password with admin permissions"
}
```

Se você quiser usar a autenticação OAuth 2.0 para o AEM on-premises, a senha é armazenada em uma estrutura JSON com as seguintes chaves:

```
{
    "aemUrl": "Adobe Experience Manager host URL",
    "clientId": "client ID",
    "clientSecret": "client secret",
    "privateKey": "private key"
}
```

Se você quiser usar a autenticação OAuth 2.0 para o Cloud Service, a senha é armazenada em uma estrutura JSON com as seguintes chaves:

```
{
    "clientId": "client ID",
    "clientSecret": "client secret",
    "privateKey": "private key",
    "orgId": "organization ID",
```

```
    "technicalAccountId": "technical account ID",
    "imsHost": "Adobe Identity Management System (IMS) host"
}
```

- IAM função — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do Adobe Experience Manager e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Adobe Experience Manager.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- ID do fuso horário — Se você usa o AEM On-Premise e o fuso horário do seu servidor é diferente do fuso horário do conector ou índice do Amazon Kendra AEM, você pode especificar o fuso horário do servidor para se alinhar ao conector ou índice do AEM.

  O fuso horário padrão para o AEM On-Premise é o fuso horário do conector ou índice do Amazon Kendra AEM. O fuso horário padrão para o AEM como serviço de nuvem é o Greenwich Mean Time.

  Para obter informações sobre os IDs de fuso horário compatíveis, consulte Esquema JSON do Adobe Experience Manager.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinadas páginas e ativos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Adobe Experience Manager para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

  > **Note**
  >
  > O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o Esquema do modelo do Adobe Experience Manager.

Alfresco

OAlfresco é um serviço de gerenciamento de conteúdo que ajuda os clientes a armazenar e gerenciar o conteúdo. Você pode usar Amazon Kendra para indexar sua biblioteca de Alfresco documentos, Wiki e blog.

Amazon Kendra oferece suporte Alfresco local e Alfresco na nuvem (plataforma como serviço).

Você pode se conectar Amazon Kendra à sua fonte de Alfresco dados usando o Amazon Kendra console ou a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados Amazon Kendra Alfresco, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

o conector de fonte de dados do Amazon Kendra Alfresco oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- OAuth 2.0 e autenticação básica
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Alfresco, faça essas alterações em seu Alfresco e. Contas da AWS

No Alfresco, verifique se você:

- Copiou o URL do repositório do Alfresco e o URL do aplicativo da Web. Se quiser indexar apenas um site específico do Alfresco, copie também o ID do site.
- Anote suas credenciais de autenticação do Alfresco, que incluem um nome de usuário e senha com no mínimo permissão de leitura. Para usar a autenticação OAuth 2.0, adicione o usuário ao grupo de administradores do Alfresco.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não

recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Opcional: configurou as credenciais do OAuth 2.0 em. Alfresco As credenciais incluem ID do cliente, senha do cliente e URL do token. Para obter mais informações sobre como configurar clientes no Alfresco on-premises, consulte a documentação do Alfresco. Para usar o Cloud (PaaS) do Alfresco, entre em contato com o suporte da Hyland para a autenticação OAuth 2.0 do Alfresco.
- Verifique se cada documento é exclusivo no Alfresco e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou as credenciais de autenticação do Alfresco em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Alfresco a. Amazon

Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Alfresco, você deve fornecer os detalhes necessários da sua fonte de dados do Alfresco para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Alfresco para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra a Alfresco

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector Alfresco e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Alfresco com a etiqueta “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:

a. Alfrescotipo — Escolha se você usa Alfresco local/servidor ou Alfresco nuvem (plataforma como serviço).

b. URL do repositório do Alfresco: insira o URL do repositório do Alfresco. Por exemplo, se você usa o Cloud (PaaS) do Alfresco, o URL do repositório pode ser https://company.alfrescocloud.com. Ou, se você usa o Alfredo on-premises, o URL do repositório pode ser https://company-alfresco-instance.company-domain.suffix:port.

c. Aplicativo do usuário do Alfresco. URL: insira o URL da interface de usuário do Alfresco. Você pode obter o URL do repositório com o administrador do Alfresco. Por exemplo, o URL da interface do usuário pode ser https://example.com.

d. Local do certificado SSL — Insira o caminho para o certificado SSL armazenado em um bucket. Amazon S3 Você usa isso para se conectar ao Alfresco on-premises com uma conexão SSL segura.

e. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

f. Autenticação: escolha a Autenticação básica ou Autenticação OAuth 2.0. Em seguida, escolha uma senha do Secrets Manager existente ou crie um novo segredo para armazenar as credenciais do Alfresco. Se você optar por criar um novo segredo, uma janela AWS Secrets Manager secreta será aberta.

   Se escolher a Autenticação básica, insira um nome para a senha, o nome de usuário e a senha do Alfresco.

   Se escolheu a Autenticação OAuth 2.0, insira um nome para a senha, o ID do cliente, a senha do cliente e a chave privada.

g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

h. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador

de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

i. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

j. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Escopo de sincronização: defina limites de crawling determinados e filtre o conteúdo usando padrões de expressão regex.

   b. i. Conteúdo: escolha se deseja rastrear conteúdo marcado com “Aspectos” no Alfresco, conteúdo em um site específico do Alfresco ou conteúdo em todos os sites do Alfresco.

      ii. (Opcional) Configurações adicionais: defina as seguintes configurações opcionais:

      - Incluir comentários: escolha incluir comentários na biblioteca de documentos e no blog do Alfresco.
      - Padrões Regex: os padrões de expressão regular para incluir ou excluir determinados arquivos.

   c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Selecione entre os campos da fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

   b. Para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Alfresco

Você deve especificar um JSON do esquema da fonte de dados usando a API TemplateConfiguration. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `ALFRESCO` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- ID do site do Alfresco: especifique o ID do site do Alfresco.
- URL do repositório do Alfresco: especifique o URL do repositório do Alfresco. Você pode obter o URL do repositório com o administrador do Alfresco. Por exemplo, se você usa o Cloud

(PaaS) do Alfresco, o URL do repositório pode ser https://company.alfrescocloud.com. Ou, se você usa o Alfredo on-premises, o URL do repositório pode ser https://company-alfresco-instance.company-domain.suffix:port.

- URL do aplicativo da Web do Alfresco: especifique o URL da interface do usuário do Alfresco. Você pode obter o URL do repositório com o administrador do Alfresco. Por exemplo, o URL da interface do usuário pode ser https://example.com.
- Tipo de autenticação: especifique qual tipo de autenticação deseja usar: `OAuth2` ou `Basic`.
- Tipo do Alfresco: especifique qual o tipo do Alfresco uado: `PAAS` (nuvem/plataforma como serviço) ou `ON_PREM` (on-premises).
- Nome do recurso da Amazon (ARN) da senha: para usar a autenticação básica, forneça uma senha que armazena as credenciais de autenticação do nome de usuário e da senha. Você fornece o Amazon Resource Name (ARN) de um AWS Secrets Manager segredo. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "username": "user name",
    "password": "password"
}
```

Para usar a autenticação OAuth 2.0, a senha é armazenada em uma estrutura JSON com as seguintes chaves:

```
{
    "clientId": "client ID",
    "clientSecret": "client secret",
    "tokenUrl": "token URL"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Alfresco e. Amazon Kendra Para obter mais informações, Funções do IAM para fontes de dados do Alfresco.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Tipo de conteúdo: o tipo de conteúdo que você deseja rastrear, seja um conteúdo marcado com “Aspectos” no Alfresco, conteúdo em um site específico do Alfresco ou conteúdo em todos os sites do Alfresco. Você também pode listar conteúdo específico de “Aspectos”.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinadas arquivos.

  > **Note**
  >
  > A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados

publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Alfresco para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o Esquema do modelo do Alfresco.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Alfresco, consulte:

- Pesquise Alfresco conteúdo de forma inteligente usando Amazon Kendra

Aurora (MySQL)

Aurora é um sistema de gerenciamento de banco de dados relacional (RDBMS) criado para a nuvem. Se você for um Aurora usuário, poderá usar Amazon Kendra para indexar sua fonte Aurora (MySQL) de dados. O conector da fonte de Amazon Kendra Aurora (MySQL) dados oferece suporte ao Aurora MySQL 3 e ao Aurora Serverless MySQL 8.0.

Você pode se conectar Amazon Kendra à sua fonte de Aurora (MySQL) dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra Aurora (MySQL) dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de Aurora (MySQL) dados, faça essas alterações em suas Aurora (MySQL) AWS contas.

Em Aurora (MySQL), verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados. Você pode encontrar essas informações no Amazon RDS console.
- Verifique se cada documento é exclusivo em Aurora (MySQL) e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de Aurora (MySQL) em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de Aurora (MySQL) dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de Aurora (MySQL) dados, você deve fornecer detalhes de suas Aurora (MySQL) credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Aurora (MySQL) para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a Aurora (MySQL)

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Aurora (MySQL)conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o Aurora (MySQL)conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. Em Fonte, insira o seguinte:
   b. Host: insira o URL do host do banco de dados; por exemplo: `http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`.
   c. Porta: insira a porta do banco de dados; por exemplo, `5432`.
   d. Instância: insira a instância do banco de dados.
   e. Autenticação: insira as seguintes informações:
      - AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de Aurora (MySQL) autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- Aurora (MySQL) -' é adicionado automaticamente ao seu nome secreto.

II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

B. Escolha Salvar.

f. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

g. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

h. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Em Sincronizar escopo, escolha uma das opções a seguir:

- Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB As consultas SQL devem ter menos de 32 KB e não conter ponto e vírgula (;). Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
- Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
- Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
- Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

   - Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
   - Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
   - Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
   - Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.
   - Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
   - Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
   - Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

   - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
   - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Aurora (MySQL)

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `mySql`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de

dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:

- `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
- `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. Aurora (MySQL) A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o Aurora (MySQL) conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do Aurora (MySQL).

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Aurora (MySQL) para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Aurora Esquema de modelo (MySQL).

Observações

- As linhas excluídas do banco de dados não serão rastreadas durante a Amazon Kendra verificação do conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.

- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Aurora (PostgreSQL)

Aurora é um sistema de gerenciamento de banco de dados relacional (RDBMS) criado para a nuvem. Se você for um Aurora usuário, poderá usar Amazon Kendra para indexar sua fonte Aurora (PostgreSQL) de dados. O conector da fonte de Amazon Kendra Aurora (PostgreSQL) dados é compatível com o Aurora PostgreSQL 1.

Você pode se conectar Amazon Kendra à sua fonte de Aurora (PostgreSQL) dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra Aurora (PostgreSQL) dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de Aurora (PostgreSQL) dados, faça essas alterações em suas Aurora (PostgreSQL) AWS contas.

Em Aurora (PostgreSQL), verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados.
- Verifique se cada documento é exclusivo em Aurora (PostgreSQL) e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de Aurora (PostgreSQL) em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de Aurora (PostgreSQL) dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de Aurora (PostgreSQL) dados, você deve fornecer detalhes de suas Aurora (PostgreSQL) credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Aurora (PostgreSQL) para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a Aurora (PostgreSQL)

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Aurora (PostgreSQL)conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o Aurora (PostgreSQL)conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Em Fonte, insira o seguinte:

   b. Host: insira o URL do host do banco de dados; por exemplo: `http://instance URL.region.rds.amazonaws.com`.

   c. Porta: insira a porta do banco de dados; por exemplo, `5432`.

   d. Instância: insira a instância do banco de dados; por exemplo, `postgres`.

   e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

   f. Em Autenticação: insira as seguintes informações:

      - AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de Aurora (PostgreSQL) autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

        A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

           I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- Aurora (PostgreSQL) -' é adicionado automaticamente ao seu nome secreto.

           II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

        B. Escolha Salvar.

   g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

   h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Em Sincronizar escopo, escolha uma das opções a seguir:

      - Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB As consultas SQL devem ter menos de 32 KB e não conter ponto e vírgula (;). Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
      - Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
      - Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
      - Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

   b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

      - Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
      - Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
      - Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
      - Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.
      - Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
      - Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.

- Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta

página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Aurora (PostgreSQL)

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `postgresql`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Modo de sincronização — Amazon Kendra especifique como atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. Aurora (PostgreSQL) A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o Aurora (PostgreSQL) conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do Aurora (PostgreSQL).

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Aurora (PostgreSQL) para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice_`document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Aurora Esquema de modelo (PostgreSQL).

Observações

- As linhas excluídas do banco de dados não serão rastreadas ao Amazon Kendra verificar o conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.
- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Amazon FSx (Windows)

Amazon FSx (Windows) é um sistema de servidor de arquivos baseado em nuvem totalmente gerenciado que oferece recursos de armazenamento compartilhado. Se você for um usuário Amazon FSx (Windows), você pode usar Amazon Kendra para indexar sua fonte de dados Amazon FSx (Windows).

> **Note**
>
> Amazon Kendra agora oferece suporte a um conector atualizado Amazon FSx (Windows).

O console foi atualizado automaticamente para você. Todos os novos conectores que você criar no console usarão a arquitetura atualizada. Se você usa a API, agora deve usar o TemplateConfigurationobjeto em vez do `FSxConfiguration` objeto para configurar seu conector.
Os conectores configurados usando o console antigo e a arquitetura de API continuarão funcionando conforme configurados. No entanto, você não poderá editá-los ou atualizá-los. Se você quiser editar ou atualizar a configuração do conector, deverá criar um novo conector. Recomendamos migrar o fluxo de trabalho do conector para a versão atualizada. O suporte para conectores configurados usando a arquitetura mais antiga está programado para terminar em junho de 2024.

Você pode se conectar Amazon Kendra à sua fonte de dados Amazon FSx (Windows) usando o Amazon Kendra console ou a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados Amazon Kendra Amazon FSx (Windows), consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra Amazon FSx O conector de fonte de dados (Windows) oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Rastreamento de identidade do usuário
- Filtros de inclusão e exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de usar Amazon Kendra para indexar sua fonte de dados Amazon FSx (Windows), verifique os detalhes de seu Amazon FSx (Windows) Contas da AWS e.

Para Amazon FSx (Windows), verifique se você tem:

- Configure Amazon FSx (Windows) com permissões de leitura e montagem.
- Anotou o ID do seu sistema de arquivos. Você pode encontrar o ID do sistema de arquivos no painel Sistemas de arquivos no console Amazon FSx (Windows).
- Configurou uma nuvem privada virtual usando Amazon VPC onde seu sistema de arquivos Amazon FSx (Windows) reside.
- Anotou suas credenciais de autenticação Amazon FSx (Windows) para uma conta de Active Directory usuário. Isso inclui seu nome de usuário do Active Directory com seu nome de domínio DNS (por exemplo, user@corp.example.com) e senha.

  > **Note**
  >
  > Use somente as credenciais necessárias para que o conector funcione. Não use credenciais privilegiadas, como administrador de domínio.

  > **Note**
  >
  > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Verificado se cada documento é exclusivo no Amazon FSx (Windows) e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.

- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação Amazon FSx (Windows) em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados Amazon FSx (Windows) Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados Amazon FSx (Windows), você deve fornecer os detalhes necessários da sua fonte de dados Amazon FSx (Windows) para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Amazon FSx (Windows) para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao seu sistema de arquivos Amazon FSx (Windows)

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha conector Amazon FSx (Windows) e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Amazon FSx (Windows) com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Amazon FSx ID do sistema de arquivos (Windows) — Selecione na lista suspensa sua ID do sistema de arquivos existente, obtida Amazon FSx em (Windows). Ou crie um sistema de arquivos Amazon FSx (Windows). Você pode encontrar o ID do sistema de arquivos no painel Sistemas de arquivos no console Amazon FSx (Windows).

   b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   c. Autenticação — escolha um AWS Secrets Manager segredo existente ou crie um novo segredo para armazenar suas credenciais do sistema de arquivos. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

Forneça um segredo que armazene suas credenciais de autenticação do seu nome de usuário e senha. O nome de usuário deve incluir seu nome de domínio DNS. Por exemplo, user@corp.example.com.

Salve e adicione seu segredo.

d. Virtual Private Cloud (VPC) — Você deve selecionar um Amazon VPC local onde seu Amazon FSx (Windows) reside. Você inclui a sub-rede e os grupos de segurança da VPC. Consulte Configurando um Amazon VPC.

e. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

f. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Escopo de sincronização, padrões Regex — adicione padrões de expressão regular para incluir ou excluir determinados arquivos.

b. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

c. Cronograma de execução da sincronização — em Frequência, escolha com que frequência sincronizar o conteúdo da fonte de dados e atualize seu índice.

d. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos padrão Amazon Kendra gerados dos seus arquivos que você deseja mapear para o seu índice. Para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

b. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao seu sistema de arquivos Amazon FSx (Windows)

Você deve especificar um JSON do esquema da fonte de dados usando a API TemplateConfiguration. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `FSX` quando você usa o esquema TemplateConfigurationJSON. Além disso, especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- ID do sistema de arquivos — O identificador do sistema de arquivos Amazon FSx (Windows). Você pode encontrar o ID do sistema de arquivos no painel Sistemas de arquivos no console Amazon FSx (Windows).
- Tipo de sistema de arquivos: especifique o tipo de sistema de arquivos como `WINDOWS`.
- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.

> **Note**
>
> Você deve selecionar um Amazon VPC local onde seu Amazon FSx (Windows) resida. Você inclui a sub-rede e os grupos de segurança da VPC.

- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta (Windows). Amazon FSx A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "username": "user@corp.example.com",
    "password": "password"
}
```

- IAM role — Especifique `RoleArn` quando você chama `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Amazon FSx (Windows) e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes de dados Amazon FSx (Windows).

Você também pode adicionar os seguintes recursos opcionais:

- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinadas arquivos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Lista de controle de acesso (ACL) — Especifique se deseja rastrear as informações da ACL para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

> **Note**
>
> Para testar a filtragem de contexto de usuário em um usuário, você deve incluir o nome de domínio DNS como parte do nome de usuário ao realizar a consulta. Você deve ter permissões administrativas do domínio do Active Directory. Você também pode testar a filtragem de contexto do usuário no nome de um grupo.

- Mapeamentos de campo — Escolha mapear seus campos de fonte de dados Amazon FSx (Windows) para seus Amazon Kendra campos de índice. Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice_`document_body`. Todos os demais campos são opcionais.

Para obter uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo Amazon FSx (Windows).

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados Amazon FSx (Windows), consulte:

- Pesquise com segurança dados não estruturados em sistemas de arquivos Windows com o Amazon Kendra conector para Amazon FSx (Windows) para. Windows File Server

Amazon FSx (EM UM NetApp TOQUE)

Amazon FSx (NetApp ONTAP) é um sistema de servidor de arquivos baseado em nuvem totalmente gerenciado que oferece recursos de armazenamento compartilhado. Se você for um usuário Amazon FSx (NetApp ONTAP), você pode usar Amazon Kendra para indexar sua fonte de dados Amazon FSx (NetApp ONTAP).

Você pode se conectar Amazon Kendra à sua fonte de dados Amazon FSx (NetApp ONTAP) usando o Amazon Kendra console ou a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados Amazon Kendra Amazon FSx (NetApp ONTAP), consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão

Atributos compatíveis

Amazon Kendra Amazon FSx O conector de fonte de dados (NetApp ONTAP) oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão e exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de usar Amazon Kendra para indexar sua fonte de dados Amazon FSx (NetApp ONTAP), verifique os detalhes de sua Amazon FSx (NetApp ONTAP) e. Contas da AWS

Para Amazon FSx (NetApp ONTAP), verifique se você tem:

- Configure Amazon FSx (NetApp ONTAP) com permissões de leitura e montagem.
- Anotou o ID do seu sistema de arquivos. Você pode encontrar o ID do sistema de arquivos no painel Sistemas de arquivos no console Amazon FSx (NetApp ONTAP).
- Anotou a ID da máquina virtual de armazenamento (SVM) usada com seu sistema de arquivos. Você pode encontrar sua ID SVM acessando o painel Sistemas de arquivos no console Amazon FSx (NetApp ONTAP), selecionando a ID do sistema de arquivos e, em seguida, selecionando Máquinas virtuais de armazenamento.
- Configurou uma nuvem privada virtual usando Amazon VPC onde seu sistema de arquivos Amazon FSx (NetApp ONTAP) reside.
- Anotou suas credenciais de autenticação Amazon FSx (NetApp ONTAP) para uma conta de Active Directory usuário. Isso inclui seu nome de usuário do Active Directory com seu nome de domínio DNS (por exemplo, user@corp.example.com) e senha. Se você usa o protocolo Network File System (NFS) para seu sistema de arquivos Amazon FSx (NetApp ONTAP), as credenciais de autenticação incluem uma ID esquerda, uma ID direita e uma chave pré-compartilhada.

> **Note**
>
> Use somente as credenciais necessárias para que o conector funcione. Não use credenciais privilegiadas, como administrador de domínio.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Verifique se cada documento é exclusivo no Amazon FSx (NetApp ONTAP) e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação Amazon FSx (NetApp ONTAP) em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não

recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados Amazon FSx (NetApp ONTAP) a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados Amazon FSx (NetApp ONTAP), você deve fornecer os detalhes necessários da sua fonte de dados Amazon FSx (NetApp ONTAP) para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Amazon FSx (NetApp ONTAP) para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao seu sistema de arquivos Amazon FSx (NetApp ONTAP)

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

   > **Note**
   >
   > Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha conector Amazon FSx (NetApp ONTAP) e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Amazon FSx (NetApp ONTAP) com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. Fonte — Forneça as informações do seu sistema de arquivos.

- Protocolo do sistema de arquivos — Escolha o protocolo do seu sistema de arquivos Amazon FSx (NetApp ONTAP). Você pode escolher o protocolo Common Internet File System (CIFS) ou o protocolo Network File System (NFS) para Linux.
- Amazon FSx ID do sistema de arquivos (NetApp ONTAP) — Selecione na lista suspensa sua ID do sistema de arquivos existente, obtida de (ONTAP). Amazon FSx NetApp Ou crie um sistema de arquivos Amazon FSx (NetApp ONTAP). Você pode encontrar o ID do sistema de arquivos no painel Sistemas de arquivos no console Amazon FSx (NetApp ONTAP).
- ID SVMAmazon FSx (NetApp ONTAPsomente para NetApp ONTAP) — Forneça a ID da máquina virtual de armazenamento (SVM) da sua Amazon FSx (ONTAP). NetApp NetApp ONTAP Você pode encontrar sua ID SVM acessando o painel Sistemas de arquivos no console Amazon FSx (NetApp ONTAP), selecionando a ID do sistema de arquivos e selecionando Máquinas virtuais de armazenamento.

b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

c. Autenticação — escolha um AWS Secrets Manager segredo existente ou crie um novo segredo para armazenar suas credenciais do sistema de arquivos. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

Forneça um segredo que armazene suas credenciais de autenticação do seu nome de usuário e senha. O nome de usuário deve incluir seu nome de domínio DNS. Por exemplo, user@corp.example.com.

Se você usa o protocolo NFS para seu sistema de arquivos Amazon FSx (NetApp ONTAP), forneça um segredo que armazene suas credenciais de autenticação da ID esquerda, da ID da direita e da chave pré-compartilhada.

Salve e adicione seu segredo.

d. Virtual Private Cloud (VPC) — Você deve selecionar um Amazon VPC local onde seu Amazon FSx (ONTAP) reside. NetApp Você inclui a sub-rede e os grupos de segurança da VPC. Consulte Configurando um Amazon VPC.

e. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

f. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Escopo de sincronização, padrões Regex — adicione padrões de expressão regular para incluir ou excluir determinados arquivos.

b. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

c. Cronograma de execução da sincronização — em Frequência, escolha com que frequência sincronizar o conteúdo da fonte de dados e atualize seu índice.

d. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos padrão Amazon Kendra gerados dos seus arquivos que você deseja mapear para o seu índice. Para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

b. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao seu sistema de arquivos Amazon FSx (NetApp ONTAP)

Você deve especificar um JSON do esquema da fonte de dados usando a API TemplateConfiguration. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `FSXONTAP` quando você usa o esquema TemplateConfigurationJSON. Além disso, especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- ID do sistema de arquivos — O identificador do sistema de arquivos Amazon FSx (NetApp ONTAP). Você pode encontrar o ID do sistema de arquivos no painel Sistemas de arquivos no console Amazon FSx (NetApp ONTAP).
- ID SVM — A ID da máquina virtual de armazenamento (SVM) usada com seu sistema de arquivos. Você pode encontrar sua ID SVM acessando o painel Sistemas de arquivos no console Amazon FSx (NetApp ONTAP), selecionando a ID do sistema de arquivos e, em seguida, selecionando Máquinas virtuais de armazenamento.
- Tipo de protocolo — Especifique se você usa o protocolo Common Internet File System (CIFS) ou o protocolo Network File System (NFS) para Linux.
- Tipo de sistema de arquivos — especifique o tipo de sistema de arquivos como qualquer um. `FSXONTAP`

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.

  > **Note**
  >
  > Você deve selecionar um Amazon VPC local onde seu Amazon FSx (NetApp ONTAP) reside. Você inclui a sub-rede e os grupos de segurança da VPC.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta (ONTAP). Amazon FSx NetApp A senha deve conter uma estrutura JSON com as seguintes chaves:

  ```
  {
      "username": "user@corp.example.com",
      "password": "password"
  }
  ```

  Se você usa o protocolo NFS para seu sistema de arquivos Amazon FSx (NetApp ONTAP), o segredo é armazenado em uma estrutura JSON com as seguintes chaves:

  ```
  {
      "leftId": "left ID",
      "rightId": "right ID",
      "preSharedKey": "pre-shared key"
  }
  ```

- IAM role — Especifique `RoleArn` quando você chama `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o NetApp conector Amazon FSx (ONTAP) e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes de dados Amazon FSx (NetApp ONTAP).

Você também pode adicionar os seguintes recursos opcionais:

- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão.

Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:

- `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
- `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinadas arquivos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Lista de controle de acesso (ACL) — Especifique se deseja rastrear as informações da ACL para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

> **Note**
>
> Para testar a filtragem de contexto de usuário em um usuário, você deve incluir o nome de domínio DNS como parte do nome de usuário ao realizar a consulta. Você deve ter permissões administrativas do domínio do Active Directory. Você também pode testar a filtragem de contexto do usuário no nome de um grupo.

- Mapeamentos de campo — Escolha mapear seus campos de fonte de dados Amazon FSx (NetApp ONTAP) para seus campos de índice. Amazon Kendra Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice `_document_body`. Todos os demais campos são opcionais.

Para obter uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo Amazon FSx (NetApp ONTAP).

Amazon RDS/Aurora

Você pode indexar documentos armazenados em um banco de dados usando uma fonte de dados do banco de dados. Depois de fornecer as informações de conexão para o banco de dados, Amazon Kendra conecta e indexa os documentos.

Amazon Kendra suporta os seguintes bancos de dados:

- Amazon Aurora MySQL
- Amazon Aurora PostgreSQL
- Amazon RDS para MySQL
- Amazon RDS para PostgreSQL

> **Note**
>
> Não há suporte para bancos de dados do Aurora com tecnologia sem servidor.

> **Important**
>
> Esse conector Amazon RDS/Aurora está programado para ser descontinuado até o final de 2023.
> Amazon Kendra agora oferece suporte a novos conectores de fonte de dados de banco de dados. Para uma experiência aprimorada, recomendamos escolher entre os seguintes novos conectores para seu caso de uso:

- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon RDS (MySQL)
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (Oráculo)
- Amazon RDS (PostgreSQL)
- IBM DB2
- Microsoft SQL Server
- MySQL
- Oracle Database
- PostgreSQL

Você pode se conectar Amazon Kendra à sua fonte de dados do banco de dados usando o Amazon Kendra console e a DatabaseConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra banco de dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão

Atributos compatíveis

Amazon Kendra o conector de fonte de dados do banco de dados oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do banco de dados, faça essas alterações no banco de dados e AWS nas contas.

No banco de dados, verifique se você:

- Anotou as credenciais básicas de autenticação do nome de usuário e senha no banco de dados.
- Copiou o nome do host, o número da porta, o endereço do host, o nome do banco de dados e o nome da tabela de dados que contém os dados do documento. Para o PostgreSQL, a tabela de dados deve ser uma tabela pública ou um esquema público.

> **Note**
>
> O host e a porta informam Amazon Kendra onde encontrar o servidor de banco de dados na Internet. O nome do banco de dados e o nome da tabela informam Amazon Kendra onde encontrar os dados do documento no servidor do banco de dados.

- Copiou os nomes das colunas na tabela de dados que contêm os dados do documento. Inclua o ID do documento, o corpo do documento, as colunas para detectar se um documento foi alterado (por exemplo, a última coluna atualizada) e as colunas opcionais da tabela de dados que foram mapeadas para campos de índice personalizados. Você também pode mapear qualquer nome de campo reservado do Amazon Kendra para uma coluna da tabela.
- Copiou as informações do tipo de mecanismo de banco de dados, como se você usa Amazon RDS para MySQL ou outro tipo.
- Verificou se cada documento é exclusivo no banco de dados e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do banco de dados em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do banco de dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à fonte de dados do banco de dados, você deve fornecer os detalhes necessários da fonte de dados do banco de dados para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o banco de dados para Amazon Kendra, consulte Pré-requisitos.

Console

Para se conectar Amazon Kendra a um banco de dados

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha conector de banco de dados e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector de banco de dados com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Endpoint: um nome de host DNS, um endereço IPv4 ou um endereço IPv6.

   b. Porta: um número de porta.

   c. Banco de dados: nome do banco de dados.

   d. Nome da tabela: nome da tabela.

   e. Em Tipo de autenticação, escolha entre Existente e Novo para armazenar as credenciais de autenticação do banco de dados. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

      - Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

        A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-database-' é adicionado automaticamente ao seu nome secreto.

B. Em Nome de usuário e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

C. Escolha Salvar autenticação.

f. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

> **Note**
>
> Você deve usar uma sub-rede privada. Se sua instância do RDS estiver em uma sub-rede pública na VPC, crie uma sub-rede privada que tenha acesso de saída a um gateway NAT na sub-rede pública. As sub-redes fornecidas na configuração da VPC devem estar nas regiões Oeste dos EUA (Oregon), Leste dos EUA (N. da Virgínia), Leste dos EUA (N. da Virgínia) e UE (Irlanda).

g. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

h. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Selecione entre Aurora MySQL, MySQL, Aurora PostgreSQL e PostgreSQL com base no seu caso de uso.

b. Coloque os identificadores SQL entre aspas duplas: selecione para colocar os identificadores SQL entre aspas duplas. Por exemplo, “ColumnName”.

c. Coluna ACL e colunas de detecção de alterações — Configure as colunas Amazon Kendra usadas para detecção de alterações (por exemplo, a última coluna atualizada) e sua lista de controle de acesso.

d. Em Cronograma de execução da sincronização, em Frequência — Escolha com que frequência Amazon Kendra será sincronizada com sua fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Amazon Kendra mapeamentos de campo padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice. Você deve adicionar os valores da coluna Banco de dados para `document_id` e `document_body`
   b. Mapeamentos de campo personalizados: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.
   c. Escolha Próximo.
9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a um banco de dados

Você deve especificar o seguinte na DatabaseConfigurationAPI:

- ColumnConfiguration—Informações sobre onde o índice deve obter as informações do documento do banco de dados. Para obter mais detalhes, consulte ColumnConfiguration. Você deve especificar os campos `DocumentDataColumnName` (corpo do documento ou texto principal), `DocumentIdColumnName` e `ChangeDetectingColumn` (por exemplo, última coluna atualizada). A coluna mapeada para o campo `DocumentIdColumnName` deve ser uma coluna inteira. O exemplo a seguir mostra uma configuração de colunas simples para uma fonte de dados de banco de dados:

```
"ColumnConfiguration": {
    "ChangeDetectingColumns": [
        "LastUpdateDate",
        "LastUpdateTime"
    ],
    "DocumentDataColumnName": "TextColumn",
    "DocumentIdColumnName": "IdentifierColumn",
    "DocoumentTitleColumnName": "TitleColumn",
    "FieldMappings": [
        {
```

```
            "DataSourceFieldName": "AbstractColumn",
            "IndexFieldName": "Abstract"
        }
    ]
}
```

- ConnectionConfiguration— Informações de configuração necessárias para se conectar a um banco de dados. Para obter mais detalhes, consulte ConnectionConfiguration.
- DatabaseEngineType— O tipo de mecanismo de banco de dados que executa o banco de dados. O `DatabaseHost` campo para `ConnectionConfiguration` deve ser o endpoint da instância Amazon Relational Database Service (Amazon RDS) do banco de dados. Não use o endpoint do cluster.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta de banco de dados. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "username": "user name",
    "password": "password"
}
```

O exemplo a seguir mostra uma configuração de banco de dados, incluindo o 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**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não

recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você chama `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do banco de dados e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do banco de dados.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique `VpcConfiguration` como parte da configuração da fonte de dados. Consulte Configuração do Amazon Kendra para usar uma VPC.

> **Note**
>
> Você deve usar somente uma sub-rede privada. Se sua instância do RDS estiver em uma sub-rede pública na VPC, crie uma sub-rede privada que tenha acesso de saída a um gateway NAT na sub-rede pública. As sub-redes fornecidas na configuração da VPC devem estar nas regiões Oeste dos EUA (Oregon), Leste dos EUA (N. da Virgínia), Leste dos EUA (N. da Virgínia) e UE (Irlanda).

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

Amazon RDS (Microsoft SQL Server)

O SQL Server é um sistema de gerenciamento de banco de dados desenvolvido pela Microsoft. Amazon RDS for SQL Server facilita a configuração, a operação e a escalabilidade das implantações do SQL Server na nuvem. Se você for um usuário Amazon RDS (Microsoft SQL Server), você pode usar Amazon Kendra para indexar sua fonte de dados Amazon RDS (Microsoft SQL Server). O conector da fonte de dados Amazon Kendra JDBC oferece suporte ao Microsoft SQL Server 2019.

Você pode se conectar Amazon Kendra à sua fonte de dados Amazon RDS (Microsoft SQL Server) usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados Amazon Kendra Amazon RDS (Microsoft SQL Server), consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados Amazon RDS (Microsoft SQL Server), faça essas alterações na sua Amazon RDS (Microsoft SQL Server) e AWS nas contas.

No Amazon RDS (Microsoft SQL Server), verifique se você tem:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados.
- Verificado se cada documento é exclusivo no Amazon RDS (Microsoft SQL Server) e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação Amazon RDS (Microsoft SQL Server) em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados Amazon RDS

(Microsoft SQL Server) Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados Amazon RDS (Microsoft SQL Server), você deve fornecer detalhes de suas credenciais Amazon RDS (Microsoft SQL Server) para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Amazon RDS (Microsoft SQL Server), Amazon Kendra consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao Amazon RDS (Microsoft SQL Server)

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha conector Amazon RDS (Microsoft SQL Server) e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Amazon RDS (Microsoft SQL Server) com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Em Fonte, insira o seguinte:

   b. Host: insira o nome do host do banco de dados.

   c. Port: insira a porta do banco de dados.

   d. Instância: insira a instância do banco de dados.

   e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

   f. Em Autenticação: insira as seguintes informações:

      - AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação Amazon RDS (Microsoft SQL Server). Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

        A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

           I. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Amazon RDS (Microsoft SQL Server) -' é adicionado automaticamente ao seu nome secreto.

           II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

        B. Escolha Salvar.

   g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

   h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

      > **Note**
      >
      > IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

   i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Em Sincronizar escopo, escolha uma das opções a seguir:

      - Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.

        > **Note**
        >
        > Se o nome de uma tabela incluir caracteres especiais (não alfanuméricos) no nome, você deverá usar colchetes ao redor do nome da tabela. Por exemplo, `selecione * em [my-database-table]`

      - Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
      - Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
      - Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

   b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

      - Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
      - Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
      - Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
      - Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.
      - Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.

- Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
- Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Amazon RDS (Microsoft SQL Server)

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `sqlserver`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.

  > **Note**
  >
  > Se o nome de uma tabela incluir caracteres especiais (não alfanuméricos) no nome, você deverá usar colchetes ao redor do nome da tabela. Por exemplo, *`selecione * em [my-database-table]`*

- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta (Amazon RDS Microsoft SQL Server). A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você chama `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector (Amazon RDS Microsoft SQL Server) e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes de dados Amazon RDS (Microsoft SQL Server).

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos.

As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

- Mapeamentos de campo — Escolha mapear seus campos de fonte de dados (Amazon RDS Microsoft SQL Server) para seus Amazon Kendra campos de índice. Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Amazon RDS Esquema de modelo (Microsoft SQL Server).

Observações

- As linhas excluídas do banco de dados não serão rastreadas durante a Amazon Kendra verificação do conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.
- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Amazon RDS (MySQL)

Amazon RDS (Amazon Relational Database Service) é um serviço web que facilita a configuração, a operação e a escalabilidade de um banco de dados relacional na AWS nuvem. Se você for um Amazon RDS usuário, poderá usar Amazon Kendra para indexar sua fonte Amazon RDS (MySQL) de dados. O conector da fonte de Amazon Kendra dados é compatível com Amazon RDS MySql 5.6, 5.7 e 8.0.

Você pode se conectar Amazon Kendra à sua fonte de Amazon RDS (MySQL) dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra Amazon RDS (MySQL) dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de Amazon RDS (MySQL) dados, faça essas alterações em suas Amazon RDS (MySQL) AWS contas.

Em Amazon RDS (MySQL), verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados. Você pode encontrar essas informações no Amazon RDS console.
- Verifique se cada documento é exclusivo em Amazon RDS (MySQL) e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de Amazon RDS (MySQL) em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de Amazon RDS (MySQL)

dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de Amazon RDS (MySQL) dados, você deve fornecer detalhes de suas Amazon RDS (MySQL) credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Amazon RDS (MySQL) para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a Amazon RDS (MySQL)

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

   > **Note**
   >
   > Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Amazon RDS (MySQL)conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o Amazon RDS (MySQL)conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. Em Fonte, insira o seguinte:

b. Host: insira o URL do host do banco de dados; por exemplo: `http://instance URL.region.rds.amazonaws.com`.

c. Porta: insira a porta do banco de dados; por exemplo, `5432`.

d. Instância: insira a instância do banco de dados; por exemplo, `postgres`.

e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

f. Em Autenticação: insira as seguintes informações:

- AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de Amazon RDS (MySQL) autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

  A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

     I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- Amazon RDS (MySQL) -' é adicionado automaticamente ao seu nome secreto.

     II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

  B. Escolha Salvar.

g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Em Sincronizar escopo, escolha uma das opções a seguir:

- Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB As consultas SQL devem ter menos de 32 KB e não conter ponto e vírgula (;). Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
- Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
- Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
- Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

- Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
- Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.
- Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
- Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
- Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

   - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
   - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
   - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Amazon RDS (MySQL)

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `mySql`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. Amazon RDS (MySQL) A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
```

```
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o Amazon RDS (MySQL) conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do Amazon RDS (MySQL).

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Amazon RDS (MySQL) para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso

do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Amazon RDS Esquema de modelo (MySQL).

Observações

- As linhas excluídas do banco de dados não serão rastreadas durante a Amazon Kendra verificação do conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.
- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Amazon RDS (Oracle)

Amazon RDS (Amazon Relational Database Service) é um serviço web que facilita a configuração, a operação e a escalabilidade de um banco de dados relacional na AWS nuvem. Se você for um Amazon RDS (Oracle) usuário, poderá usar Amazon Kendra para indexar sua fonte Amazon RDS (Oracle) de dados. O conector da fonte de Amazon Kendra Amazon RDS (Oracle) dados é compatível com Amazon RDS Oracle Database 21c, Oracle Database 19c, Oracle Database 12c.

Você pode se conectar Amazon Kendra à sua fonte de Amazon RDS (Oracle) dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra Amazon RDS (Oracle) dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis

- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de Amazon RDS (Oracle) dados, faça essas alterações em suas Amazon RDS (Oracle) AWS contas.

Em Amazon RDS (Oracle), verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados.
- Verifique se cada documento é exclusivo em Amazon RDS (Oracle) e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.

- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de Amazon RDS (Oracle) em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de Amazon RDS (Oracle) dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de Amazon RDS (Oracle) dados, você deve fornecer detalhes de suas Amazon RDS (Oracle) credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Amazon RDS (Oracle) para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a Amazon RDS (Oracle)

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Amazon RDS (Oracle)conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o Amazon RDS (Oracle)conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Em Fonte, insira o seguinte:

   b. Host: insira o nome do host do banco de dados.

   c. Port: insira a porta do banco de dados.

   d. Instância: insira a instância do banco de dados.

   e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

   f. Em Autenticação: insira as seguintes informações:

      - AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de Amazon RDS (Oracle) autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- Amazon RDS (Oracle) -' é adicionado automaticamente ao seu nome secreto.

II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

B. Escolha Salvar.

g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Em Sincronizar escopo, escolha uma das opções a seguir:

- Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
- Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
- Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

- Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
- Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.
- Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
- Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
- Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear

alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Amazon RDS (Oracle)

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `oracle`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Modo de sincronização — Amazon Kendra especifique como atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo

que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:

- `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
- `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. Amazon RDS (Oracle) A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o Amazon RDS (Oracle) conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do Amazon RDS (Oracle).

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Amazon RDS (Oracle) para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Amazon RDS Esquema de modelo (Oracle).

Observações

- As linhas excluídas do banco de dados não serão rastreadas ao Amazon Kendra verificar o conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.

- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Amazon RDS (PostgreSQL)

Amazon RDS é um serviço web que facilita a configuração, a operação e a escalabilidade de um banco de dados relacional na AWS nuvem. Se você for um Amazon RDS usuário, poderá usar Amazon Kendra para indexar sua fonte Amazon RDS (PostgreSQL) de dados. O conector da fonte de Amazon Kendra Amazon RDS (PostgreSQL) dados é compatível com o PostgreSQL 9.6.

Você pode se conectar Amazon Kendra à sua fonte de Amazon RDS (PostgreSQL) dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra Amazon RDS (PostgreSQL) dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de Amazon RDS (PostgreSQL) dados, faça essas alterações em suas Amazon RDS (PostgreSQL) AWS contas.

Em Amazon RDS (PostgreSQL), verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados. Você pode encontrar essas informações no Amazon RDS console.
- Verifique se cada documento é exclusivo em Amazon RDS (PostgreSQL) e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de Amazon RDS (PostgreSQL) em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de Amazon RDS (PostgreSQL) dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de Amazon RDS (PostgreSQL) dados, você deve fornecer detalhes de suas Amazon RDS (PostgreSQL) credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Amazon RDS (PostgreSQL) para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a Amazon RDS (PostgreSQL)

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Amazon RDS (PostgreSQL)conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o Amazon RDS (PostgreSQL)conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. Em Fonte, insira o seguinte:

b. Host: insira o URL do host do banco de dados; por exemplo: `http://instance URL.region.rds.amazonaws.com`.

c. Porta: insira a porta do banco de dados; por exemplo, `5432`.

d. Instância: insira a instância do banco de dados; por exemplo, `postgres`.

e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

f. Em Autenticação: insira as seguintes informações:

- AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de Amazon RDS (PostgreSQL) autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

  A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

     I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- Amazon RDS (PostgreSQL) -' é adicionado automaticamente ao seu nome secreto.

     II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

  B. Escolha Salvar.

g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Em Sincronizar escopo, escolha uma das opções a seguir:

   - Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ter menos de 32 KB As consultas SQL devem ter menos de 32 KB e não conter ponto e vírgula (;). Amazon Kendra rastreará todo o conteúdo do banco de dados que corresponda à sua consulta.
   - Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
   - Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
   - Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

   b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

   - Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
   - Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
   - Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
   - Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.

- Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
- Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
- Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Amazon RDS (PostgreSQL)

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `postgresql`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte

de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. Amazon RDS (PostgreSQL) A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o Amazon RDS (PostgreSQL) conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do Amazon RDS (PostgreSQL).

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Amazon RDS (PostgreSQL) para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice_`document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Amazon RDS Esquema de modelo (PostgreSQL).

Observações

- As linhas excluídas do banco de dados não serão rastreadas durante a Amazon Kendra verificação do conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.
- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Amazon S3

Amazon S3 é um serviço de armazenamento de objetos que armazena dados como objetos dentro de buckets. Você pode usar Amazon Kendra para indexar seu repositório de documentos em Amazon S3 bucket.

> **Warning**
>
> Amazon Kendra não usa uma política de bucket que conceda permissões a um Amazon Kendra principal para interagir com um bucket do S3. Em vez disso, ele usa IAM funções. Certifique-se de que isso Amazon Kendra não esteja incluído como membro confiável em sua política de bucket para evitar problemas de segurança de dados ao conceder permissões acidentalmente a diretores arbitrários. No entanto, você pode adicionar uma política de bucket para usar um bucket do Amazon S3 em contas diferentes. Para obter mais informações, consulte Políticas para uso do Amazon S3 em todas as contas (na guia Funções do IAM S3, em Funções para fontes de dados do IAM). Para obter informações sobre IAM funções para fontes de dados do S3, consulte IAM funções.

> **Note**
>
> Amazon Kendra agora oferece suporte a um Amazon S3 conector atualizado.
> O console foi atualizado automaticamente para você. Todos os novos conectores que você criar no console usarão a arquitetura atualizada. Se você usa a API, agora deve usar o TemplateConfigurationobjeto em vez do `S3DataSourceConfiguration` objeto para configurar seu conector.
> Os conectores configurados usando o console antigo e a arquitetura de API continuarão funcionando conforme configurados. No entanto, você não poderá editá-los ou atualizá-los. Se você quiser editar ou atualizar a configuração do conector, deverá criar um novo conector. Recomendamos migrar o fluxo de trabalho do conector para a versão atualizada. O suporte para conectores configurados usando a arquitetura mais antiga está programado para terminar em junho de 2024.

Você pode se conectar à sua fonte de Amazon S3 dados usando o Amazon Kendra console ou a TemplateConfigurationAPI.

> **Note**
>
> Para gerar um relatório de status de sincronização para sua fonte de Amazon S3 dados, consulte Solução de problemas com fontes de dados.

Para solucionar problemas do conector da fonte de dados Amazon Kendra S3, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Criação de uma fonte Amazon S3 de dados
- Amazon S3 metadados do documento
- Controle de acesso para fontes Amazon S3 de dados
- Usando Amazon VPC com uma fonte Amazon S3 de dados

Atributos compatíveis

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do S3, faça essas alterações no S3 e AWS nas contas.

No S3, verifique se você:

- Copiou o nome do seu Amazon S3 bucket.

> **Note**
>
> Seu bucket deve estar na mesma região do seu Amazon Kendra índice e seu índice deve ter permissão para acessar o bucket que contém seus documentos.

- Verifique se cada documento é exclusivo no S3 e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter

o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

Em sua AWS conta, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

Se você não tiver uma IAM função existente, poderá usar o console para criar uma nova IAM função ao conectar sua fonte de dados do S3 a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função existente e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do S3, você deve fornecer os detalhes necessários da sua fonte de dados do S3 para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o S3 para Amazon Kendra, consulte Pré-requisitos.

Console

Para se conectar Amazon Kendra a Amazon S3

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector S3 e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector S3 com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações opcionais a seguir:

a. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

b. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

c. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Para localização da fonte de dados — especifique o caminho para o Amazon S3 bucket em que seus dados estão armazenados. Selecione Procurar no S3 para escolher seu bucket do S3.

b. Para Tamanho máximo do arquivo — especifique um limite em MB para rastrear somente arquivos abaixo desse limite. O tamanho Amazon Kendra máximo de arquivo permitido é de 50 MB.

c. Para arquivos de metadados (opcional), prefixe a localização da pasta — especifique o caminho para a pasta na qual seus campos/atributos e outros metadados do documento estão armazenados. Selecione Procurar no S3 para localizar a pasta de metadados.

d. Para (opcional) localização do arquivo de configuração da lista de controle de acesso — especifique o caminho para o arquivo que contém uma estrutura JSON de seus usuários e seu acesso aos documentos. Selecione Procurar S3 para localizar o arquivo ACL.

e. (Opcional) Selecione a chave de decodificação: selecione para usar uma chave de decodificação. Você pode optar por usar uma AWS KMS chave existente.

f. Para configuração adicional (opcional) — adicione padrões para incluir ou excluir determinados arquivos. Todos os caminhos são relativos ao bucket S3 da localização da fonte de dados.

g. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

   - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
   - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

h. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

i. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações opcionais:

   a. Mapeamentos de campo padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta

página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Amazon S3

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `S3` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- BucketName— O nome do bucket que contém os documentos.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- IAM role — Especifique `RoleArn` quando você chama `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector S3 e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do S3.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.

- Filtros de inclusão e exclusão — especifique se deseja incluir ou excluir determinados nomes de arquivo, tipos de arquivo e caminhos de arquivo. Você usa padrões globais (padrões que podem expandir um padrão curinga em uma lista de nomes de caminhos que correspondem ao padrão fornecido). Para ver exemplos, consulte Uso de filtros de exclusão e inclusão na referência de comandos da AWS CLI.
- Configuração de metadados do documento e controle de acesso — adicione metadados do documento e arquivos de controle de acesso que contenham informações como o URI de origem, o autor do documento ou os atributos/campos personalizados do documento e seus usuários e quais documentos eles podem acessar. Cada arquivo de metadados contém metadados sobre um único documento.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do S3 para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

  > **Note**
  >
  > O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o Esquema do modelo do S3.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do S3, consulte:

- Pesquise respostas com precisão usando o Amazon Kendra S3 Connector com suporte a VPC

Criação de uma fonte Amazon S3 de dados

Os exemplos a seguir demonstram a criação de uma fonte de Amazon S3 dados. Os exemplos pressupõem que você já tenha criado um índice e uma IAM função com permissão para ler os dados do índice. Para obter mais informações sobre a IAM função, consulte funções de IAM acesso. Para obter mais informações sobre como criar um índice, consulte Como criar um 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

O trecho de código Python a seguir cria uma fonte de dados. Amazon S3 Para um exemplo completo, consulte Conceitos básicos (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
    )
```

Pode levar algum tempo para criar sua fonte de dados. Você pode monitorar o progresso usando a DescribeDataSourceAPI. Quando o status da fonte de dados é `ACTIVE`, a fonte de dados está pronta para uso.

Os exemplos a seguir demonstram como obter o status de uma fonte de dados.

CLI

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

Python

O trecho de código Python a seguir obtém informações sobre uma fonte de dados do S3. Para um exemplo completo, consulte Conceitos básicos (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
```

Essa fonte de dados não tem um agendamento e, portanto, não é executada automaticamente. Para indexar a fonte de dados, você chama StartDataSourceSyncJobpara sincronizar o índice com a fonte de dados.

Os exemplos a seguir demonstram a sincronização de uma fonte de dados.

CLI

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

```
--id data source ID
```

Python

O trecho de código Python a seguir sincroniza uma fonte de dados do Amazon S3 . Para um exemplo completo, consulte Conceitos básicos (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 metadados do documento

Você pode adicionar metadados, informações adicionais sobre um documento, aos documentos em um bucket do Amazon S3 usando um arquivo de metadados. Cada arquivo de metadados está associado a um documento indexado.

Os arquivos de metadados devem ser armazenados no mesmo bucket dos arquivos indexados. Você pode especificar um local dentro do bucket para seus arquivos de metadados usando o console ou o `S3Prefix` campo do `DocumentsMetadataConfiguration` parâmetro ao criar uma fonte de Amazon S3 dados. Se você não especificar um prefixo do Amazon S3 , os arquivos de metadados deverão ser armazenados no mesmo local dos documentos indexados.

Se você especificar um Amazon S3 prefixo para seus arquivos de metadados, eles estarão em uma estrutura de diretórios paralela aos seus documentos indexados. Amazon Kendra procura seus metadados somente no diretório especificado. Se os metadados não forem lidos, verifique se a localização do diretório corresponde à localização dos metadados.

Os seguintes exemplos mostram como a localização do documento indexado é mapeada para a localização do arquivo de metadados: Observe que a Amazon S3 chave do documento é anexada ao Amazon S3 prefixo dos metadados e depois sufixada com `.metadata.json` para formar o caminho do arquivo de metadados. Amazon S3 A Amazon S3 chave combinada, com o Amazon S3 prefixo e o `.metadata.json` sufixo dos metadados, não deve ter mais do que um total de 1024 caracteres. É recomendável que você mantenha sua Amazon S3 chave abaixo de 1000 caracteres para considerar caracteres adicionais ao combinar sua chave com o prefixo e o sufixo.

```
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
```

Os metadados do documento são definidos em um arquivo JSON. O arquivo deve ser um arquivo de texto UTF-8 sem um marcador BOM. O nome do arquivo JSON deve ser `<document>.<extension>.metadata.json`. Neste exemplo, o "documento" é o nome do documento ao qual os metadados se aplicam e a "extensão" é a extensão do arquivo do documento. O ID do documento deve ser exclusivo na `<document>.<extension>.metadata.json`.

O conteúdo do arquivo JSON segue esse modelo. Todos os atributos/campos são opcionais, portanto, não é necessário incluir todos os atributos. Você deve fornecer um valor para cada atributo que deseja incluir; o valor não pode estar vazio. Se você não especificar o`_source_uri`, os links retornados Amazon Kendra nos resultados da pesquisa apontarão para o Amazon S3 bucket que contém o documento. `DocumentId`é mapeado para o campo `s3_document_id` e é o caminho absoluto para o documento no 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."
}
```

Os campos de metadados `_created_at` e `_last_updated_at` são datas codificadas no ISO 8601. Por exemplo, 2012-03-25T12:30:10+01:00 é o formato de data e hora do ISO 8601 para 25 de março de 2012 às 12h30 (mais 10 segundos) no horário da Europa Central.

Você pode adicionar informações adicionais ao campo `Attributes` sobre um documento que você usa para filtrar consultas ou agrupar respostas de consultas. Para ter mais informações, consulte Criação de campos de documentos personalizados.

Você pode usar o campo `AccessControlList` para filtrar a resposta de uma consulta. Dessa forma, somente determinados usuários e grupos têm acesso aos documentos. Para ter mais informações, consulte Filtragem no contexto do usuário.

Controle de acesso para fontes Amazon S3 de dados

Você pode controlar o acesso aos documentos em uma fonte Amazon S3 de dados usando um arquivo de configuração. Você especifica o arquivo no console ou como `AccessControlListConfiguration` parâmetro ao chamar a UpdateDataSourceAPI CreateDataSourceou.

O arquivo de configuração contém uma estrutura JSON que identifica um prefixo S3 e lista as configurações de acesso para o prefixo. O prefixo pode ser um caminho ou um arquivo individual. Se o prefixo for um caminho, as configurações de acesso se aplicarão a todos os arquivos nesse caminho. Há um número máximo de prefixos S3 no arquivo de configuração JSON e um tamanho máximo de arquivo padrão. Para mais informações, consulte Cotas para Amazon Kendra.

Você pode especificar usuários e grupos nas configurações de acesso. Ao consultar o índice, especifique informações do usuário e do grupo. Para ter mais informações, consulte Filtrando por atributo do usuário.

A estrutura JSON do arquivo de configuração deve estar no seguinte 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"
            }
        ]
    }
]
```

Usando Amazon VPC com uma fonte Amazon S3 de dados

Este tópico fornece um step-by-step exemplo que mostra como se conectar a um bucket do Amazon S3 usando um conector do Amazon S3 por meio do Amazon VPC. O exemplo pressupõe que você esteja começando com um bucket S3 existente. Recomendamos que você faça upload de apenas alguns documentos em seu bucket do S3 para testar o exemplo.

Você pode se conectar Amazon Kendra ao seu Amazon S3 bucket por meio de Amazon VPC. Para fazer isso, você deve especificar a Amazon VPC sub-rede e os grupos de Amazon VPC segurança ao criar seu conector de fonte de Amazon S3 dados.

> **Important**
>
> Para que um Amazon Kendra Amazon S3 conector possa acessar seu Amazon S3 bucket, certifique-se de ter atribuído um Amazon S3 endpoint à sua nuvem privada virtual (VPC).

Amazon Kendra Para sincronizar documentos do seu Amazon S3 bucket Amazon VPC, você deve concluir as seguintes etapas:

- Configure um Amazon S3 endpoint para Amazon VPC. Para obter mais informações sobre como configurar um Amazon S3 endpoint, consulte Endpoints do Gateway Amazon S3 no AWS PrivateLink Guia.
- (Opcional) Verificou suas políticas Amazon S3 de bucket para garantir que o Amazon S3 bucket seja acessível a partir da nuvem privada virtual (VPC) à qual você atribuiu. Amazon KendraPara obter mais informações, consulte Como controlar o acesso de VPC endpoints com políticas de bucket no Guia do usuário do Amazon S3

Etapas

- Etapa 1: configurar um Amazon VPC
- (Opcional) Etapa 2: configurar a política Amazon S3 de bucket
- Etapa 3: criar um conector de fonte Amazon S3 de dados de teste

Etapa 1: configurar um Amazon VPC

Crie uma rede VPC, incluindo uma sub-rede privada com um endpoint de Amazon S3 gateway e um grupo de segurança para Amazon Kendra uso posterior.

Para configurar uma VPC com uma sub-rede privada, um endpoint S3 e um grupo de segurança

1. Faça login no AWS Management Console e abra o Amazon VPC console emhttps://console.aws.amazon.com/vpc/.
2. Crie uma VPC com uma sub-rede privada e um endpoint S3 para usar: Amazon Kendra

   No painel de navegação, escolha Suas VPCs e, em seguida, escolha Criar VPC.

   a. Em Resources to create (Recursos a serem criados), escolha VPC and more (VPC e mais).
   b. Em Etiqueta de nome, ative Geração automática e, em seguida, insira**`kendra-s3-example`**.
   c. Para o bloco CIDR IPv4/IPv6, mantenha os valores padrão.
   d. Em Número de zonas de disponibilidade (AZs), escolha o número 1.
   e. Selecione Personalizar AZs e, em seguida, selecione uma zona de disponibilidade na lista Primeira zona de disponibilidade.

      Amazon Kendra só oferece suporte a um conjunto específico de zonas de disponibilidade.
   f. Em Número de sub-redes públicas, escolha o número 0.
   g. Em Número de sub-redes privadas, escolha o número 1.
   h. Em NAT gateways (Gateways NAT), escolha None (Nenhum).
   i. Para endpoints de VPC, escolha gateway.Amazon S3 .
   j. Deixe o resto dos valores em suas configurações padrão.
   k. Selecione Create VPC (Criar VPC).

      Espere até que o fluxo de trabalho Create VPC termine. Em seguida, escolha Exibir VPC para verificar a VPC que você acabou de criar.

   Agora você criou uma rede VPC com uma sub-rede privada, que não tem acesso à Internet pública.
3. Copie o ID do endpoint VPC do seu endpoint Amazon S3:

   a. No painel de navegação, escolha Endpoints.
   b. Na lista de endpoints, encontre o `kendra-s3-example-vpce-s3` endpoint Amazon S3 que você acabou de criar junto com sua VPC.
   c. Anote o ID do VPC endpoint.

Agora você criou um endpoint de gateway do Amazon S3 para acessar seu bucket do Amazon S3 por meio de uma sub-rede.

4. Crie um grupo de segurança Amazon Kendra para usar:

   a. No painel de navegação, escolha Grupos de segurança e, em seguida, selecione Criar grupo de segurança.

   b. Em Nome do grupo de segurança, insira **`s3-data-source-security-group`**.

   c. Escolha sua VPC na Amazon VPClista.

   d. Deixe as regras de entrada e de saída como padrão.

   e. Escolha Create security group (Criar grupo de segurança).

   Agora você criou um grupo de segurança de VPC.

Você atribui a sub-rede e o grupo de segurança que você criou ao seu conector de fonte de dados do Amazon Kendra Amazon S3 durante o processo de configuração do conector.

(Opcional) Etapa 2: configurar a política Amazon S3 de bucket

Nesta etapa opcional, aprenda a configurar uma política de bucket do Amazon S3 para que seu bucket do Amazon S3 só possa ser acessado a partir da VPC à qual você atribui. Amazon Kendra

Amazon Kendra usa funções do IAM para acessar seu bucket do Amazon S3 e não exige que você configure uma política de bucket do Amazon S3. No entanto, talvez seja útil criar uma política de bucket se quiser configurar um Amazon S3 conector usando um bucket do Amazon S3 que tenha políticas existentes que restringem o acesso a ele pela Internet pública.

Para configurar sua política Amazon S3 de bucket

1. Abra o console do Amazon S3 em https://console.aws.amazon.com/s3/.
2. No painel de navegação, escolha Buckets.
3. Escolha o nome do bucket do Amazon S3 com o qual você deseja sincronizar. Amazon Kendra
4. Escolha a guia Permissões, role para baixo até a política do Bucket e clique em Editar.
5. Adicione ou modifique sua política de bucket para permitir acesso somente do VPC endpoint que você criou.

A seguir há um exemplo de política de bucket. Substitua *bucket-name* e *vpce-id* pelo nome do bucket do Amazon S3 e pelo ID do endpoint do Amazon S3 que você anotou anteriormente.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::bucket-name/*",
      "Condition": {
        "StringNotEquals": {
          "aws:SourceVpce": "vpce-id"
        }
      }
    }
  ]
}
```

6. Selecione Save Changes (Salvar alterações).

Seu bucket do S3 agora está acessível somente a partir da VPC específica que você criou.

Etapa 3: criar um conector de fonte Amazon S3 de dados de teste

Para testar sua Amazon VPC configuração, crie um Amazon S3 conector. Em seguida, configure-a com a VPC que você criou seguindo as etapas descritas em. Amazon S3

Para valores de Amazon VPC configuração, escolha os valores que você criou durante este exemplo:

- Amazon VPC(VPC) — `kendra-s3-example-vpc`
- Sub-redes — `kendra-s3-example-subnet-private1-[availability zone]`
- Grupos de segurança — `s3-data-source-security-group`

Aguarde até que seu conector termine de criar. Depois que o Amazon S3 conector for criado, escolha Sincronizar agora para iniciar uma sincronização.

Pode levar de alguns minutos a várias horas para concluir a sincronização, dependendo de quantos documentos estão em seu Amazon S3 bucket. Para testar o exemplo, recomendamos que você faça upload de apenas alguns documentos em seu bucket do S3. Se sua configuração estiver correta, você eventualmente verá o status de Sincronização de Concluída.

Se você encontrar algum erro, consulte Solução de problemas de Amazon VPC conexão.

Amazon Kendra Rastreador da Web

Você pode usar o Amazon Kendra Web Crawler para rastrear e indexar páginas da Web.

Você só pode rastrear sites públicos ou internos de empresas que usam o protocolo de comunicação segura do Hypertext Transfer Protocol Secure (HTTPS). Um erro recebido durante o crawling pode indicar que o site está bloqueado para crawling. Para rastrear sites internos, você pode configurar um proxy da web. O proxy da web deve estar voltado para o público. Você também pode usar a autenticação para acessar e rastrear sites.

Ao selecionar sites para indexar, você precisa aderir à Política de uso aceitável da Amazon e a todos os outros termos da Amazon. Lembre-se de que você só deve usar o Amazon Kendra Web Crawler para indexar suas próprias páginas da Web ou páginas da Web que você tenha autorização para indexar. Para saber como impedir que o Amazon Kendra Web Crawler indexe seu (s) site (s), consulte. Configurando o arquivo do *robots.txt* para o Web Crawler do Amazon Kendra

> **Note**
>
> Abusar do Amazon Kendra Web Crawler para rastrear agressivamente sites ou páginas da web que você não possui não é considerado uso aceitável.

Amazon Kendra tem duas versões do web crawler conector. Os recursos suportados de cada versão incluem:

Amazon Kendra Conector Web Crawler v1.0/API WebCrawlerConfiguration

- Proxy da Web
- Filtros de inclusão/exclusão

Amazon Kendra Conector Web Crawler v2.0/API TemplateConfiguration

- Mapeamentos de campos
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Proxy da Web
- Autenticação básica, NTLM/Kerberos, SAML e de formulários para os sites
- Nuvem privada virtual (VPC)

> **Important**
>
> A criação do conector Web Crawler v2.0 não é suportada pelo. AWS CloudFormation Use o conector Web Crawler v1.0 se precisar de suporte. AWS CloudFormation

Para solucionar problemas do conector da fonte de dados do Amazon Kendra web crawler, consulte. Solucionar problemas de origens de dados

Tópicos

- Amazon Kendra Conector Web Crawler v1.0
- Amazon Kendra Conector Web Crawler v2.0
- Configurando o arquivo do robots.txt para o Web Crawler do Amazon Kendra

Amazon Kendra Conector Web Crawler v1.0

Você pode usar o Amazon Kendra Web Crawler para rastrear e indexar páginas da Web.

Você só pode rastrear sites públicos e sites que usam o protocolo de comunicação segura do Hypertext Transfer Protocol Secure (HTTPS). Um erro recebido durante o crawling pode indicar que o site está bloqueado para crawling. Para rastrear sites internos, você pode configurar um proxy da web. O proxy da web deve estar voltado para o público.

Ao selecionar sites para indexar, você precisa aderir à Política de uso aceitável da Amazon e a todos os outros termos da Amazon. Lembre-se de que você só deve usar o Amazon Kendra Web Crawler para indexar suas próprias páginas da Web ou páginas da Web que você tenha autorização para indexar. Para saber como impedir que o Amazon Kendra Web Crawler indexe seu (s) site (s), consulte. Configurando o arquivo do *robots.txt* para o Web Crawler do Amazon Kendra

> **Note**
>
> Abusar do Amazon Kendra Web Crawler para rastrear agressivamente sites ou páginas da web que você não possui não é considerado uso aceitável.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra web crawler, consulte. Solucionar problemas de origens de dados

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

- Proxy da Web
- Filtros de inclusão/exclusão

Pré-requisitos

Antes de usar Amazon Kendra para indexar seus sites, verifique os detalhes de seus sites e AWS contas.

Para os sites, verifique se você tem:

- Copiou os URLs semente ou mapa dos sites nos quais você deseja fazer o crawling.
- Para sites que exigem autenticação básica: anotou o nome de usuário e a senha e copiou o nome do host do site e o número da porta.
- Opcional: copiou o nome do host do site e o número da porta se quiser usar um proxy da web para se conectar aos sites internos que você deseja rastrear. O proxy da web deve estar voltado para o público. O Amazon Kendra suporta a conexão com servidores proxy da web que são apoiados pela autenticação básica ou você pode se conectar sem autenticação.
- Verificou se cada documento de página da Web que você deseja indexar é único e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você

deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

Em sua AWS conta, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Para sites que exigem autenticação, ou se estiverem usando um proxy da web com autenticação, armazenaram suas credenciais de autenticação em um AWS Secrets Manager segredo e, se estiverem usando a API, anotaram o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de web crawler dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de web crawler dados, você deve fornecer os detalhes necessários da sua fonte de web crawler dados para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou web crawler para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a web crawler

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha conector do web crawler e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector do web crawler com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. Em Fonte, escolha entre URLs de origem e mapas de sites de origem, dependendo do caso de uso, e insira os valores de cada um.

      Você pode adicionar até 10 URLs de origem e 3 mapas de sites.

> **Note**
>
> Se quiser rastrear um sitemap, verifique se o URL base ou raiz é o mesmo que os URLs listados na página do mapa do site. Por exemplo, se o URL do mapa do site para https://example.com/sitemap-page.html, os URLs listados nessa página do mapa do site também devem usar o URL base "https://example.com/".

b. (Opcional) Para o proxy da Web, insira as seguintes informações:

   i. Nome do host: o nome do host em que o proxy da web é necessário.

   ii. Número da porta: o número da porta usado pelo protocolo de transporte de URL do host. O número da porta deve ser um valor numérico entre 0 e 65535.

   iii. Para credenciais de proxy da Web: se a conexão do proxy da Web exigir autenticação, escolha uma senha existente ou crie uma nova senha para armazenar as credenciais de autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

   iv. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager Secrets Manager :

      A. Senha: um nome para sua senha. O prefixo “AmazonKendra-WebCrawler-” é adicionado automaticamente à senha.

      B. Em Nome de usuário e senha, insira essas credenciais básicas de autenticação para seus sites.

      C. Escolha Salvar.

c. (Opcional) Hosts com autenticação: selecione para adicionar outros hosts com autenticação.

d. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

e. Escolha Próximo.

7. Na página Configurações de sincronização, insira as seguintes informações:

    a. Intervalo de rastreamento: escolha o tipo de página da Web que você deseja rastrear.

    b. Profundidade do rastreamento — Selecione o número de níveis do URL inicial que Amazon Kendra devem ser rastreados.

    c. As configurações avançadas de rastreamento e a configuração adicional inserem as seguintes informações:

        i. Tamanho máximo do arquivo: o tamanho máximo da página da Web ou do anexo a ser rastreado. Mínimo de 0,000001 MB (1 byte). Máximo de 50 MB.

        ii. Máximo de links por página: o número máximo de links rastreados por página. Os links passam pelo crawling por ordem de exibição. Mínimo de 1 link/página. Máximo de 1000 links/página.

        iii. Controle de utilização máxima: o número máximo de URLs que o crawling percorre por nome de host por minuto. Mínimo de 1 URL/nome do host/minuto. Máximo de 300 URLs/nome do host/minuto.

        iv. Padrões Regex: adicionar padrões de expressão regular para incluir ou excluir determinados URL. Você pode adicionar até 100 padrões.

    d. Em Cronograma de execução da sincronização, em Frequência — Escolha com que frequência Amazon Kendra será sincronizada com sua fonte de dados.

    e. Escolha Próximo.

8. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a web crawler

Você deve especificar o seguinte usando a WebCrawlerConfigurationAPI:

- URLs: especifica os URLs semente, ou de partida, ou os URLs de mapa dos sites nos quais você deseja fazer o crawling usando SeedUrlConfiguration e SiteMapsConfiguration.

> **Note**
>
> Se quiser rastrear um sitemap, verifique se o URL base ou raiz é o mesmo que os URLs listados na página do mapa do site. Por exemplo, se o URL do mapa do site para https://example.com/sitemap-page.html, os URLs listados nessa página do mapa do site também devem usar o URL base "https://example.com/".

- Nome do recurso da Amazon (ARN) da senha: se for necessário usar a autenticação básica do site, forneça o nome do host, o número da porta e uma senha que armazena as credenciais básicas de autenticação do seu nome de usuário e senha. Você fornece o ARN secreto usando a API AuthenticationConfiguration. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "username": "user name",
    "password": "password"
}
```

  Você também pode fornecer credenciais de proxy da web usando uma senha do AWS Secrets Manager . Você usa a API ProxyConfiguration para fornecer o nome do host e o número da porta do site e, opcionalmente, a senha que armazena as credenciais de proxy da web.
- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do web crawler e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do web crawler.

Você também pode adicionar os seguintes recursos opcionais:

- Modo de rastreamento: escolha se deseja rastrear somente nomes de host de sites ou nomes de host com subdomínios ou também rastrear outros domínios aos quais as páginas da Web estão vinculadas.
- A “profundidade” ou número de níveis do nível semente para crawling. Por exemplo, a página de URL semente tem profundidade 1 e todos os hiperlinks nessa página que também são rastreados têm profundidade 2.
- O número máximo de URLs em uma página da Web a serem incluídos no crawling.
- O tamanho máximo (em MB) de uma página da Web para crawling.

- O número máximo de URLs que o crawling percorre por host de site por minuto.
- O host do proxy da web e o número da porta para se conectar e rastrear sites internos. Por exemplo, o nome do host https://a.example.com/page1.html é "a.example.com" e o número da porta é 443, a porta padrão para HTTPS. Se o proxy da Web exigir credenciais para se conectar ao host de um site, crie um AWS Secrets Manager que armazene as credenciais.
- As informações de autenticação para acessar e fazer o crawling de sites que exigem autenticação do usuário.
- Você pode extrair metatags HTML como campos usando a ferramenta de Enriquecimento de documentos personalizados. Para obter mais informações, consulte Personalização de metadados de documentos durante o processo de ingestão. Para obter um exemplo de extração de metatags HTML, consulte exemplos de CDE.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados URLs..

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte web crawler de dados, consulte:

- Reimagine a descoberta de conhecimento usando o Web Amazon Kendra Crawler

Amazon Kendra Conector Web Crawler v2.0

Você pode usar o Amazon Kendra Web Crawler para rastrear e indexar páginas da Web.

Você só pode rastrear sites públicos ou internos de empresas que usam o protocolo de comunicação segura do Hypertext Transfer Protocol Secure (HTTPS). Um erro recebido durante o crawling pode indicar que o site está bloqueado para crawling. Para rastrear sites internos, você pode configurar

um proxy da web. O proxy da web deve estar voltado para o público. Você também pode usar a autenticação para acessar e rastrear sites.

Amazon Kendra O Web Crawler v2.0 usa o pacote Selenium web crawler e um driver Chromium. Amazon Kendra atualiza automaticamente a versão do Selenium e do driver Chromium usando a Integração Contínua (CI).

Ao selecionar sites para indexar, você precisa aderir à Política de uso aceitável da Amazon e a todos os outros termos da Amazon. Lembre-se de que você só deve usar o Amazon Kendra Web Crawler para indexar suas próprias páginas da Web ou páginas da Web que você tenha autorização para indexar. Para saber como impedir que o Amazon Kendra Web Crawler indexe seu (s) site (s), consulte. Configurando o arquivo do *robots.txt* para o Web Crawler do Amazon Kendra . Abusar do Amazon Kendra Web Crawler para rastrear agressivamente sites ou páginas da web que você não possui não é considerado uso aceitável.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra web crawler, consulte. Solucionar problemas de origens de dados

> **Note**
>
> O conector Web Crawler v2.0 não suporta o rastreamento de listas de sites da Web a partir de buckets criptografados. AWS KMS Amazon S3 Ele suporta somente criptografia do lado do servidor com chaves gerenciadas Amazon S3 .

> **Important**
>
> A criação do conector Web Crawler v2.0 não é suportada pelo. AWS CloudFormation Use o conector Web Crawler v1.0 se precisar de suporte. AWS CloudFormation

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão

Atributos compatíveis

- Mapeamentos de campos
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Proxy da Web
- Autenticação básica, NTLM/Kerberos, SAML e de formulários para os sites
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de usar Amazon Kendra para indexar seus sites, verifique os detalhes de seus sites e AWS contas.

Para os sites, verifique se você tem:

- Copiou os URLs semente ou mapa dos sites nos quais você deseja fazer o crawling. Você pode armazenar os URLs em um arquivo de texto e enviá-los para um bucket do Amazon S3 . Cada URL no arquivo de texto deve ser formatado em uma linha separada. Se você quiser armazenar seus sitemaps em um Amazon S3 bucket, certifique-se de ter copiado o XML do sitemap e salvado em um arquivo XML. Você também pode agrupar vários arquivos XML de mapa do site em um arquivo ZIP.

  > **Note**
  >
  > (Local/servidor) Amazon Kendra verifica se as informações do endpoint incluídas são iguais às informações do endpoint especificadas nos AWS Secrets Manager detalhes de configuração da fonte de dados. Isso ajuda a proteger contra o problema de assistência confusa, que é um problema de segurança em que um usuário não tem permissão para realizar uma ação, mas usa o Amazon Kendra como proxy para acessar a senha configurada e realizar a ação. Se você alterar posteriormente as informações do endpoint, crie uma nova senha para sincronizar essas informações.

- Para sites que exigem autenticação básica, NTLM ou Kerberos:
  - Anote suas credenciais de autenticação do site, que incluem um nome de usuário e senha.

> **Note**
>
> Amazon Kendra O Web Crawler v2.0 suporta o protocolo de autenticação NTLM, que inclui hash de senha, e o protocolo de autenticação Kerberos, que inclui criptografia de senha.

- Para sites que exigem autenticação por SAML ou formulário de login:
  - Anote suas credenciais de autenticação do site, que incluem um nome de usuário e senha.
  - Copiou o XPaths (XML Path Language) do campo do nome do usuário (e o botão do nome do usuário se estiver usando SAML), do campo e do botão da senha e copiou o URL da página de login. Você pode encontrar os XPaths dos elementos usando as ferramentas de desenvolvedor do navegador da Web. Os XPaths geralmente seguem este formato:`//tagname[@Attribute='Value']`.

> **Note**
>
> Amazon Kendra O Web Crawler v2.0 usa um navegador Chrome sem cabeçalho e as informações do formulário para autenticar e autorizar o acesso com um URL protegido pelo OAuth 2.0.

- Opcional: copiou o nome do host e o número da porta se quiser usar um servidor do proxy da web para se conectar aos sites internos que você deseja rastrear. O proxy da web deve estar voltado para o público. Amazon Kendra suporta a conexão com servidores proxy da web que são apoiados pela autenticação básica ou você pode se conectar sem autenticação.
- Opcional: copiou o ID da sub-rede da nuvem privada virtual (VPC) se você quiser usar uma VPC para se conectar aos sites internos que deseja rastrear. Para obter mais informações, consulte Configurando um Amazon VPC.
- Verificou se cada documento de página da Web que você deseja indexar é único e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

Em sua AWS conta, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.

- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o nome de recurso da Amazon da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Para sites que exigem autenticação, ou se estiverem usando um proxy da web com autenticação, armazenaram suas credenciais de autenticação em um AWS Secrets Manager segredo e, se estiverem usando a API, anotaram o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de web crawler dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de web crawler dados, você deve fornecer os detalhes necessários da sua fonte de web crawler dados para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou web crawler para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a web crawler

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha conector do web crawler e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector do web crawler com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Fonte : escolha os URLs de fonte, mapas de sites de fonte, arquivo de URLs de fonte e arquivo de mapas de site de fonte. Se você optar por usar um arquivo de texto que inclua uma lista de até 100 URLs iniciais, especifique o caminho para o Amazon S3 bucket em que seu arquivo está armazenado. Se você optar por usar um arquivo XML de mapa do site, especifique o caminho para o bucket do Amazon S3 em que o arquivo está armazenado. Você também pode agrupar vários arquivos XML de mapa do site em um arquivo ZIP. Caso contrário, você pode inserir manualmente até 10 URLs semente ou de partida e até 3 URLs de mapa.

   > **Note**
   >
   > Se quiser rastrear um sitemap, verifique se o URL base ou raiz é o mesmo que os URLs listados na página do mapa do site. Por exemplo, se o URL do mapa do

site para https://example.com/sitemap-page.html, os URLs listados nessa página do mapa do site também devem usar o URL base "https://example.com/".

Se os sites precisarem de autenticação para acessar os sites, você poderá escolher a autenticação básica, NTLM/Kerberos, SAML ou de formulário. Caso contrário, escolha a opção sem autenticação.

> **Note**
>
> Se você quiser editar posteriormente a fonte de dados para alterar os URLs iniciais com autenticação em mapas de site, você deve crie uma nova fonte de dados. O Amazon Kendra configura a fonte de dados usando as informações do endpoint dos URLs iniciais na senha do Secrets Manager para autenticação e, portanto, não pode reconfigurar a fonte de dados ao mudar para os mapas do site.

- AWS Secrets Manager segredo — Se seus sites precisarem da mesma autenticação para acessar os sites, escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar as credenciais do seu site. Se você optar por criar um novo segredo, uma janela AWS Secrets Manager secreta será aberta.

  Se escolher a Autenticação básica ou NTML/Kerberos, insira um nome para o segredo o nome de usuário e a senha. O protocolo de autenticação NTLM inclui hash de senha e o protocolo de autenticação Kerberos inclui criptografia de senha.

  Se escolher a Autenticação Formulário ou SAML, insira um nome para o segredo o nome de usuário e a senha. Use XPath para o campo de nome de usuário (e XPath para o botão de nome de usuário se estiver usando SAML). Use XPaths para o campo e botão de senha e URL da página de login. Você pode encontrar os XPaths (XML Path Language) dos elementos usando as ferramentas de desenvolvedor do navegador. Os XPaths geralmente seguem este formato: `//tagname[@Attribute='Value']`.

b. (Opcional) Proxy da Web: insira o nome do host e o número da porta do servidor proxy que deseja usar para se conectar aos sites internos. Por exemplo, o nome do host https://a.example.com/page1.html é "a.example.com" e o número da porta é 443, a

porta padrão para HTTPS. Se forem necessárias credenciais de proxy da web para se conectar a um host de site, você poderá criar uma AWS Secrets Manager que armazene as credenciais.

c. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

d. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

e. Escolha Próximo.

7. Na página Configurações de sincronização, insira as seguintes informações:

   a. Escopo de sincronização: defina limites para rastrear páginas da Web, incluindo domínios, tamanhos de arquivo e links, e filtre URLs usando padrões de regex.

      i. (Opcional) Intervalo de domínios de rastreamento: escolha se deseja rastrear somente domínios de sites, domínios com subdomínios ou também rastrear outros domínios aos quais as páginas da Web estão vinculadas. Por padrão, rastreia Amazon Kendra apenas os domínios dos sites que você deseja rastrear.

      ii. (Opcional) Configurações adicionais: defina as seguintes configurações opcionais:

      - Profundidade do crawling: a “profundidade” ou número de níveis do nível semente para crawling. Por exemplo, a página de URL semente tem profundidade 1 e todos os hiperlinks nessa página que também são rastreados têm profundidade 2.
      - Tamanho máximo do arquivo: o tamanho máximo em MB da página da Web ou do anexo a ser rastreado.
      - Máximo de links por página: o número máximo de URLs em uma página da Web a serem incluídos no crawling.
      - Controle de utilização e velociade de crawling máximos: o número máximo de URLs que o crawling percorre por nome de host por minuto.

- Arquivos: escolha para rastrear arquivos aos quais as páginas da web estão vinculadas.
- URLs para indexar e rastrear: uma lista de padrões de expressão regular que inclui o crawling de determinados URLs e a indexação de quaisquer hiperlinks nessas páginas da Web com URL.

b. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

c. Cronograma de execução da sincronização: em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

d. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos padrão Amazon Kendra gerados de páginas da Web e arquivos que você deseja mapear para o seu índice.

b. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a web crawler

Você deve especificar um JSON do esquema da fonte de dados usando a API TemplateConfiguration. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `WEBCRAWLERV2` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- URLs: especifica os URLs semente, ou de partida, ou os URLs de mapa dos sites nos quais você deseja fazer o crawling usando e . Você pode especificar o caminho para um Amazon S3 bucket que armazena sua lista de URLs iniciais. Cada URL no arquivo de texto para os URLS semente deve ser formatado em uma linha separada. Você também pode especificar o caminho para um Amazon S3 bucket que armazena os arquivos XML do seu sitemap. Você pode agrupar vários arquivos do mapa do site em um arquivo ZIP e armazená-lo em seu bucket do Amazon S3 .

  > **Note**
  >
  > Se quiser rastrear um sitemap, verifique se o URL base ou raiz é o mesmo que os URLs listados na página do mapa do site. Por exemplo, se o URL do mapa do site para https://example.com/sitemap-page.html, os URLs listados nessa página do mapa do site também devem usar o URL base "https://example.com/".

- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Autenticação: se os sites exigirem a mesma autenticação, especifique a autenticação `BasicAuth`, `NTLM_Kerberos`, `SAML` ou `Form`. Se os sites não precisarem de autenticação, especifique `NoAuthentication`.

- Nome do recurso da Amazon (ARN) da senha: se os sites exigirem autenticação básica, NTLM ou Kerberos, você fornecerá uma senha que armazena as credenciais de autenticação de nome de usuário e senha. Forneça o nome do recurso da Amazon (ARN) da senha do AWS Secrets Manager . A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "seedUrlsHash": "Hash representation of all seed URLs",
    "userName": "user name",
    "password": "password"
}
```

Se o site precisar de uma autenticação OAuth2, a senha deverá conter uma estrutura JSON com as seguintes chaves:

```
{
    "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"
}
```

Se o site precisar de uma autenticação, a senha deverá conter uma estrutura JSON com as seguintes chaves:

```
{
    "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"
}
```

Você pode encontrar os XPaths (XML Path Language) dos elementos usando as ferramentas de desenvolvedor do navegador. Os XPaths geralmente seguem este formato: `//tagname[@Attribute='Value']`.

Você também pode fornecer credenciais de proxy da web usando uma senha do AWS Secrets Manager .

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do web crawler e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do web crawler.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Faixa de domínio: escolha se deseja rastrear somente domínios de sites, domínios com subdomínios ou também rastrear outros domínios aos quais as páginas da Web estão vinculadas. Por padrão, rastreia Amazon Kendra apenas os domínios dos sites que você deseja rastrear.
- A “profundidade” ou número de níveis do nível semente para crawling. Por exemplo, a página de URL semente tem profundidade 1 e todos os hiperlinks nessa página que também são rastreados têm profundidade 2.
- O número máximo de URLs em uma página da Web a serem incluídos no crawling.
- O tamanho máximo em MB de uma página da Web ou anexo para crawling.
- O número máximo de URLs que o crawling percorre por host de site por minuto.
- O host do proxy da web e o número da porta para se conectar e rastrear sites internos. Por exemplo, o nome do host https://a.example.com/page1.html é "a.example.com" e o número da porta é 443, a porta padrão para HTTPS. Se o proxy da Web exigir credenciais para se conectar ao host de um site, crie um AWS Secrets Manager que armazene as credenciais.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir o rastreamento de determinados URLs e a indexação de quaisquer hiperlinks nessas páginas da Web com URL.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Mapeamentos de campo — Escolha mapear os campos de páginas da Web e arquivos de páginas da Web para seus Amazon Kendra campos de índice. Para obter mais informações, consulte Mapear campos de fonte de dados.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o Esquema do modelo do Web Crawler do Amazon Kendra.

Configurando o arquivo do **`robots.txt`** para o Web Crawler do Amazon Kendra

Amazon Kendra é um serviço de pesquisa inteligente que AWS os clientes usam para indexar e pesquisar documentos de sua escolha. Para indexar documentos na web, os clientes podem usar o Amazon Kendra Web Crawler, indicando quais URLs devem ser indexados e outros parâmetros operacionais. Amazon Kendra os clientes precisam obter autorização antes de indexar qualquer site específico.

Amazon Kendra O Web Crawler respeita as diretivas padrão do robots.txt, como e. `Allow` `Disallow` Você pode modificar o `robots.txt` arquivo do seu site para controlar como o Amazon Kendra Web Crawler rastreia seu site.

Configurando como o Amazon Kendra Web Crawler acessa seu site

Você pode controlar como o Amazon Kendra Web Crawler indexa o uso `Allow` e as diretrizes do seu site. `Disallow` Você também pode controlar quais páginas da Web são indexadas e quais páginas da Web não são rastreadas.

Para permitir que o Amazon Kendra Web Crawler rastreie todas as páginas da Web, exceto páginas da Web não permitidas, use a seguinte diretiva:

```
User-agent: amazon-kendra    # Amazon Kendra Web Crawler
```

```
Disallow: /credential-pages/ # disallow access to specific pages
```

Para permitir que o Amazon Kendra Web Crawler rastreie somente páginas da Web específicas, use a seguinte diretiva:

```
User-agent: amazon-kendra    # Amazon Kendra Web Crawler
Allow: /pages/ # allow access to specific pages
```

Para permitir que o Amazon Kendra Web Crawler rastreie todo o conteúdo do site e proibir o rastreamento de outros robôs, use a seguinte diretiva:

```
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
```

Impedindo que o Amazon Kendra Web Crawler rastreie seu site

Você pode impedir que o Amazon Kendra Web Crawler indexe seu site usando a diretiva. `Disallow` Você também pode controlar quais páginas da Web são rastreadas ou não.

Para impedir que o Amazon Kendra Web Crawler rastreie o site, use a seguinte diretiva:

```
User-agent: amazon-kendra # Amazon Kendra Web Crawler
Disallow: / # disallow access to any pages
```

Amazon Kendra O Web Crawler também suporta robôs `noindex` e `nofollow` diretivas em metatags em páginas HTML. Essas diretivas impedem que o rastreador da Web indexe uma página da Web e pare de seguir qualquer link na página da Web. Você coloca as metatags na seção do documento para especificar as regras das regras dos robôs.

Por exemplo, a página da web abaixo inclui as diretivas robôs `noindex` e`nofollow`:

```
<html>
<head>
    <meta name="robots" content="noindex, nofollow"/>
    ...
</head>
```

```
<body>...</body>
</html>
```

Se você tiver alguma dúvida ou preocupação em relação ao Amazon Kendra Web Crawler, entre em contato com a equipe de AWS suporte.

Amazon WorkDocs

Amazon WorkDocs é um serviço seguro de colaboração de conteúdo para criar, editar, armazenar e compartilhar conteúdo. Você pode usar Amazon Kendra para indexar sua fonte Amazon WorkDocs de dados.

Você pode se conectar Amazon Kendra à sua fonte de Amazon WorkDocs dados usando o Amazon Kendra console e a WorkDocsConfigurationAPI.

Amazon WorkDocs está disponível nas regiões de Oregon, Virgínia do Norte, Sydney, Cingapura e Irlanda.

Para solucionar problemas do conector da fonte de Amazon Kendra WorkDocs dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra WorkDocs o conector de fonte de dados oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Log de alterações

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de WorkDocs dados, faça essas alterações em suas WorkDocs AWS contas.

Em WorkDocs, verifique se você tem:

- Anote o ID do Amazon WorkDocs diretório (ID da organização) do seu Amazon WorkDocs repositório.
- Verifique se cada documento é exclusivo em WorkDocs e entre outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

Em sua AWS conta, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

Se você não tiver uma IAM função existente, poderá usar o console para criar uma nova IAM função ao conectar sua fonte de WorkDocs dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função existente e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de WorkDocs dados, você deve fornecer os detalhes necessários da sua fonte de WorkDocs dados para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou WorkDocs para Amazon Kendra, consulte Pré-requisitos.

Console

Para se conectar Amazon Kendra a Amazon WorkDocs

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha WorkDocs conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o WorkDocs conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. ID da organização específica Amazon WorkDocs do seu site — Selecione o ID do Amazon WorkDocs site que você deseja indexar. Você já deve ter criado um site.
   b. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

      > **Note**
      >
      > IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

   c. Escolha Próximo.
7. Na página Configurações de sincronização, insira as seguintes informações:

a. Rastrear comentários do documento: as entidades do Amazon WorkDocs ou tipos de conteúdo que você deseja rastrear.

b. Usar registros de alterações — Selecione para atualizar seu índice somente com conteúdo novo ou modificado, em vez de sincronizar todos os seus arquivos.

c. Padrões Regex: os padrões de expressão regular para incluir ou excluir determinados arquivos.

d. Em Sincronização, cronograma de execução para frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Amazon WorkDocs

Você deve especificar o seguinte usando a WorkDocsConfigurationAPI:

- Amazon WorkDocs ID do diretório — Especifique o ID da organização do seu Amazon WorkDocs diretório. Você pode encontrar o ID da organização no AWS Directory Service: vá para Active Directory e, em seguida, para Diretórios.
- Função do IAM — especifique `RoleArn` quando você chama `CreateDataSource` para fornecer uma IAM função com permissões para acessar o WorkDocs diretório e chamar as APIs públicas necessárias para o conector e. WorkDocs Amazon Kendra Para obter mais informações, consulte Funções do IAM para fontes de WorkDocs dados.

Você também pode adicionar os seguintes recursos opcionais:

- Registro de alterações — Se Amazon Kendra deve usar o mecanismo de registro de alterações da fonte de WorkDocs dados para determinar se um documento deve ser atualizado no índice.

  > **Note**
  >
  > Use o log de alterações se o Amazon Kendra não quiser digitalizar todos os documentos. Se o registro de alterações for grande, talvez leve Amazon Kendra menos tempo para digitalizar os documentos na fonte de WorkDocs dados do que para processar o registro de alterações. Se você estiver sincronizando sua fonte de WorkDocs dados com seu índice pela primeira vez, todos os documentos serão digitalizados.

- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados documentos e comentários de documentos. Cada comentário é indexado como um documento separado.

  > **Note**
  >
  > A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo — Escolha mapear os campos da fonte de WorkDocs dados para os Amazon Kendra campos de índice. Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte WorkDocs de dados, consulte:

- Comece a usar o WorkDocs conector Amazon Kendra Amazon

Box

O Box é um serviço de armazenamento em nuvem que oferece recursos de hospedagem de arquivos. Você pode usar Amazon Kendra para indexar o conteúdo do seu Box, incluindo comentários, tarefas e links da web.

Você pode se conectar Amazon Kendra à sua fonte de dados do Box usando o Amazon Kendra console e a BoxConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Box, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra O conector de fonte de dados Box oferece suporte aos seguintes recursos:

- Mapeamentos de campos

- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Registro de alterações, sincronizações completas e incrementais de conteúdo
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Box, faça essas alterações no Box e AWS nas contas.

No Box, verifique se você:

- Tem uma conta Box Enterprise ou Box Enterprise Plus.
- Configurei um aplicativo personalizado do Box no Box Developer Console, com autenticação do lado do servidor usando JSON Web Tokens (JWT). Consultou a Documentação do Box sobre a criação de um aplicativo personalizado e a Documentação do Box sobre a configuração do JWT Auth para obter mais detalhes.
- Definiu seu nível de acesso ao Aplicativo e ao Enterprise Access e permita que ele Faça chamadas de API usando o cabeçalho como usuário.
- Usou o usuário administrador para adicionar os seguintes Escopos de aplicativo no aplicativo Box:
  - Gravou todos os arquivos e pastas armazenados no Box
  - Gerenciar usuários
  - Gerenciar grupos
  - Gerenciar propriedades corporativas
- Par de chaves pública/privada configurado, incluindo um ID do cliente, um segredo do cliente, um ID de chave pública, um ID de chave privada, uma frase secreta e um ID corporativo para usar como suas credenciais de autenticação. Consulte Par de chaves públicas e privadas para obter mais detalhes.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Copiou a ID do Enterprise do Box das configurações do Box Developer Console ou do aplicativo Box. Por exemplo, *801234567*.
- Verifique se cada documento é exclusivo no Box e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Box em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Box Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Box, você deve fornecer os detalhes necessários da sua fonte de dados do Box para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Box para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao Box

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector Box e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Box com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. ID do Box Enterprise: insira seu ID do Box Enterprise. Por exemplo, *801234567*.

b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

c. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Box. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

   i. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Box-' é adicionado automaticamente ao seu nome secreto.

   ii. Para ID do cliente, segredo do cliente, ID da chave pública, ID da chave privada e frase secreta, insira os valores da chave pública/privada que você configurou no Box.

   iii. Adicione e salve seu segredo.

d. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

e. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

f. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

g. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. IDs da pasta Box — Insira determinados IDs da pasta Box que você deseja rastrear, caso contrário, o conteúdo de todas as pastas será rastreado.

   b. Arquivos Box — escolha se deseja rastrear links da Web, comentários e tarefas.

   c. Para configuração adicional — Adicione padrões de expressão regular para incluir ou excluir determinado conteúdo.

   d. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

      - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
      - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
      - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

   e. Em Sincronização, cronograma de execução para frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

   f. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Box

Você deve especificar o seguinte usando a BoxConfigurationAPI:

ID do Box Enterprise: insira seu ID do Box Enterprise. Você pode encontrar a ID corporativa nas configurações do Box Developer Console ou ao configurar um aplicativo no Box.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta Box. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "clientID": "client-id",
    "clientSecret": "client-secret",
    "publicKeyID": "public-key-id",
    "privateKey": "private-key",
    "passphrase": "pass-phrase"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Box e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para fontes de dados do Box.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique `VpcConfiguration` como parte da configuração da fonte de dados. Consulte Configuração do Amazon Kendra para usar uma VPC.
- Registro de alterações — Se Amazon Kendra deve usar o mecanismo de registro de alterações da fonte de dados Box para determinar se um documento deve ser atualizado no índice.

> **Note**
>
> Use o log de alterações se o Amazon Kendra não quiser digitalizar todos os documentos. Se o registro de alterações for grande, talvez leve Amazon Kendra menos tempo para digitalizar os documentos na fonte de dados do Box do que para processar o registro de alterações. Se estiver sincronizando a fonte de dados do Box com o índice pela primeira vez, todos os documentos serão digitalizados.

- Comentários, tarefas, links da web — especifique se esses tipos de conteúdo devem ser rastreados.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Filtros de inclusão e exclusão — especifique se deseja incluir ou excluir determinados arquivos e pastas do Box.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Box para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Box, consulte:

- Introdução ao conector Amazon Kendra Box

Confluence

O Confluence é uma ferramenta colaborativa de gerenciamento de trabalho projetada para compartilhar, armazenar e trabalhar no planejamento de projetos, no desenvolvimento de software e no gerenciamento de produtos. Você pode usar Amazon Kendra para indexar seus espaços, páginas (incluindo páginas aninhadas), blogs e comentários e anexos em páginas e blogs indexados.

Amazon Kendra é compatível com o Confluence Server/Data Center e o Confluence Cloud.

> **Note**
>
> Por padrão, Amazon Kendra não indexa arquivos e espaços pessoais do Confluence. Você pode optar por indexá-los ao criar a fonte de dados. Se você não quiser Amazon Kendra indexar um espaço, marque-o como privado no Confluence.

Você pode se conectar Amazon Kendra à sua fonte de dados do Confluence usando o Amazon Kendra console, a TemplateConfigurationAPI ou a ConfluenceConfigurationAPI.

Amazon Kendra tem duas versões do conector Confluence. Os recursos suportados de cada versão incluem:

Conector Confluence V1.0/API ConfluenceConfiguration

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- (Somente para o Confluence Server) Nuvem privada virtual (VPC)

Conector Confluence V2.0/API TemplateConfiguration

- Mapeamentos de campos
- Controle de acesso do usuário
- Padrões de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

> **Note**
>
> O suporte para o conector V1.0 ConfluenceConfiguration /API do Confluence está programado para terminar em 2023. Recomendamos migrar ou usar o conector V2.0/API do Confluence. TemplateConfiguration

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Confluence, consulte. Solucionar problemas de origens de dados

Tópicos

- Conector o Confluence v1.0
- Conector Confluence v2.0

Conector o Confluence v1.0

O Confluence é uma ferramenta colaborativa de gerenciamento de trabalho projetada para compartilhar, armazenar e trabalhar no planejamento de projetos, no desenvolvimento de software e no gerenciamento de produtos. Você pode usar o Amazon Kendra para indexar espaços, páginas (incluindo páginas aninhadas), blogs e comentários e anexos em páginas e blogs indexados.

> **Note**
>
> O suporte para o conector V1.0 ConfluenceConfiguration /API do Confluence está programado para terminar em 2023. Recomendamos migrar ou usar o conector V2.0/API do Confluence. TemplateConfiguration

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Confluence, consulte. Solucionar problemas de origens de dados

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra O conector de fonte de dados do Confluence oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- (Somente para o Confluence Server) Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Confluence, faça essas alterações no Confluence e nas contas. AWS

No Confluence, verifique se você tem:

- Amazon Kendra Permissões concedidas para visualizar todo o conteúdo em sua instância do Confluence por:
  - Tornando Amazon Kendra um membro do `confluence-administrators` grupo.
  - Conceder permissões de administrador do site para todos os espaços, blogs e páginas existentes.
- Copiar o URL da sua instância do Confluence.
- Para usuários de SSO (autenticação única): ative Exibir na página de login para o nome de usuário e a senha ao configurar os Métodos de autenticação do Confluence no Confluence Data Center.
- Para o Confluence Server
  - Anote suas credenciais básicas de autenticação contendo o nome de usuário e a senha da conta administrativa do Confluence para se conectar ao Amazon Kendra.

    > **Note**
    >
    > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

  - Opcional: gerou um token de acesso pessoal na sua conta do Confluence para se conectar ao Amazon Kendra. Para obter mais informações, consulte a documentação do Confluence sobre a geração de tokens de acesso pessoal.
- Para o Confluence Cloud
  - Anote suas credenciais básicas de autenticação contendo o nome de usuário e a senha da conta administrativa do Confluence para se conectar ao Amazon Kendra.
- Verifique se cada documento é exclusivo no Confluence e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.

- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Confluence em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados do Confluence a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Confluence, você deve fornecer detalhes das suas credenciais do Confluence para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Confluence para Amazon Kendra ver. Pré-requisitos

Console

Para se conectar Amazon Kendra ao Confluence

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha o Conector do Confluence V1.0 e, em seguida, escolha Adicionar fonte de .dados.
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. Escolha entre o Confluence Cloud e o Confluence Server.
   b. Se você escolher o Confluence Cloud, insira as seguintes informações:
      i. URL do Confluence: seu URL do Confluence.
      ii. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Confluence. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.
         - Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :
            I. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Confluence-' é adicionado automaticamente ao seu nome secreto.

II. Em Nome de usuário e senha, insira seu nome de usuário e senha do Confluence.

III. Escolha Salvar autenticação.

c. Se você escolher o Confluence Server, insira as seguintes informações:

i. URL do Confluence: seu nome de usuário e senha do Confluence.

ii. (Opcional) Para o proxy da Web, insira as seguintes informações:

A. Nome do hos: nome do host da sua conta do Confluence.

B. número da porta: o número da publicação usado pelo protocolo de transporte de URL do host.

iii. Para Autenticação, escolha Autenticação básica ou Token de acesso pessoal (somente para o Confluence Server).

iv. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Confluence. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

- Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

I. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Confluence-' é adicionado automaticamente ao seu nome secreto.

II. Em Nome de usuário e senha, insira os valores da credencial de autenticação que você configurou no Confluence. Se estiver usando a autenticação básica, use seu nome de usuário (ID de e-mail) e senha (token de API) do Confluence. Se estiver usando o token de acesso pessoal, insira os detalhes do token de acesso pessoal que você configurou na conta do Confluence.

III. Salve e adicione seu segredo.

d. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

e. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Em Incluir espaços pessoais e Incluir espaços arquivados, escolha os tipos de espaço opcionais a serem incluídos nessa fonte de dados.

   b. Para configuração adicional, especifique padrões de expressão regular para incluir ou excluir determinado conteúdo. Você pode adicionar até 100 padrões.

   c. Você também pode optar por Rastrear anexos nos espaços escolhidos.

   d. Em Cronograma de execução da sincronização, em Frequência — Escolha com que frequência Amazon Kendra será sincronizada com sua fonte de dados.

   e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Para Espaço, Página, Blog — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados ou mapeamentos de campo adicionais sugeridos para adicionar campos de índice.

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Confluence

Você deve especificar o seguinte usando a ConfluenceConfigurationAPI:

- Versão do Confluence: especifique a versão da instância do Confluence usada: `CLOUD` ou `SERVER`.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha suas credenciais de autenticação do Confluence.

  Se você usa o Confluence Server, pode usar seu nome de usuário e senha do Confluence ou seu token de acesso pessoal como credenciais de autenticação.

  Se você usar seu nome de usuário e senha do Confluence como credenciais de autenticação, armazene as seguintes credenciais como uma estrutura JSON em seu segredo: Secrets Manager

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  Se você usa um token de acesso pessoal para se conectar ao Confluence Server Amazon Kendra, armazene as seguintes credenciais como uma estrutura JSON em seu segredo: Secrets Manager

  ```
  {
      "patToken": "personal access token"
  }
  ```

  Se você usa o Confluence Cloud, usa seu nome de usuário do Confluence e um token de API, configurado no Confluence, como sua senha. Você armazena as seguintes credenciais como uma estrutura JSON em seu Secrets Manager segredo:

  ```
  {
      "username": "user name",
      "password": "API token"
  }
  ```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do Confluence e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Confluence.

Você também pode adicionar os seguintes recursos opcionais:

- Proxy da Web: para conectar à instância de URL do Confluence por meio de um proxy da Web. Você pode usar essa opção para o Confluence Server.
- (Apenas para o Servidor Confluence) Nuvem privada virtual (VPC): especifique o `VpcConfiguration` como parte da configuração da fonte de dados. Consulte Configuração Amazon Kendra para usar uma VPC.
- Filtros de inclusão e exclusão: especifique padrões de expressões regulares para incluir ou excluir determinados espaços, publicações de blog, páginas, espaços e anexos. Se você optar por indexar anexos, somente os anexos das páginas indexadas e dos blogs serão indexados.

  > **Note**
  >
  > A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Confluence para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

  > **Note**
  >
  > O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Confluence, consulte:

- Configurando seu conector do Amazon Kendra Confluence Server

Conector Confluence v2.0

O Confluence é uma ferramenta colaborativa de gerenciamento de trabalho projetada para compartilhar, armazenar e trabalhar no planejamento de projetos, no desenvolvimento de software e no gerenciamento de produtos. Você pode usar o Amazon Kendra para indexar espaços, páginas (incluindo páginas aninhadas), blogs e comentários e anexos em páginas e blogs indexados.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Confluence, consulte. Solucionar problemas de origens de dados

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão

Atributos compatíveis

Amazon Kendra O conector de fonte de dados do Confluence oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Padrões de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Confluence, faça essas alterações no Confluence e nas contas. AWS

No Confluence, verifique se você tem:

- Copiar o URL da sua instância do Confluence. Por exemplo: *https://example.confluence.com*, *https://www.example.confluence.com/* ou *https://atlassian.net/*. O URL da instância do Confluence é necessário para se conectar ao Amazon Kendra.

  Se você estiver usando o Confluence Cloud, o URL do seu host deve terminar com atlassian.net/.

  > **Note**
  >
  > Os seguintes formatos de URL não são compatíveis:
  >
  > - *https://example.confluence.com/xyz*
  > - *https://www.example.confluence.com//wiki/spacekey/xxx*
  > - *https://atlassian.net/xyz*

  > **Note**
  >
  > (Local/servidor) Amazon Kendra verifica se as informações do endpoint incluídas são as mesmas especificadas nos AWS Secrets Manager detalhes de configuração da fonte de dados. Isso ajuda a proteger contra o problema de assistência confusa, que é um problema de segurança em que um usuário não tem permissão para realizar uma ação, mas usa o Amazon Kendra como proxy para acessar a senha configurada e realizar a ação. Se você alterar posteriormente as informações do endpoint, crie uma nova senha para sincronizar essas informações.

- Credenciais básicas de autenticação configuradas contendo um nome de usuário (ID de e-mail usado para fazer login no Confluence) e senha (token da API do Confluence como senha). Consulte Gerenciar tokens de API para sua conta Atlassian.

  > **Note**
  >
  > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não

recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Opcional: credenciais do OAuth 2.0 configuradas contendo uma chave do aplicativo Confluence, um segredo do aplicativo Confluence, um token de acesso do Confluence e um token de atualização do Confluence para permitir a conexão com sua instância do Confluence. Amazon Kendra Se o token de acesso expirar, você poderá usar o token de atualização para regenerar o token de acesso e o par de tokens de atualização. Ou você pode repetir o processo de autorização. Para obter mais informações sobre tokens de acesso, consulte Gerenciar tokens de acesso OAuth.
- (Somente para servidor/data center do Confluence) Opcional: configurou um token de acesso pessoal (PAT) no Confluence. Consulte Uso de tokens de acesso pessoal.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Confluence em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados do Confluence a.

Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Confluence, você deve fornecer os detalhes necessários da sua fonte de dados do Confluence para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Confluence para o Amazon Kendra , consulte Pré-requisitos.

Console

Para se conectar Amazon Kendra ao Confluence

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector do Confluence e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Confluence com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Em Source, escolha Confluence Cloud ou Confluence Server/Data Center.

   b. URL do Confluence insira o URL do host do Confluence. Por exemplo, *https://example.confluence.com*.

   c. (Somente para o Confluence Server/Data Center) Local do certificado SSL - opcional — Insira o Amazon S3 caminho para seu arquivo de certificado SSL para o Confluence Server.

   d. (Somente para o Confluence Server/Data Center) Proxy Web - opcional — Insira o nome do host do proxy web (sem o `https://` protocolo `http://` or) e o número da porta (porta usada pelo protocolo de transporte de URL do host). O número da porta deve ser um valor numérico entre 0 e 65535.

   e. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   f. Autenticação — escolha entre autenticação básica, autenticação Oauth 2.0 ou (somente para servidor/data center do Confluence) autenticação de token de acesso pessoal.

   g. Senha do AWS Secrets Manager : escolha uma senha existente ou crie uma nova senha do Secrets Manager para armazenar as credenciais do Confluence. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta. Insira as seguintes informações na janela:

      i. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Confluence-' é adicionado automaticamente ao seu nome secreto.

      ii. Se estiver usando a autenticação básica, insira o nome secreto, o nome de usuário e a senha (token da API do Confluence como senha) que você configurou no Confluence.

      Se estiver usando a autenticação OAuth2.0 — insira o nome secreto, a chave do aplicativo, o segredo do aplicativo, o token de acesso e o token de atualização que você configurou no Confluence.

(Somente Servidor/Data Center do Confluence) Se estiver usando a autenticação do Personal Access Token, insira o nome secreto e o token do Confluence que você configurou no seu Confluence.

iii. Salve e adicione seu segredo.

h. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

i. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

j. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

k. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Em Escopo de sincronização, para Sincronizar conteúdo — Escolha sincronizar entre os seguintes tipos de conteúdo: páginas, comentários de páginas, anexos de páginas, blogs, comentários de blog, anexos de blog, espaços pessoais e espaços arquivados.

> **Note**
>
> Comentários e anexos de página só podem ser selecionados se você optar por sincronizar páginas. Comentários e anexos do blog só podem ser selecionados se você optar por sincronizar blogs.

> **Important**
>
> Se você não especificar um padrão de regex de chave de espaço na Configuração adicional, todas as páginas e blogs serão rastreados por padrão.

b. Em Configuração adicional, em Tamanho máximo do arquivo — Especifique o limite de tamanho do arquivo em MBs que Amazon Kendra será rastreado. Amazon Kendra rastreará somente os arquivos dentro do limite de tamanho que você definir. O tamanho padrão do arquivo é 50 MB. O tamanho máximo do arquivo deve ser maior que 0 MB e menor ou igual a 50 MB.

   Para padrões regex de espaços — especifique se deseja incluir ou excluir espaços específicos em seu índice usando:

   - Tecla de espaço (por exemplo, *my-space-123*)

     > **Note**
     >
     > Se você não especificar um padrão de regex de chave de espaço, todas as páginas e blogs serão rastreados por padrão.

   - URL (por exemplo,*. *//MySiteMyDocuments/)*
   - Tipo de arquivo (por exemplo, *.*\ .pdf, .*\ .txt)*

   Para padrões de regex de títulos de entidades — especifique padrões de expressão regular para incluir ou excluir determinados blogs, páginas, comentários e anexos por títulos.

> **Note**
>
> Se quiser incluir ou excluir o rastreamento de uma página ou subpágina específica, você pode usar padrões de regex do título da página.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos da fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice. Para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

b. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Confluence

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `CONFLUENCEV2` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- URL do host — especifique a instância do URL do host do Confluence. Por exemplo, *`https://example.confluence.com`*.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Tipo de autenticação — especifique o tipo de autenticação, se, `Basic0Auth2`, (somente no Confluence Server). `Personal-token`
- (Opcional: somente para o Confluence Server) Local do certificado SSL: especifique o `S3bucketName` e `s3certificateName` usados para armazenar o certificado SSL.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contém as credenciais de autenticação que você configurou no Confluence. Se você usar uma autenticação básica, a senha deverá conter uma estrutura JSON com as seguintes chaves:

```
{
    "username": "email ID or user name",
    "password": "Confluence API token"
```

```
}
```

Para usar a autenticação OAuth 2.0, a senha é armazenada em uma estrutura JSON com as seguintes chaves:

```
{
    "confluenceAppKey": "app key",
    "confluenceAppSecret": "app secret",
    "confluenceAccessToken": "access token",
    "confluenceRefreshToken": "refresh token"
}
```

Somente para Servidor Confluence) Se você usar uma autenticação básica, a senha deverá ser armazenada em uma estrutura JSON com as seguintes chaves:

```
{
    "hostUrl": "Confluence Server host URL",
    "username": "Confluence Server user name",
    "password": "Confluence Server password"
}
```

Somente para Servidor Confluence) Se você usar a autenticação de token de acesso pessoal, a senha será armazenada em uma estrutura JSON com as seguintes chaves:

```
{
    "hostUrl": "Confluence Server host URL",
    "patToken": "personal access token"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do Confluence e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Confluence.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.

- Tamanho do arquivo — Especifique o tamanho máximo do arquivo a ser rastreado.
- Tipos de documento/conteúdo — especifique se deseja rastrear páginas, comentários de páginas, anexos de páginas, blogs, comentários do blog, anexos do blog, espaços e espaços arquivados.
- Filtros de inclusão e exclusão — especifique se deseja incluir ou excluir determinados espaços, páginas, blogs e seus comentários e anexos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Proxy da Web — Especifique as informações do proxy da Web se quiser se conectar à sua instância de URL do Confluence por meio de um proxy da Web. Você pode usar essa opção para o Confluence Server.
- Lista de controle de acesso (ACL) — Especifique se deseja rastrear as informações da ACL para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Confluence para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

  > **Note**
  >
  > O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice_`document_body`. Todos os demais campos são opcionais.

  Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema do modelo do Confluence.

Observações

- O token de acesso pessoal (PAT) não está disponível para o Confluence Cloud.

Conectores de fontes de dados personalizados

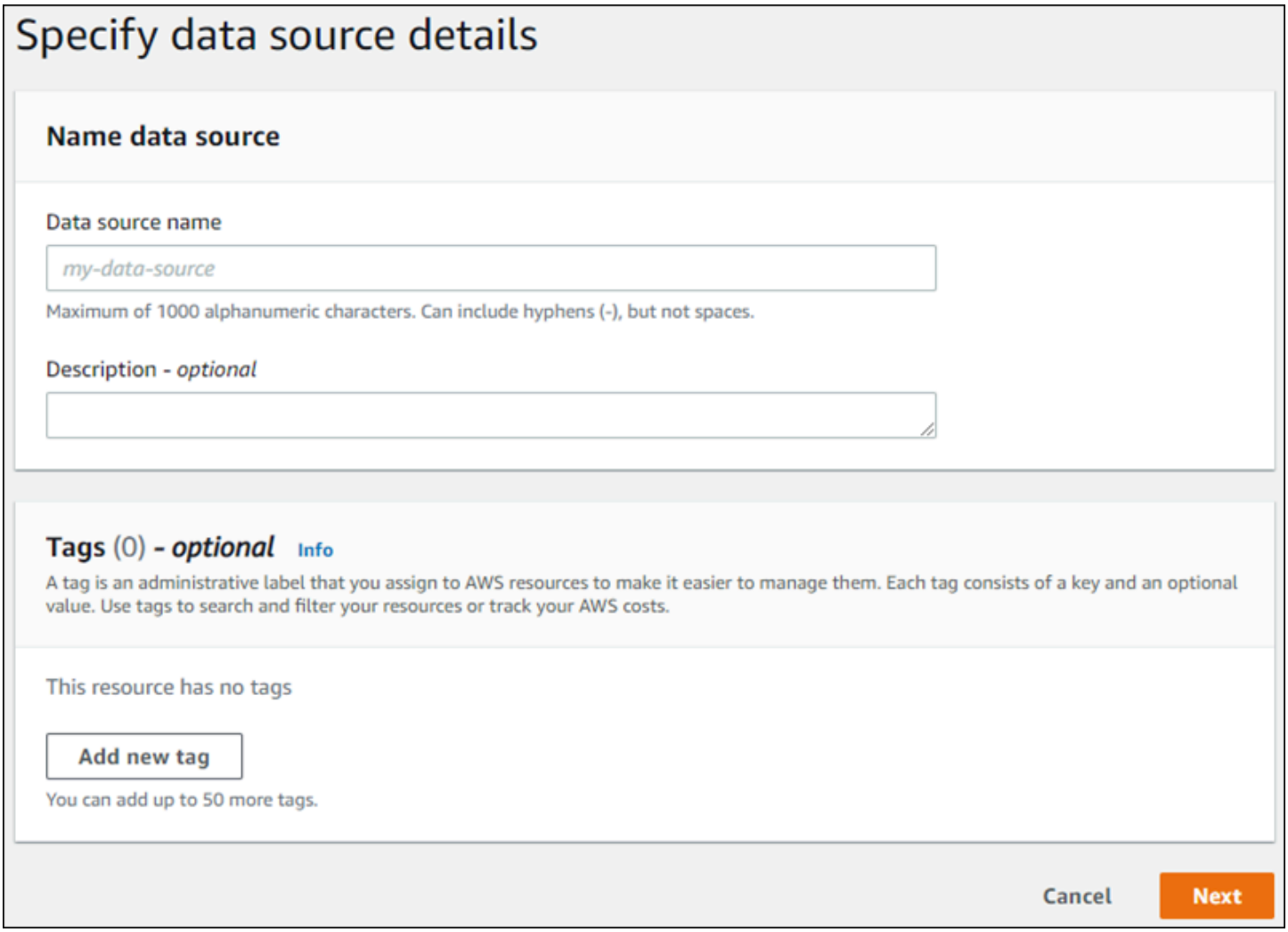
Use uma fonte de dados personalizada quando você tiver um repositório que ainda Amazon Kendra não fornece um conector de fonte de dados para. Você pode usá-lo para ver as mesmas métricas de histórico de execução que as fontes de Amazon Kendra dados fornecem, mesmo quando você não pode usar as fontes Amazon Kendra de dados para sincronizar seus repositórios. Use isso para criar uma experiência consistente de monitoramento de sincronização entre fontes de Amazon Kendra dados e fontes personalizadas. Especificamente, use uma fonte de dados personalizada para ver as métricas de sincronização de um conector de fonte de dados que você criou usando as APIs BatchPutBatchDeleteDocument e Document.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra, consulte Solucionar problemas de origens de dados.

Ao criar uma fonte de dados personalizada, você tem controle total sobre como os documentos a serem indexados são selecionados. Amazon Kendra fornece somente informações métricas que você pode usar para monitorar seus trabalhos de sincronização de fontes de dados. Você deve criar e executar o crawler que determina os documentos indexados pela fonte de dados.

Você deve especificar o título principal de seus documentos usando o objeto Documento `DocumentTitle` e para `DocumentURI` incluí-lo na resposta do `Query` resultado. `_source_uri` DocumentAttribute

Você cria um identificador para sua fonte de dados personalizada usando o console ou usando a API CreateDataSource. Para usar o console, dê um nome à sua fonte de dados e, opcionalmente, uma descrição e tags de recursos. Depois que a fonte de dados é criada, um ID da fonte de dados é exibida. Copie esse ID para usar ao sincronizar a fonte de dados com o índice.



Também é possível criar uma fonte de dados personalizada usando a API `CreateDataSource`. A API retorna um ID para ser usado quando ao sincronizar a fonte de dados. Ao usar a API `CreateDataSource` para criar uma fonte de dados personalizada, não é possível definir os parâmetros `Configuration`, `RoleArn` ou `Schedule`. Se você definir esses parâmetros, Amazon Kendra retornará uma `ValidationException` exceção.

Para usar uma fonte de dados personalizada, crie um aplicativo responsável pela atualização do índice do Amazon Kendra . O aplicativo depende de um crawler criado por você. O crawler lê os documentos em seu repositório e determina quais devem ser enviados para Amazon Kendra. O aplicativo deve executar as seguintes etapas:

1. Rastreie o repositório e faça uma lista dos documentos no repositório que foram adicionados, atualizados ou excluídos.
2. Chame a StartDataSourceSyncJob API para sinalizar que um trabalho de sincronização está começando. Você fornece uma ID da fonte de dados para identificar a fonte de dados que está sincronizando. Amazon Kendra retorna um ID de execução para identificar um trabalho de sincronização específico.
3. Chame a API de BatchDeletedocumentos para remover documentos do índice. Forneça o ID da fonte de dados e o ID de execução para identificar a fonte de dados que está sendo sincronizada e o trabalho ao qual essa atualização está associada.
4. Chame a StopDataSourceSyncJob API para sinalizar o fim do trabalho de sincronização. Depois de chamar a API `StopDataSourceSyncJob`, o ID de execução associado não é mais válido.
5. Chame a API ListDataSourceSyncJobs com os identificadores de índice e fonte de dados para listar os trabalhos de sincronização da fonte de dados e ver as métricas dos trabalhos de sincronização.

Depois de finalizar um trabalho de sincronização, você pode iniciar um novo trabalho de sincronização. Pode haver um período de tempo até que todos os documentos enviados sejam adicionados ao índice. Use a API `ListDataSourceSyncJobs` para ver o status da tarefa de sincronização. Se o `Status` retornado para o trabalho de sincronização for `SYNCING_INDEXING`, alguns documentos ainda estão sendo indexados. Você pode iniciar um novo trabalho de sincronização quando o status do trabalho anterior for `FAILED` ou`SUCCEEDED`.

Depois de chamar a API `StopDataSourceSyncJob`, você não pode usar um identificador de trabalho de sincronização em uma chamada para as APIs `BatchPutDocument` ou `BatchDeleteDocument`. Se você fizer isso, todos os documentos enviados serão retornados na mensagem de resposta `FailedDocuments` da API.

Atributos obrigatórios

Quando você envia um documento para Amazon Kendra usar a `BatchPutDocument` API, cada documento exige dois atributos para identificar a fonte de dados e a execução de sincronização à

qual ele pertence. Você deve fornecer os dois seguintes atributos para mapear documentos da sua fonte de dados personalizada corretamente para um índice do Amazon Kendra :

- `_data_source_id`: o identificador da fonte de dados. Ele é retornado quando você cria a fonte de dados com o console ou a API `CreateDataSource`.
- `_data_source_sync_job_execution_id`: o identificador da execução da sincronização. Ele é retornado quando você inicia a sincronização do índice com a API `StartDataSourceSyncJob`.

Veja a seguir o JSON necessário para indexar um documento usando uma fonte de dados personalizada.

```
{
    "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"
}
```

Ao remover um documento do índice usando a API `BatchDeleteDocument`, especifique os dois campos a seguir no parâmetro `DataSourceSyncJobMetricTarget`:

- `DataSourceId`: o identificador da fonte de dados. Ele é retornado quando você cria a fonte de dados com o console ou a API `CreateDataSource`.
- `DataSourceSyncJobId`: o identificador da execução da sincronização. Ele é retornado quando você inicia a sincronização do índice com a API `StartDataSourceSyncJob`.

Veja a seguir o JSON necessário para excluir um documento do índice usando API `BatchDeleteDocument`.

```
{
    "DataSourceSyncJobMetricTarget": {
        "DataSourceId": "data source identifier",
        "DataSourceSyncJobId": "sync job identifier"
    },
    "DocumentIdList": [
        "document identifier"
    ],
    "IndexId": "index identifier"
}
```

Visualizar métricas

Depois que um trabalho de sincronização for concluído, você poderá usar a API de DataSourceSyncJobmétricas para obter as métricas associadas ao trabalho de sincronização. Use ela para monitorar as sincronizações de fontes de dados personalizadas.

Se você enviar o mesmo documento várias vezes, seja como parte da API `BatchPutDocument`, da `BatchDeleteDocument` API ou se o documento for enviado para adição e exclusão, o documento será contado apenas uma vez nas métricas.

- `DocumentsAdded`: o número de documentos enviados usando a API `BatchPutDocument` associada a esse trabalho de sincronização adicionados ao índice pela primeira vez. Se um documento for enviado para adição mais de uma vez em uma sincronização, ele será contabilizado apenas uma vez nas métricas.
- `DocumentsDeleted`: o número de documentos enviados usando a API `BatchDeleteDocument` associada a esse trabalho de sincronização excluídos pelo índice.. Se um documento for enviado para exclusão mais de uma vez em uma sincronização, ele será contabilizado apenas uma vez nas métricas.

- `DocumentsFailed`: o número de documentos associados a essa tarefa de sincronização que falharam na indexação. Esses são documentos que foram aceitos por Amazon Kendra para indexação, mas que não puderam ser indexados ou excluídos. Se um documento não for aceito por Amazon Kendra, o identificador do documento será retornado na propriedade de `FailedDocuments` resposta das `BatchDeleteDocument` APIs `BatchPutDocument` e.
- `DocumentsModified`—O número de documentos modificados enviados usando a `BatchPutDocument` API associada a esse trabalho de sincronização que foram modificados no Amazon Kendra índice.

Amazon Kendra também emite Amazon CloudWatch métricas ao indexar documentos. Para obter mais informações, consulte Monitoramento Amazon Kendra com Amazon CloudWatch.

Amazon Kendra não retorna a `DocumentsScanned` métrica para fontes de dados personalizadas. Ele também emite as CloudWatch métricas listadas no documento Métricas para fontes Amazon Kendra de dados.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados personalizada, consulte:

- Adicionar fontes de dados personalizadas ao Amazon Kendra

Fontes de dados personalizadas (Java)

O código a seguir fornece um exemplo de implementação de uma fonte de dados personalizada usando Java. Primeiro, o programa cria uma fonte de dados personalizada e depois sincroniza os documentos recém-adicionados ao índice com a fonte de dados personalizada.

O código a seguir demonstra a criação e o uso de uma fonte de dados personalizada. Ao usar uma fonte de dados personalizada no aplicativo, não é necessário criar uma nova fonte de dados (processo único) toda vez que o índice é sincronizado com a fonte de dados. Você usa o ID do índice e o ID da fonte de dados para sincronizar os dados.

```
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

O Dropbox é um serviço de hospedagem de arquivos que oferece armazenamento em nuvem, organização de documentos e serviços de modelagem de documentos. Se você for usuário do Dropbox, você pode usar Amazon Kendra para indexar seus arquivos do Dropbox, Dropbox Paper, modelos do Dropbox Paper e atalhos armazenados para páginas da web. Você também pode configurar Amazon Kendra para indexar arquivos específicos do Dropbox, Dropbox Paper, modelos do Dropbox Paper e atalhos armazenados em páginas da web.

Amazon Kendra é compatível com o Dropbox e o Dropbox Advanced para o Dropbox Business.

Você pode se conectar Amazon Kendra à sua fonte de dados do Dropbox usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados Amazon Kendra do Dropbox, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra O conector de fonte de dados do Dropbox é compatível com os seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais

- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Dropbox, faça essas alterações em seu Dropbox e AWS em suas contas.

No Dropbox, verifique se você:

- Criou uma conta do Dropbox Advanced e configurou um usuário administrador.
- Configurei um aplicativo do Dropbox com um nome de aplicativo exclusivo, ativou o Scoped Access. Consulte a Documentação do Dropbox sobre a criação de um aplicativo.
- Ativou as permissões completas do Dropbox no console do Dropbox e adicionou as seguintes permissões:
  - files.content.read
  - files.metadata.read
  - sharing.read
  - file_requests.read
  - groups.read
  - team_info.read
  - team_data.content.read
- Anote a chave do aplicativo do Dropbox, a senha do aplicativo do Dropbox e o token de acesso do Dropbox para credenciais básicas de autenticação.

  > **Note**
  >
  > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Configurei e copiei um token de acesso temporário do Oauth 2.0 para seu aplicativo do Dropbox. Esse token é temporário e expira após 4 horas. Consulte a Documentação do Dropbox sobre autenticação OAuth.

> **Note**
>
> É recomendável criar um token de acesso de atualização que nunca expire no Dropbox, em vez de confiar em um token de acesso único que expira após quatro horas. Um token de acesso de atualização é permanente e nunca expira, para que você possa continuar sincronizando a fonte de dados no futuro.

- Recomendado: configurou um token de atualização permanente do Dropbox que nunca expira Amazon Kendra para permitir que você continue sincronizando sua fonte de dados sem interrupções. Consulte a Documentação do Dropbox sobre tokens de atualização.
- Verificou se cada documento é exclusivo no banco de dados e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Dropbox em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Dropbox a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Dropbox, você deve fornecer os detalhes necessários da sua fonte de dados do Dropbox para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Dropbox para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao Dropbox

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector do Dropbox e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector do Dropbox com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   b. Tipo de token de autenticação — Escolha um token permanente (recomendado) ou um token de acesso temporário.

   c. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Dropbox. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

      i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

         A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Dropbox' é adicionado automaticamente ao seu nome secreto.

         B. Para informações de chave do aplicativo, segredo do aplicativo e token (permanentes ou temporárias), insira os valores da credencial de autenticação configurados no Dropbox.

      ii. Salve e adicione seu segredo.

   d. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

   e. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade

estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

f. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

g. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Em Selecionar entidades ou tipos de conteúdo — Escolha entidades do Dropbox ou tipos de conteúdo que você deseja rastrear.

b. Em Configurações adicionais para Padrões Regex: adicione os padrões de expressão regular para incluir ou excluir determinados arquivos.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Arquivos, Dropbox Paper e modelos do Dropbox Paper — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Dropbox

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `DROPBOX` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo

da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Tipo de token de acesso — especifique se você deseja usar um token de acesso permanente ou temporário para seu AWS Secrets Manager segredo, que armazena suas credenciais de autenticação.

> **Note**
>
> É recomendável criar um token de acesso de atualização que nunca expire no Dropbox, em vez de confiar em um token de acesso único que expira após quatro horas. Crie um aplicativo e um token de acesso de atualização no console do desenvolvedor do Dropbox e forneça o token de acesso na senha.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Dropbox. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "appKey": "Dropbox app key",
    "appSecret": "Dropbox app secret",
    "accesstoken": "temporary access token or refresh access token"
}
```

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

- IAM função — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e para chamar as APIs públicas necessárias para o conector do Dropbox e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para fontes de dados do Dropbox.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Tipos de documento/conteúdo — especifique se deseja rastrear arquivos em seu Dropbox, documentos do Dropbox Paper, modelos do Dropbox Paper e atalhos de páginas da web armazenados em seu Dropbox.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinadas arquivos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Lista de controle de acesso (ACL) — Especifique se deseja rastrear as informações da ACL para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Dropbox para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você

deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice_`document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes para configurar, consulte Esquema de modelos do Dropbox.

Saiba mais

Para saber mais sobre a integração do Amazon Kendra com a fonte de dados do Box, consulte:

- Indexe o conteúdo do Dropbox usando o conector do Dropbox para o Amazon Kendra

Drupal

O Drupal é um sistema de gerenciamento de conteúdo de código aberto (CMS) que pode ser usado para criar sites e aplicativos da Web. Você pode usar Amazon Kendra para indexar o seguinte no Drupal:

- Conteúdo: artigos, páginas básicas, blocos básicos, tipos de conteúdo definidos pelo usuário, tipos de blocos definidos pelo usuário, tipos de conteúdo personalizados e tipos de blocos personalizados
- Comentário: para qualquer tipo de conteúdo e tipo de bloco
- Anexos: para qualquer tipo de conteúdo e tipo de bloco

Você pode se conectar Amazon Kendra à sua fonte de dados do Drupal usando o Amazon Kendra console ou a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Drupal, consulte. Solucionar problemas de origens de dados

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

Amazon Kendra O conector de fonte de dados Drupal oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Drupal, faça essas alterações no Drupal e nas contas. AWS

No Drupal, verifique se você:

- Criou uma conta do Drupal (Standard) Suite e um usuário com função de administrador.
- Copiou o nome do site do Drupal e configurou um URL de host. Por exemplo, *https:///*<hostname><drupalsitename>.
- Credenciais básicas de autenticação configuradas contendo um nome de usuário (nome de usuário de login do site do Drupal) e senha (senha do site do Drupal).
- Recomendado: configurou um token de credencial do OAuth 2.0. Use esse token junto com a concessão de senha do Drupal, o ID do cliente, a senha do cliente, o nome de usuário (nome de usuário de login do site do Drupal) e senha (senha do site do Drupal) para se conectar ao Amazon Kendra.
- As seguintes permissões foram adicionadas à sua conta do Drupal usando uma função de administrador:
  - administrar blocos
  - administrar blocos_exibição de conteúdo
  - administrar blocos_campos de conteúdo
  - administrar blocos_exibição de formulário de conteúdo
  - administrar visualizações
  - visualizar endereços de e-mail do usuário
  - ver conteúdo próprio não publicado

- ver revisões da página
- ver revisões do artigo
- ver todas as revisões
- ver o tema de administração
- acessar conteúdo
- visão geral do conteúdo do acesso
- acessar comentários
- pesquisar conteúdo
- visão geral dos arquivos de acesso
- acessar links contextuais

> **Note**
>
> Se houver tipos de conteúdo definidos pelo usuário, tipos de blocos definidos pelo usuário ou se quaisquer visualizações e blocos forem adicionados ao site do Drupal, eles deverão receber acesso de administrador.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Drupal em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não

recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Drupal a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Drupal, você deve fornecer detalhes de suas credenciais do Drupal para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Drupal para Amazon Kendra ver. Pré-requisitos

Console

Para se conectar Amazon Kendra ao Drupal

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

   > **Note**
   >
   > Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector Drupal e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Drupal com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. Em Fonte, para URL do host: o URL do host do site do Drupal. Por exemplo, *https:///*<hostname><drupalsitename>.

b. Para o Local do certificado SSL, insira o caminho para o certificado SSL armazenado em um bucket do Amazon S3 .

c. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

d. Para a Autenticação, escolha entre Autenticação básica e Autenticação Oauth 2.0 com base no seu caso de uso.

e. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Drupal. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

A. Se você escolher a Autenticação básica, digite uma senha, o Nome de usuário (nome de usuário do site do Drupal) e a Senha (senha do site do Drupal) que você copiou e escolha Salvar e adicionar senha.

B. Se você escolheu a Autenticação OAuth 2.0, insira uma senha, Nome de usuário (nome de usuário do site do Drupal), Senha (senha do site do Drupal), ID do cliente e Senha do cliente gerados na conta do Drupal e escolha Salvar e adicionar senha.

ii. Escolha Salvar.

f. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

g. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Em Sincronizar escopo, escolha uma das opções a seguir:

> **Note**
>
> Ao escolher rastrear Artigos, Páginas básicas e Blocos básicos, os campos padrão serão sincronizados automaticamente. Você também pode optar por sincronizar comentários, anexos, campos personalizados e outras entidades personalizadas.

- Para Selecionar entidades:

- Artigos: escolha se deseja rastrear Artigos, seus comentários em Comentários e Anexos.
- Páginas básicas: escolha se deseja rastrear as Páginas básicas, seus comentários em Comentários e Anexos.
- Blocos básicos: escolha se deseja rastrear os Blocos básicos, seus comentários em Comentários e Anexos.
- Também é possível adicionar Tipos de conteúdo personalizados e Blocos personalizados.

b. Para Opções de configuração opcionais:

- Para o padrão Regex, adicione padrões de expressão regular para incluir ou excluir títulos de entidades e nomes de arquivos específicos. Você pode adicionar até 100 padrões.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Para Conteúdo, Comentários e Anexos — Selecione entre os campos da fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Drupal

Você deve especificar um JSON do esquema da fonte de dados usando a API TemplateConfiguration. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `DRUPAL` quando você usa o esquema TemplateConfigurationJSON. Além disso, especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou na sua conta do Drupal.

  Se você usar uma autenticação básica, a senha deverá conter uma estrutura JSON com as seguintes chaves:

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  Para usar a autenticação OAuth 2.0, a senha é armazenada em uma estrutura JSON com as seguintes chaves:

  ```
  {
      "username": "user name",
      "password": "password",
      "clientId": "client id",
      "clientSecret": "client secret"
  }
  ```

  > **Note**
  >
  > > **Note**
  > >
  > > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Drupal e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para fontes de dados do Drupal.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: você pode especificar se deseja incluir conteúdo, comentários e anexos. Você também pode especificar padrões de expressão regular para incluir ou excluir conteúdos, comentários e anexos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Drupal para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelos do Drupal.

Observações

- As APIs do Drupal não têm limites oficiais de controle de utilização.
- Os SDKs Java não estão disponíveis para o Drupal.
- Os dados do Drupal só podem ser obtidos usando as APIs JSON nativas.
- Os tipos de conteúdo não associados a nenhuma visualização do Drupal não podem ser rastreados.
- Você precisa de acesso de administrador para rastrear dados do dos Blocos do Drupal.
- Não há API JSON disponível para criar o tipo de conteúdo definido pelo usuário usando verbos HTTP.
- O corpo do documento e os comentários para Artigos, Páginas básicas, Blocos básicos, tipo de conteúdo definido pelo usuário e tipo de bloco definido pelo usuário são exibidos no formato HTML. Se o conteúdo HTML não estiver bem formado, as tags relacionadas ao HTML aparecerão no corpo do documento e nos comentários e ficarão visíveis nos resultados da pesquisa do Amazon Kendra .
- Os tipos de conteúdo e os tipos de bloco sem descrição ou corpo não serão Amazon Kendra incorporados. Somente comentários e anexos desses tipos de conteúdo ou bloco serão inseridos em seu índice. Amazon Kendra

GitHub

GitHub é um serviço de hospedagem baseado na web para desenvolvimento de software que fornece serviços de armazenamento e gerenciamento de código com controle de versão. Você pode usar Amazon Kendra para indexar seus arquivos de repositório GitHub Enterprise Cloud (SaaS) e GitHub Enterprise Server (On Prem), solicitações de problemas e pull, comentários de problemas e pull requests e anexos de comentários de problemas e pull requests. Você também pode optar por incluir ou excluir determinados arquivos.

> **Note**
>
> Amazon Kendra agora suporta um GitHub conector atualizado.

O console foi atualizado automaticamente para você. Todos os novos conectores que você criar no console usarão a arquitetura atualizada. Se você usa a API, agora deve usar o TemplateConfigurationobjeto em vez do `GitHubConfiguration` objeto para configurar seu conector.
Os conectores configurados usando o console antigo e a arquitetura de API continuarão funcionando conforme configurados. No entanto, você não poderá editá-los ou atualizá-los. Se você quiser editar ou atualizar a configuração do conector, deverá criar um novo conector. Recomendamos migrar o fluxo de trabalho do conector para a versão atualizada. O suporte para conectores configurados usando a arquitetura mais antiga está programado para terminar em junho de 2024.

Você pode se conectar Amazon Kendra à sua fonte de GitHub dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra GitHub dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra GitHub o conector de fonte de dados oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de GitHub dados, faça essas alterações em suas GitHub AWS contas.

Em GitHub, verifique se você tem:

- Criou um GitHub usuário com permissões administrativas para a GitHub organização.
- Configurei um token de acesso pessoal no Git Hub para usar como suas credenciais de autenticação. Consulte a GitHub documentação sobre como criar um token de acesso pessoal.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Recomendado: configurou um token OAuth para credenciais de autenticação. Use o token OAuth para melhorar os limites de controle de utilização da API e o desempenho do conector. Consulte a GitHub documentação sobre autorização do OAuth.
- Anote o URL do GitHub host para o tipo de GitHub serviço que você usa. Por exemplo, a URL do host da GitHub nuvem pode ser *`https://api.github.com`* e a URL do host do GitHub servidor pode ser *`https://on-prem-host-url/api/v3/`*.
- Anote o nome da sua organização para GitHub a conta do GitHub Enterprise Cloud (SaaS) ou da conta do GitHub Enterprise Server (local) à qual você deseja se conectar. Você pode encontrar o nome da sua organização fazendo login no GitHub desktop e selecionando Suas organizações na lista suspensa da foto do perfil.
- Opcional (somente servidor): gerou um certificado SSL e copiou o caminho para o certificado armazenado em um Amazon S3 bucket. Você usa isso para GitHub se conectar se precisar de uma conexão SSL segura. Você pode simplesmente gerar um certificado autoassinado X509 em qualquer computador usando o OpenSSL. Para ver um exemplo de uso do OpenSSL para criar um certificado X509, consulte Criar e assinar um certificado X509.
- Adicionou as seguintes permissões:

  Para nuvem GitHub corporativa (SaaS)

- `repo:status`— Concede acesso de leitura/gravação aos status de confirmação em repositórios públicos e privados. Esse escopo só é necessário para conceder a outros usuários ou serviços acesso aos status de confirmação do repositório privado sem conceder acesso ao código.
- `repo_deployment`— Concede acesso aos status de implantação de repositórios públicos e privados. Esse escopo só é necessário para conceder a outros usuários ou serviços acesso aos status de implantação, sem conceder acesso ao código.
- `public_repo`— Limita o acesso aos repositórios públicos. Isso inclui acesso de leitura/gravação ao código, status de confirmação, projetos de repositório, colaboradores e status de implantação para repositórios e organizações públicas. Também é necessário para marcar repositórios públicos como favoritos.
- `repo:invite`— Concede habilidades de aceitação/recusa de convites para colaborar em um repositório. Esse escopo só é necessário para conceder a outros usuários ou serviços acesso aos convites sem conceder acesso ao código.
- `security_events`— Concede: acesso de leitura e gravação a eventos de segurança na API de digitalização de código. Esse escopo só é necessário para conceder a outros usuários ou serviços acesso a eventos de segurança sem conceder acesso ao código.
- `read:org`— Acesso somente para leitura à associação à organização, aos projetos da organização e à associação à equipe.
- `user:email`— Concede acesso de leitura aos endereços de e-mail de um usuário. Exigido pela Amazon Kendra para rastrear ACLs.
- `user:follow`— Concede acesso para seguir ou deixar de seguir outros usuários. Exigido pela Amazon Kendra para rastrear ACLs.
- `read:user`— Concede acesso para ler os dados do perfil de um usuário. Exigido pela Amazon Kendra para rastrear ACLs.
- `workflow`— Concede a capacidade de adicionar e atualizar arquivos de fluxo de trabalho do GitHub Actions. Arquivos de fluxo de trabalho podem ser confirmados sem esse escopo se o mesmo arquivo (com o mesmo caminho e conteúdo) existir em outra ramificação no mesmo repositório.

Para obter mais informações, consulte Escopos para aplicativos OAuth no Docs. GitHub

Para servidor GitHub corporativo (no local)

- `repo:status`— Concede acesso de leitura/gravação aos status de confirmação em repositórios públicos e privados. Esse escopo só é necessário para conceder a outros usuários

ou serviços acesso aos status de confirmação do repositório privado sem conceder acesso ao código.

- `repo_deployment`— Concede acesso aos status de implantação de repositórios públicos e privados. Esse escopo só é necessário para conceder a outros usuários ou serviços acesso aos status de implantação, sem conceder acesso ao código.
- `public_repo`— Limita o acesso aos repositórios públicos. Isso inclui acesso de leitura/gravação ao código, status de confirmação, projetos de repositório, colaboradores e status de implantação para repositórios e organizações públicas. Também é necessário para marcar repositórios públicos como favoritos.
- `repo:invite`— Concede habilidades de aceitação/recusa de convites para colaborar em um repositório. Esse escopo só é necessário para conceder a outros usuários ou serviços acesso aos convites sem conceder acesso ao código.
- `security_events`— Concede: acesso de leitura e gravação a eventos de segurança na API de digitalização de código. Esse escopo só é necessário para conceder a outros usuários ou serviços acesso a eventos de segurança sem conceder acesso ao código.
- `read:user`— Concede acesso para ler os dados do perfil de um usuário. Exigido pelo Amazon Q Business para rastrear ACLs.
- `user:email`— Concede acesso de leitura aos endereços de e-mail de um usuário. Exigido pelo Amazon Q Business para rastrear ACLs.
- `user:follow`— Concede acesso para seguir ou deixar de seguir outros usuários. Exigido pelo Amazon Q Business para rastrear ACLs.
- `site_admin`— Concede aos administradores do site acesso aos endpoints da API GitHub Enterprise Server Administration.
- `workflow`— Concede a capacidade de adicionar e atualizar arquivos de fluxo de trabalho do GitHub Actions. Arquivos de fluxo de trabalho podem ser confirmados sem esse escopo se o mesmo arquivo (com o mesmo caminho e conteúdo) existir em outra ramificação no mesmo repositório.

Para obter mais informações, consulte Escopos para aplicativos OAuth no GitHub Documentos e Entendendo os escopos para aplicativos OAuth no Developer. GitHub

- Verifique se cada documento é exclusivo em GitHub e entre outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de GitHub autenticação em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de GitHub dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de GitHub dados, você deve fornecer os detalhes necessários da sua fonte de GitHub dados para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou GitHub para Amazon Kendra, consulte Pré-requisitos.

Console

Para se conectar Amazon Kendra a GitHub

1. Faça login no AWS Management Console e abra o Amazon Kendra console.

2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha GitHub conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o GitHub conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. GitHubfonte — Escolha entre GitHub Enterprise Cloud e GitHubEnterprise Server.

   b. GitHub URL do host — Por exemplo, o URL do host para a GitHub nuvem pode ser *`https://api.github.com`* e o URL do host do GitHub servidor pode ser *`https://on-prem-host-url/api/v3/`*.

   c. GitHub nome da organização — Insira o nome GitHub da sua organização. Você pode encontrar as informações da sua organização em sua GitHub conta.

   d. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no

acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

e. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de GitHub autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

   i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

      A. Senha: um nome para sua senha. O prefixo 'AmazonKendra- GitHub -' é adicionado automaticamente ao seu nome secreto.

      B. Para GitHubtoken — Insira o valor da credencial de autenticação configurado em. GitHub

   ii. Salve e adicione seu segredo.

f. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

g. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Selecionar repositórios — Escolha rastrear todos os repositórios ou selecionar.

      Se você optar por rastrear repositórios selecionados, adicione os nomes dos repositórios e, opcionalmente, o nome de qualquer ramificação específica.

   b. Tipos de conteúdo — escolha os tipos de conteúdo que você deseja rastrear a partir de arquivos, problemas, pull requests e muito mais.

   c. Padrões Regex: adicionar padrões de expressão regular para incluir ou excluir determinados arquivos.

   d. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

      - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
      - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
      - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

   e. Em Sincronização, cronograma de execução para frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

   f. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a GitHub

Você deve especificar um JSON do esquema da fonte de dados usando a API TemplateConfiguration. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `GITHUB` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- GitHubtipo — Especifique o tipo como `SAAS` ou`ON_PREMISE`.
- URL do host — especifique o URL do GitHub host ou o URL do endpoint da API. Por exemplo, se você usa GitHub SaaS/Enterprise Cloud, o URL do host pode ser e`https://api.github.com`, para o servidor GitHub local/corporativo, o URL do host pode ser. `https://on-prem-host-url/api/v3/`
- Nome da organização — Especifique o nome da organização da GitHub conta. Você pode encontrar o nome da sua organização fazendo login no GitHub desktop e selecionando Suas organizações na lista suspensa da foto do perfil.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo

da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta. GitHub A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "personalToken": "token"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o GitHub conector e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes GitHub de dados.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.

> **Note**
>
> Se você usa GitHub servidor, você deve usar um Amazon VPC para se conectar ao seu GitHub servidor.

- Filtro de repositório — filtra repositórios por seus nomes e nomes de ramificações.
- Tipos de documento/conteúdo — especifique se deseja rastrear documentos do repositório, problemas, comentários de problemas, anexos de comentários de problemas, pull requests, comentários de pull request, anexos de comentários de pull request.
- Filtros de inclusão e exclusão — especifique se deseja incluir ou excluir determinados arquivos e pastas.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Lista de controle de acesso (ACL) — Especifique se deseja rastrear as informações da ACL para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo — Escolha mapear os campos da fonte de GitHub dados para os Amazon Kendra campos de índice. Você pode incluir campos de documentos, confirmações, problemas, anexos de problemas, comentários de problemas, pull requests, anexos de pull request, comentários de pull request. Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é necessário para que a Amazon Kendra pesquise seus documentos.

Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice `_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o Esquema do modelo do GitHub.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte GitHub de dados, consulte:

- Reinvente a pesquisa em GitHub repositórios com a potência do conector Amazon Kendra GitHub

Gmail

O Gmail é um cliente de e-mail desenvolvido pelo Google em que você pode enviar mensagens de e-mail com anexos de arquivos. As mensagens do Gmail podem ser classificadas e armazenadas na caixa de entrada do e-mail usando pastas e marcadores. Você pode usar Amazon Kendra para indexar suas mensagens de e-mail e anexos de mensagens. Você também pode configurar Amazon Kendra para incluir ou excluir mensagens de e-mail, anexos de mensagens e rótulos específicos para indexação.

Você pode se conectar Amazon Kendra à sua fonte de dados do Gmail usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Gmail, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Gmail, faça essas alterações no Gmail e AWS nas contas.

No Gmail, verifique se você:

- Criou uma conta de administrador do Google Cloud Platform e um projeto do Google Cloud.
- Ativou a API do Gmail e a API do SDK Admin na conta de administrador.
- Criou uma conta de serviço e baixou uma chave privada JSON para o Gmail. Para obter informações sobre como criar e acessar a chave privada, consulte a documentação do Google Cloud sobre como criar uma chave de conta de serviço e as credenciais da conta de serviço.
- Copiou o e-mail da sua conta de administrador, o e-mail da sua conta de serviço e sua chave privada para usar como suas credenciais de autenticação.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Foram adicionados os seguintes escopos do Oauth (usando uma função de administrador) para o usuário e os diretórios compartilhados que você deseja indexar:
  - https://www.googleapis.com/auth/admin.directory.user.readonly
  - https://www.googleapis.com/auth/gmail.readonly
- Verifique se cada documento é exclusivo no Gmail e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve

conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Gmail em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados do Gmail a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Gmail, você deve fornecer detalhes das suas credenciais do Gmail para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Gmail para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao Gmail

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector do Gmail e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector do Gmail com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   b. Em Autenticação por AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Gmail. Se você optar por criar um novo segredo, uma janela AWS Secrets Manager secreta será aberta.

- Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

  A. Senha: um nome para sua senha.

  B. E-mail do cliente: o e-mail do cliente copiado da conta de serviço do Google.

  C. E-mail da conta do administrador: o e-mail da conta do administrador que você gostaria de usar.

  D. Chave privada: a chave privada que você copiou da conta de serviço do Google.

  E. Salve e adicione seu segredo.

c. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

d. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

e. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Para tipos de entidade — Escolha sincronizar anexos de mensagens.

   b. (Opcional) Em Configurações adicionais, insira as seguintes informações:

      i. Intervalo de datas — insira um intervalo de datas para especificar a data de início e término dos e-mails que você deseja rastrear.

      ii. Domínios de e-mail — inclua ou exclua determinados e-mails com base nos domínios de e-mail “para”, “de”, “cc” e “bcc”.

      iii. Palavras-chave nos assuntos — inclua ou exclua e-mails com base nas palavras-chave nos assuntos dos e-mails.

> **Note**
>
> Você também pode optar por incluir qualquer documento que corresponda a todas as palavras-chave do assunto inseridas

iv. Marcadores — Adicione padrões de expressão regular para incluir ou excluir determinados rótulos de e-mail.

v. Anexos — Adicione padrões de expressão regular para incluir ou excluir determinados anexos de e-mail.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

> **Important**
>
> Como não há uma API para atualizar mensagens do Gmail excluídas permanentemente, sincronize conteúdo novo, modificado ou excluído:
>
> - Não removerá mensagens que foram excluídas permanentemente do Gmail do seu índice Amazon Kendra
> - Não sincronizará alterações nas etiquetas de e-mail do Gmail

Para sincronizar as alterações no rótulo da fonte de dados do Gmail e as mensagens de e-mail excluídas permanentemente com o índice do Amazon Kendra , execute rastreamentos completos periodicamente.

d. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

> **Note**
>
> Amazon Kendra O conector de fonte de dados do Gmail não é compatível com a criação de campos de índice personalizados devido às limitações da API.

b. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Gmail

Você deve especificar um JSON do esquema da fonte de dados usando a API TemplateConfiguration. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `GMAIL` quando você usa o esquema TemplateConfigurationJSON. Além disso, especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão.

Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:

- `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
- `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

> **Important**
>
> Como não há uma API para atualizar mensagens do Gmail excluídas permanentemente, sincronize conteúdo novo, modificado ou excluído:
>
> - Não removerá mensagens que foram excluídas permanentemente do Gmail do seu índice Amazon Kendra
> - Não sincronizará alterações nas etiquetas de e-mail do Gmail
>
> Para sincronizar as alterações no rótulo da fonte de dados do Gmail e as mensagens de e-mail excluídas permanentemente com seu Amazon Kendra índice, você deve executar rastreamentos completos periodicamente.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Gmail. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "adminAccountEmailId": "service account email",
    "clientEmailId": "user account email",
    "privateKey": "private key"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do Gmail e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para fontes de dados do Gmail.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão — especifique se deseja incluir ou excluir determinados e-mails “para”, “de”, “cc”, “bcc”.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Gmail para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

> **Note**
>
> Amazon Kendra O conector de fonte de dados do Gmail não é compatível com a criação de campos de índice personalizados devido às limitações da API.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o Esquema do modelo do Gmail.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Gmail, consulte:

- Faça pesquisas inteligentes em e-mails no seu espaço de trabalho do Google usando o conector do Gmail para o Amazon Kendra.

Observações

- Como não há uma API para atualizar mensagens do Gmail excluídas permanentemente, uma `FULL_CRAWL`/Sincronização de conteúdo novo, modificado ou excluído:
  - Não removerá mensagens que foram excluídas permanentemente do Gmail do seu índice Amazon Kendra
  - Não sincronize alterações nas etiquetas de e-mail do Gmail

  Para sincronizar as alterações no rótulo da fonte de dados do Gmail e as mensagens de e-mail excluídas permanentemente com seu Amazon Kendra índice, você deve executar rastreamentos completos periodicamente.
- Amazon Kendra O conector de fonte de dados do Gmail não é compatível com a criação de campos de índice personalizados devido às limitações da API.

Google Drive

O Google Drive é um serviço de armazenamento de arquivos baseado em nuvem. Você pode usar o Amazon Kendra para indexar documentos armazenados nos drives compartilhados, Meus Drives e Compartilhado comigo na fonte de dados do Google Drive. Você pode indexar os documentos do Google Workspace e os documentos listados em Tipos de documentação. Você também pode usar filtros de inclusão e exclusão para indexar o conteúdo por nome, tipo e caminho do arquivo.

Você pode se conectar Amazon Kendra à sua fonte de dados do Google Drive usando o Amazon Kendra console, a TemplateConfigurationAPI ou a GoogleDriveConfigurationAPI.

Amazon Kendra tem duas versões do conector do Google Drive. Os recursos suportados de cada versão incluem:

Conector do Google Drive V1.0/API GoogleDriveConfiguration

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão

Conector do Google Drive V2.0/API TemplateConfiguration

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

> **Note**
>
> O suporte para o conector do Google Drive V1.0 DriveConfiguration /API do Google está programado para terminar em 2023. Recomendamos migrar ou usar o conector V2.0 TemplateConfiguration /API do Google Drive.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Google Drive, consulteSolucionar problemas de origens de dados.

Tópicos

- Conector Google Drive V1.0
- Conector Google Drive V2.0

Conector Google Drive V1.0

O Google Drive é um serviço de armazenamento de arquivos baseado em nuvem. Você pode usar Amazon Kendra para indexar documentos e comentários armazenados nas pastas Drives compartilhados, Meus Drives e Compartilhado comigo na sua fonte de dados do Google Drive. Você pode indexar os documentos do Google Workspace e os documentos listados em Tipos de

documentação. Você também pode usar filtros de inclusão e exclusão para indexar o conteúdo por nome, tipo e caminho do arquivo.

> **Note**
>
> O suporte para o conector do Google Drive V1.0 DriveConfiguration /API do Google está programado para terminar em 2023. Recomendamos migrar ou usar o conector V2.0 TemplateConfiguration /API do Google Drive.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Google Drive, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão

Pré-requisitos

Antes de usar Amazon Kendra para indexar sua fonte de dados do Google Drive, faça essas alterações no Google Drive e AWS nas contas.

No Google Drive, verifique se você:

- Recebeu acesso para uma função de superadministrador ou é um usuário com privilégios administrativos. Você não precisa de uma função de superadministrador para você se tiver recebido acesso de uma função de superadministrador.
- Criou uma conta de serviço com a opção Ativar a delegação em todo o domínio do G Suite ativada e uma chave JSON como chave privada usando a conta.

- Copiou o e-mail da conta de usuário e o e-mail da conta de serviço. Ao se conectar, Amazon Kendra você insere o e-mail da sua conta de usuário como e-mail da conta de administrador e o e-mail da sua conta de serviço como e-mail do cliente em seu AWS Secrets Manager segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- A API Admin SDK e a API do Google Drive foram adicionadas à sua conta.
- Adicionou (ou solicitou que um usuário com uma função de superadministrador adicionasse) as seguintes permissões à conta de serviço usando uma função 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
- Verificou se cada documento é exclusivo no Google Drive e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Google Drive em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Google Drive Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Google Drive, você deve fornecer os detalhes necessários da sua fonte de dados do Google Drive para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Google Drive para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra ao Google Drive

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

   > **Note**
   >
   > Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha o Conector do Google Drive V1.0 e, em seguida, escolha Adicionar conector.
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. Para Tipo de autenticação: escolha entre Existente e Novo. Se você optar por usar um segredo existente, use Selecionar senha para escolher a senha.

b. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

- Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Google Drive-' é adicionado automaticamente ao seu nome secreto.

B. Para o e-mail da conta de administrador, o e-mail do cliente e a chave privada, insira os valores da credencial de autenticação que você gerou e baixou da conta do Google Drive.

C. Escolha Salvar autenticação.

c. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

d. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Excluir contas de usuário: os usuários do Google Drive que você deseja excluir do índice. Você pode adicionar até 100 contas de usuário.

   b. Excluir drives compartilhados: os drives compartilhados do Google Drive que você deseja excluir do índice. Você pode adicionar até 100 drives compartilhados.

   c. Excluir drives de tipos de arquivos: os drives compartilhados do Google Drive que você deseja excluir do índice. Você também pode optar por editar as seleções do tipo MIME.

   d. Configurações adicionais: especifique padrões de expressão regular para incluir ou excluir determinado conteúdo. Você pode adicionar até 100 padrões.

   e. Frequência: escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

   f. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Para nome do GoogleDrive campo e mapeamentos adicionais de campo sugeridos — Selecione entre os campos da fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Google Drive

Você deve especificar o seguinte usando a GoogleDriveConfigurationAPI:

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Google Drive. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "clientAccount": "service account email",
    "adminAccount": "user account email"",
    "privateKey": "private key"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer um IAM papel com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do Google Drive e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para fontes de dados do Google Drive.

Você também pode adicionar os seguintes recursos opcionais:

- Filtros de inclusão e exclusão, por padrão, o Amazon Kendra indexa todos os documentos no Google Drive. Você pode especificar se deseja incluir ou excluir determinados conteúdos em drives compartilhados, contas de usuário, tipos MIME de documentos e arquivos. Se você optar por excluir contas de usuário, nenhum dos arquivos no My Drive pertencentes à conta será indexado. Os arquivos compartilhados com o usuário são indexados, a menos que o proprietário do arquivo também seja excluído.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Google Drive para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você

deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice_`document_body`. Todos os demais campos são opcionais.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Google Drive, consulte:

- Primeiros passos com o conector Amazon Kendra do Google Drive

Conector Google Drive V2.0

O Google Drive é um serviço de armazenamento de arquivos baseado em nuvem. Você pode usar Amazon Kendra para indexar documentos e comentários armazenados nas pastas Drives compartilhados, Meus Drives e Compartilhado comigo na sua fonte de dados do Google Drive. Você pode indexar os documentos do Google Workspace e os documentos listados em Tipos de documentação. Você também pode usar filtros de inclusão e exclusão para indexar o conteúdo por nome, tipo e caminho do arquivo.

> **Note**
>
> O suporte para o conector do Google Drive V1.0 DriveConfiguration /API do Google está programado para terminar em 2023. Recomendamos migrar ou usar o conector V2.0 TemplateConfiguration /API do Google Drive.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Google Drive, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis

- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de usar Amazon Kendra para indexar sua fonte de dados do Google Drive, faça essas alterações no Google Drive e AWS nas contas.

No Google Drive, verifique se você:

- Recebeu acesso para uma função de superadministrador ou é um usuário com privilégios administrativos. Você não precisa de uma função de superadministrador para você se tiver recebido acesso de uma função de superadministrador.
- Credenciais de conexão da conta de serviço do Google Drive configuradas contendo o e-mail da conta de administrador, e-mail do cliente (e-mail da conta de serviço) e chave privada. Consulte a documentação do Google Cloud sobre como criar e excluir chaves de contas de serviço.

  > **Note**
  >
  > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Criou uma conta de serviço do Google Cloud (uma conta com autoridade delegada para assumir a identidade do usuário) com a ativação da Delegação em todo o domínio do G Suite para server-to-server autenticação e, em seguida, gerou uma chave privada JSON usando a conta.

> **Note**
>
> A chave privada deve ser gerada após a criação da conta de serviço.

- A API Admin SDK e a API do Google Drive foram adicionadas à sua conta de usuário.
- Opcional: configurou as credenciais de conexão do Google Drive OAuth 2.0 com o ID do cliente, a senha do cliente e o token de atualização como credenciais de conexão para um usuário específico. Você precisa disso para rastrear dados de contas individuais. Consulte a documentação do Google sobre como usar o OAuth 2.0 para acessar as APIs.
- Adicionou (ou solicitou que um usuário com uma função de superadministrador adicionasse) os seguintes escopos do OAuth à conta de serviço usando uma função de superadministrador: Esses escopos de API são necessários para rastrear todos os documentos e as informações de controle de acesso (ACL) de todos os usuários em um domínio do Google Workspace:
  - https://www.googleapis.com/auth/drive.readonly: visualize e baixe todos os arquivos do Google Drive
  - https://www.googleapis.com/auth/drive.readonly: visualize e baixe todos os arquivos do Google Drive
  - https://www.googleapis.com/auth/admin.directory.group.readonly: escopo para recuperar somente informações sobre grupos, alias de grupos e membros. Isso é necessário para o Amazon Kendra Identity Crawler.
  - https://www.googleapis.com/auth/admin.directory.user.readonly: escopo para recuperar somente usuários ou aliases de usuários. Isso é necessário para listar usuários no Amazon Kendra Identity Crawler e para definir ACLs.
  - https://www.googleapis.com/auth/cloud-platform: escopo para gerar token de acesso para buscar conteúdo de arquivos grandes do Google Drive.
  - https://www.googleapis.com/auth/forms.body.readonly: escopo para buscar dados do Google Forms.

  Para oferecer suporte à API de formulários, adicione o seguinte escopo adicional:

  - https://www.googleapis.com/auth/forms.body.readonly
- Verificou se cada documento é exclusivo no Google Drive e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Google Drive em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Google Drive Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Google Drive, você deve fornecer os detalhes necessários da sua fonte de dados do Google Drive para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Google Drive para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra ao Google Drive

1. Faça login no AWS Management Console e abra o Amazon Kendra console.

2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector do Google Drive e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector do Google Drive com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   b. Para a Autenticação, escolha entre Conta de serviço do Google e Autenticação Oauth 2.0 com base no seu caso de uso.

   c. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Google

Drive. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

i. Se você escolheu a conta de serviço do Google, insira um nome para seu segredo, o ID de e-mail do usuário administrador ou “Usuário da conta de serviço” na configuração da sua conta de serviço (e-mail do administrador), o ID de e-mail da conta do serviço (e-mail do cliente) e a chave privada que você criou na sua conta de serviço.

   Salve e adicione seu segredo

ii. Se você escolheu a autenticação OAuth 2.0, insira um nome para seu segredo, ID do cliente, segredo do cliente e token de atualização que você criou na sua conta OAuth.

   Salve e adicione seu segredo.

d. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

e. (Somente para usuários de autenticação da conta de serviço do Google)

   Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

f. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

g. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Sincronizar conteúdo — Selecione quais opções ou o conteúdo que você deseja rastrear. Você pode escolher rastrear Meu Drive (pastas pessoais), Drive compartilhado (pastas compartilhadas com você) ou ambos. Você também pode incluir comentários no arquivo.

   b. Em Configuração adicional - opcional Você também pode inserir as seguintes informações opcionais:

      i. Públicos-alvo — adicione públicos-alvo específicos para os documentos que você deseja rastrear.

      ii. Tamanho máximo do arquivo — Defina o limite máximo de tamanho em MBs de arquivos a serem rastreados.

      iii. E-mail do usuário — Adicione e-mails do usuário que você deseja incluir ou excluir.

      iv. Drives compartilhados — adicione os nomes dos drives compartilhados que você deseja incluir ou excluir.

      v. Tipos MIME — Adicione os tipos de MIME que você deseja incluir ou excluir.

      vi. Padrões de expressão regular de entidades — adicione padrões de expressão regular para incluir ou excluir determinados anexos de todas as entidades suportadas. Você pode adicionar até 100 padrões.

   c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

> **Important**
>
> A API do Google Drive não é compatível com a recuperação de comentários de um arquivo excluído permanentemente. Os comentários dos arquivos descartados podem ser recuperados. Quando um arquivo é descartado, o conector exclui os comentários do Amazon Kendra índice.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Em Histórico de execução da sincronização, escolha armazenar relatórios gerados automaticamente em um Amazon S3 ao sincronizar sua fonte de dados. Isso é útil para rastrear problemas ao sincronizar sua fonte de dados.

f. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Para arquivos — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

> **Note**
>
> A API do Google Drive não é compatível com a criação de campos personalizados. O mapeamento de campo personalizado não está disponível para o conector do Google Drive.

b. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Google Drive

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `GOOGLEDRIVEV2` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de autenticação — especifique se deseja usar a autenticação da conta de serviço ou a autenticação OAuth 2.0.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

> **Important**
>
> A API do Google Drive não é compatível com a recuperação de comentários de um arquivo excluído permanentemente. Os comentários dos arquivos descartados podem ser recuperados. Quando um arquivo é descartado, o conector exclui os comentários do Amazon Kendra índice.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contém as credenciais de autenticação que você criou na sua conta do Google Drive. Se você usar uma autenticação da conta de serviço do Google, a senha deverá conter uma estrutura JSON com as seguintes chaves:

```
{
    "clientEmail": "user account email",
    "adminAccountEmail": "service account email",
    "privateKey": "private key"
}
```

  Para usar a autenticação OAuth 2.0, a senha é armazenada em uma estrutura JSON com as seguintes chaves:

```
{
    "clientID": "OAuth client ID",
    "clientSecret": "client secret",
    "refreshToken": "refresh token"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer um IAM papel com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do Google Drive e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para fontes de dados do Google Drive.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.

- Meus drives, drives compartilhados, comentários — você pode especificar se deseja rastrear esses tipos de conteúdo.
- Filtros de inclusão e exclusão — você pode especificar se deseja incluir ou excluir determinadas contas de usuário, drives compartilhados e tipos de MIME.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Lista de controle de acesso (ACL) — Especifique se deseja rastrear as informações da ACL para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Google Drive para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você

deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice_`document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o Esquema do modelo do Google Drive.

Observações

- O mapeamento de campo personalizado não está disponível para o conector do Google Drive, pois a interface do usuário do Google Drive não é compatível com a criação de campos personalizados.
- A API do Google Drive não é compatível com a recuperação de comentários de um arquivo excluído permanentemente. Os comentários dos arquivos na lixeira podem ser recuperados. Quando um arquivo é descartado, o Amazon Kendra conector exclui os comentários do Amazon Kendra índice.
- A API do Google Drive não retorna comentários presentes em um arquivo.docx.

IBM DB2

O IBM DB2 é um sistema de gerenciamento de banco de dados relacional desenvolvido pelo IBM. Se você for um usuário do IBM DB2, poderá usar o Amazon Kendra para indexar a fonte de dados do IBM DB2. O conector da fonte de Amazon Kendra IBM DB2 dados oferece suporte ao DB2 11.5.7.

Você pode se conectar Amazon Kendra à sua fonte de IBM DB2 dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra IBM DB2 dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de IBM DB2 dados, faça essas alterações em suas IBM DB2 AWS contas.

Em IBM DB2, verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados.
- Verifique se cada documento é exclusivo em IBM DB2 e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de IBM DB2 em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de IBM DB2 dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de IBM DB2 dados, você deve fornecer detalhes de suas IBM DB2 credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou IBM DB2 para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a IBM DB2

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha IBM DB2conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o IBM DB2conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Em Fonte, insira o seguinte:

   b. Host: insira o nome do host do banco de dados.

   c. Port: insira a porta do banco de dados.

   d. Instância: insira a instância do banco de dados.

   e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

   f. Em Autenticação: insira as seguintes informações:

      - AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de IBM DB2 autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- IBM DB2 -' é adicionado automaticamente ao seu nome secreto.

II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

B. Escolha Salvar.

g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Em Sincronizar escopo, escolha uma das opções a seguir:

- Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
- Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
- Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

- Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
- Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.
- Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de registro de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
- Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
- Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear

alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a IBM DB2

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Além disso, especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `db2`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial

falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:

- `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
- `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. IBM DB2 A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o IBM DB2 conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do IBM DB2.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do IBM DB2 para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo do IBM DB2.

Observações

- As linhas excluídas do banco de dados não serão rastreadas durante a Amazon Kendra verificação do conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.

- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Jira

O Jira é uma ferramenta de gerenciamento de projetos para desenvolvimento de software, gerenciamento de produtos e rastreamento de bugs. Você pode usar Amazon Kendra para indexar seus projetos, problemas, comentários, anexos, registros de trabalho e status do Jira.

Amazon Kendra atualmente só é compatível com o Jira Cloud.

Você pode se conectar Amazon Kendra à sua fonte de dados do Jira usando o Amazon Kendra console ou a JiraConfiguration API. Para obter uma lista de recursos compatíveis por cada um, consulte Atributos compatíveis.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Jira, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra O conector de fonte de dados do Jira é compatível com os seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Jira, faça essas alterações no Jira e AWS nas contas.

No Jira, verifique se você:

- Credenciais de autenticação de token de API configuradas, que incluem um ID do Jira (nome de usuário ou e-mail) e uma credencial do Jira (token da API do Jira). Consulte a documentação da Atlassian sobre o gerenciamento de tokens de API.

  > **Note**
  >
  > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Anote o URL da conta do Jira nas configurações da sua conta do Jira. Por exemplo, *`https://company.atlassian.net/`*.
- Verifique se cada documento é exclusivo no Jira e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

  > **Note**
  >
  > Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Jira em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados do Jira a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Jira, você deve fornecer os detalhes necessários da sua fonte de dados do Jira para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Jira para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao Jira

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

   > **Note**
   >
   > Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector Jira e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Jira com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. URL da conta do Jira — Insira a URL da sua conta do Jira. Por exemplo, *https://company.atlassian.net/*.

b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

c. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Jira. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Jira-' é adicionado automaticamente ao seu nome secreto.

B. Para Jira ID: insira o nome de usuário ou e-mail do Jira.

C. Em Senha/Token — Insira o token da API do Jira configurado no Jira.

ii. Salve e adicione seu segredo.

d. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

e. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

f. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

g. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Selecione quais projetos do Jira indexar — Escolha rastrear todos os projetos ou projetos específicos.

b. Configuração adicional — especifique determinados status e tipos de problemas. Escolha rastrear comentários, anexos e registros de trabalho. Use padrões de expressão regular para incluir ou excluir determinados conteúdos.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.

- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Jira

Você deve especificar o seguinte usando a JiraConfiguration API:

- URL da fonte de dados: especifique o URL da conta do Jira. Por exemplo, *https://company.atlassian.net/*.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Jira. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "jiraId": "Jira user name or email",
    "jiraCredential": "Jira API token"
}
```

- IAM role — Especifique `RoleArn` quando você chama `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Jira e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para a fontes de dados do Jira.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique `VpcConfiguration` como parte da configuração da fonte de dados. Consulte Configuração do Amazon Kendra para usar uma VPC.
- Registro de alterações — Se Amazon Kendra deve usar o mecanismo de registro de alterações da fonte de dados do Jira para determinar se um documento deve ser atualizado no índice.

> **Note**
>
> Use o log de alterações se o Amazon Kendra não quiser digitalizar todos os documentos. Se o registro de alterações for grande, talvez leve Amazon Kendra menos tempo para digitalizar os documentos na fonte de dados do Jira do que para processar o registro de alterações. Se estiver sincronizando a fonte de dados do Jira com o índice pela primeira vez, todos os documentos serão digitalizados.

- Filtros de inclusão e exclusão — Você pode especificar se deseja incluir ou excluir determinados arquivos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Comentários, anexos e registros de trabalho — você pode especificar se deseja rastrear determinados comentários, anexos e registros de trabalho de problemas.
- Projetos, problemas, status — você pode especificar se deseja rastrear determinados IDs de projetos, tipos de problemas e status.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Jira para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Jira, consulte:

- Pesquise de forma inteligente seus projetos do Jira com o conector Amazon Kendra Jira Cloud

Microsoft Exchange

O Microsoft Exchange é uma ferramenta de colaboração corporativa para mensagens, reuniões e compartilhamento de arquivos. Se você for um usuário do Microsoft Exchange, poderá usar Amazon Kendra para indexar sua fonte de dados do Microsoft Exchange.

Você pode se conectar Amazon Kendra à sua fonte de dados do Microsoft Exchange usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Microsoft Exchange, consulteSolucionar problemas de origens de dados.

Atributos compatíveis

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Microsoft Exchange, faça essas alterações no Microsoft Exchange e AWS nas contas.

No Microsoft Exchange, verifique se você:

- Criou uma conta do Microsoft Exchange no Office 365.
- Anotou o ID de inquilino do Microsoft 365. Encontre o ID de inquilino nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
- Configurei um aplicativo OAuth no portal do Azure e anotei a ID e o segredo do cliente ou as credenciais do cliente. Consulte o tutorial da Microsoft e o exemplo de aplicativo registrado para obter mais informações.

> **Note**
>
> Quando você cria ou registra um aplicativo no portal do Azure, a ID secreta representa o valor secreto real. Você deve anotar ou salvar o valor real do segredo imediatamente ao criar o segredo e o aplicativo. Você pode acessar seu segredo selecionando o nome do seu aplicativo no portal do Azure e navegando até a opção de menu sobre certificados e segredos.
> Você pode acessar sua ID de cliente selecionando o nome do seu aplicativo no portal do Azure e navegando até a página de visão geral. O ID do aplicativo (cliente) é o ID do cliente.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Foram adicionadas as seguintes permissões para o aplicativo conector:

Microsoft Graph	Office 365 Exchange Online
• Mail.Read (Aplicativo) • Correio. ReadBasic (Aplicação) • Correio. ReadBasic.All (Aplicativo) • Calendars.Read (Aplicativo) • User.Read.All (Aplicativo) • Contacts.Read (Aplicativo) • Notes.Read.All (Aplicativo) • Directory.Read.All (Aplicativo) • NOTÍCIAS. AccessAsUser.Tudo (delegado)	full_access_as_app (Aplicativo)

- Verifique se cada documento é exclusivo no Microsoft Exchange e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Microsoft Exchange em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Microsoft Exchange Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Microsoft Exchange, você deve fornecer os detalhes necessários da sua fonte de dados do Microsoft Exchange para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Microsoft Exchange para Amazon Kendra, consulte Pré-requisitos.

Console

Para se conectar Amazon Kendra ao Microsoft Exchange

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector do Microsoft Exchange e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Microsoft Exchange com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. ID do inquilino — insira seu ID de inquilino do Microsoft 365. Encontre o ID de inquilino nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
   b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
   c. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Microsoft Exchange. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Microsoft Exchange

B. Para ID do cliente, segredo do cliente — insira as credenciais de autenticação configuradas no Microsoft Exchange no portal do Azure.

ii. Salve e adicione seu segredo.

d. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

e. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

f. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. IDs de usuário — Forneça os e-mails do usuário se quiser filtrar o conteúdo por determinados e-mails.

b. Configuração adicional — especifique os tipos de conteúdo que você deseja rastrear.

- Tipos de entidade — Você pode escolher rastrear o conteúdo do calendário ou dos OneNotes contatos.
- Rastreamento do calendário — insira as datas de início e término para rastrear o conteúdo entre determinadas datas.
- Incluir e-mail — Digite “para”, “de” e linhas de assunto do e-mail para filtrar determinados e-mails que você deseja rastrear.
- Acesso a pastas compartilhadas — Escolha ativar o rastreamento da lista de controle de acesso para controle de acesso à sua fonte de dados do Microsoft Exchange.

- Regex para domínios — adicione padrões de expressão regular para incluir ou excluir determinados domínios de e-mail.
- Padrões Regex: adicionar padrões de expressão regular para incluir ou excluir determinados arquivos.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

> **Note**
>
> O conector da fonte de dados do Amazon Kendra Microsoft Exchange não oferece suporte a mapeamentos de campo personalizados.

b. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Microsoft Exchange

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `MSEXCHANGE` quando você usa o esquema TemplateConfigurationJSON. Além disso, especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- ID do inquilino: encontre o ID de inquilino nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Microsoft Exchange. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do Microsoft Exchange e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Microsoft Exchange.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados tipos de conteúdo.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Lista de controle de acesso (ACL) — Especifique se deseja rastrear as informações da ACL para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Microsoft Exchange para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para obter uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo do Microsoft Exchange.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Microsoft Exchange, consulte:

- Indexe o conteúdo do Microsoft Exchange usando o conector do Exchange para o Amazon Kendra

Microsoft OneDrive

A Microsoft OneDrive é um serviço de armazenamento baseado em nuvem que você pode usar para armazenar, compartilhar e hospedar seu conteúdo. Você pode usar Amazon Kendra para indexar sua fonte OneDrive de dados.

Você pode se conectar Amazon Kendra à sua fonte de OneDrive dados usando o Amazon Kendra console e a OneDriveConfigurationAPI.

Amazon Kendra tem duas versões do OneDrive conector. Os recursos suportados de cada versão incluem:

OneDrive Conector Microsoft V1.0/API OneDriveConfiguration

- Mapeamentos de campos
- Filtros de inclusão/exclusão

OneDrive Conector Microsoft V2.0/API TemplateConfiguration

- Filtragem de contexto do usuário
- Rastreador de identidade de usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

> **Note**
>
> O suporte para o OneDrive conector V1.0 OneDriveConfiguration /API está programado para terminar em junho de 2023. Recomendamos usar o OneDrive conector V2.0/ TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra OneDrive dados, consulteSolucionar problemas de origens de dados.

Tópicos

- OneDrive Conector Microsoft V1.0
- OneDrive Conector Microsoft V2.0
- Saiba mais

OneDrive Conector Microsoft V1.0

A Microsoft OneDrive é um serviço de armazenamento baseado em nuvem que você pode usar para armazenar, compartilhar e hospedar seu conteúdo. Você pode usar Amazon Kendra para indexar sua fonte de OneDrive dados da Microsoft.

> **Note**
>
> O suporte para o OneDrive conector V1.0 OneDrive /API da Microsft está programado para terminar em junho de 2023. Recomendamos usar o OneDrive conector V2.0/ TemplateConfiguration API.

Para solucionar problemas do conector da fonte de Amazon Kendra OneDrive dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão

Atributos compatíveis

- Mapeamentos de campos
- Filtros de inclusão/exclusão

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de OneDrive dados, faça essas alterações em suas OneDrive AWS contas.

No Azure Active Directory (AD), verifique se você:

- Criou um aplicativo Azure Active Directory (AD).
- Usou o ID do aplicativo AD para registrar uma chave secreta para o aplicativo no site do AD. A chave secreta deve conter o ID do aplicativo e uma chave secreta.
- Copiou o domínio AD da organização.
- As seguintes permissões de aplicativo foram adicionadas ao seu aplicativo AD na opção Microsoft Graph:
  - Leia arquivos em todos os conjuntos de sites (File.Read.All)
  - Leia o perfil completo de todos os usuários (User.Read.All)
  - Leia os dados do diretório (Directory.Read.All)
  - Leia todos os grupos (Group.Read.All)
  - Leia itens em todos os conjuntos de sites (Site.Read.All)
- Copiou a lista de usuários cujos documentos devem ser indexados. Você pode optar por fornecer uma lista de nomes de usuário ou pode fornecer os nomes de usuário em um arquivo armazenado em um Amazon S3. Depois de criar a fonte de dados, você poderá:
  - Modifique a lista de usuários.
  - Mude de uma lista de usuários para uma lista armazenada em um Amazon S3 bucket.

- Altere a localização do Amazon S3 bucket de uma lista de usuários. Se você alterar a localização do bucket, também deverá atualizar a IAM função da fonte de dados para que ela tenha acesso ao bucket.

  > **Note**
  >
  > Se você armazenar a lista de nomes de usuário em um Amazon S3 bucket, a IAM política da fonte de dados deverá fornecer acesso ao bucket e acesso à chave com a qual o bucket foi criptografado, se houver.

- Verifique se cada documento é exclusivo em OneDrive e entre outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

  > **Note**
  >
  > Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de OneDrive autenticação em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

  > **Note**
  >
  > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de OneDrive dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de OneDrive dados, você deve fornecer detalhes de suas OneDrive credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou OneDrive para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a OneDrive

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha OneDrive conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o OneDrive conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. OneDrive ID do inquilino — Insira o ID do OneDrive inquilino sem o protocolo.

   b. Tipo de autenticação: escolha entre Novo e Existente.

   c. i. Se você escolher Existente, selecione uma senha existente em Selecionar senha.

      ii. Se você escolher Novo, insira as seguintes informações na seção Nova senha AWS Secrets Manager :

         A. Senha: um nome para sua senha. O prefixo 'AmazonKendra- OneDrive -' é adicionado automaticamente ao seu nome secreto.

         B. Para ID do aplicativo e senha do aplicativo — insira os valores da credencial de autenticação da sua OneDrive conta e escolha Salvar autenticação.

   d. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

   e. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Escolha entre Arquivo da lista e Lista de nomes com base no caso de uso.

      i. Se você escolher o Arquivo da lista, insira as seguintes informações:

         - Selecione o local e insira o caminho para o bucket do Amazon S3 .

           Adicionar arquivo de lista de usuários a Amazon S3 —Selecione para adicionar seus arquivos de lista de usuários ao seu Amazon S3 bucket.

           Mapeamentos de grupos locais de usuários: selecione para usar o mapeamento de grupos locais para filtrar seu conteúdo.

      ii. Se você escolher o Lista de nomes, insira as seguintes informações:

- Nome de usuário: insira até 10 drives de usuário para indexar. Para adicionar mais de 10 usuários, crie um arquivo que contenha os nomes.

  Selecione Adicionar outro: escolha adicionar mais usuários.

  Mapeamentos de grupos locais de usuários: selecione para usar o mapeamento de grupos locais para filtrar seu conteúdo.

b. Para configuração adicional, especifique padrões de expressão regular para incluir ou excluir determinados arquivos. Você pode adicionar até 100 padrões.

c. Em Cronograma de execução da sincronização, em Frequência — Escolha com que frequência Amazon Kendra será sincronizada com sua fonte de dados.

d. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Para Campos de fonte de dados padrão e mapeamentos de campo adicionais sugeridos, selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

   b. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a OneDrive

Você deve especificar o seguinte usando a OneDriveConfigurationAPI:

- ID do inquilino: o domínio do Azure Active Directory da organização.
- OneDrive Usuários — especifique a lista de contas de usuário cujos documentos devem ser indexados.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta. OneDrive A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "username": "OAuth client ID",
    "password": "client secret"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o OneDrive conector e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes OneDrive de dados.

Você também pode adicionar os seguintes recursos opcionais:

- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados documentos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Mapeamentos de campo — Escolha mapear os campos da fonte de OneDrive dados para os Amazon Kendra campos de índice. Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos.

As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

OneDrive Conector Microsoft V2.0

A Microsoft OneDrive é um serviço de armazenamento baseado em nuvem que você pode usar para armazenar, compartilhar e hospedar seu conteúdo. Você pode usar Amazon Kendra para indexar sua fonte OneDrive de dados.

Você pode se conectar Amazon Kendra à sua fonte de OneDrive dados usando o Amazon Kendra console e a OneDriveConfigurationAPI.

> **Note**
>
> O suporte para o OneDrive Connector V1.0/ OneDriveConfiguration API está programado para terminar em junho de 2023. Recomendamos usar o OneDrive Connector V2.0/ TemplateConfiguration API. A versão 2.0 fornece ACLs adicionais e funcionalidade de crawler de identidade.

Para solucionar problemas do conector da fonte de Amazon Kendra OneDrive dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão

Atributos compatíveis

Amazon Kendra OneDrive o conector de fonte de dados oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais

- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de OneDrive dados, faça essas alterações em suas OneDrive AWS contas.

Em OneDrive, verifique se você tem:

- Criou uma OneDrive conta no Office 365.
- Anotou o ID de inquilino do Microsoft 365. Encontre o ID de inquilino nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
- Criou um aplicativo OAuth no portal do Azure e anotou a ID do cliente e o segredo do cliente ou as credenciais do cliente usadas para autenticação com um segredo. AWS Secrets Manager Consulte o tutorial da Microsoft e o exemplo de aplicativo registrado para obter mais informações.

> **Note**
>
> Quando você cria ou registra um aplicativo no portal do Azure, a ID secreta representa o valor secreto real. Você deve anotar ou salvar o valor real do segredo imediatamente ao criar o segredo e o aplicativo. Você pode acessar seu segredo selecionando o nome do seu aplicativo no portal do Azure e navegando até a opção de menu sobre certificados e segredos.
> Você pode acessar sua ID de cliente selecionando o nome do seu aplicativo no portal do Azure e navegando até a página de visão geral. O ID do aplicativo (cliente) é o ID do cliente.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Usou o ID do aplicativo AD para registrar uma chave secreta para o aplicativo no site do AD. A chave secreta deve conter o ID do aplicativo e uma chave secreta.

- Copiou o domínio AD da organização.
- As seguintes permissões foram adicionadas ao seu aplicativo AD na opção Microsoft Graph:
  - Leia arquivos em todos os conjuntos de sites (File.Read.All)
  - Leia o perfil completo de todos os usuários (User.Read.All)
  - Leia todos os grupos (Group.Read.All)
  - Leia todas as notas (Notes.Read.All)
- Copiou a lista de usuários cujos documentos devem ser indexados. Você pode optar por fornecer uma lista de nomes de usuário ou pode fornecer os nomes de usuário em um arquivo armazenado em um Amazon S3. Depois de criar a fonte de dados, você poderá:
  - Modifique a lista de usuários.
  - Mude de uma lista de usuários para uma lista armazenada em um Amazon S3 bucket.
  - Altere a localização do Amazon S3 bucket de uma lista de usuários. Se você alterar a localização do bucket, também deverá atualizar a IAM função da fonte de dados para que ela tenha acesso ao bucket.

> **Note**
>
> Se você armazenar a lista de nomes de usuário em um Amazon S3 bucket, a IAM política da fonte de dados deverá fornecer acesso ao bucket e acesso à chave com a qual o bucket foi criptografado, se houver.
> O OneDrive conector usa e-mail das informações de contato presentes nas propriedades do usuário do Onedrive. Certifique-se de que o usuário cujos dados você deseja rastrear tenha o campo de e-mail configurado na página Informações de contato, pois para novos usuários, isso pode estar em branco.

Em sua AWS conta, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.
- Armazenou suas credenciais de OneDrive autenticação em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de OneDrive dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de OneDrive dados, você deve fornecer detalhes de suas OneDrive credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou OneDrive para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra a OneDrive

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha OneDrive conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o OneDrive conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. OneDrive ID do inquilino — Insira o ID do OneDrive inquilino sem o protocolo.

   b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   c. Em Autenticação: escolha entre Novo e Existente..

   d. i. Se você escolher Existente, selecione uma senha existente em Selecionar senha.

      ii. Se você escolher Novo, insira as seguintes informações na seção Nova senha AWS Secrets Manager :

         A. Senha: um nome para sua senha. O prefixo 'AmazonKendra- OneDrive -' é adicionado automaticamente ao seu nome secreto.

         B. Para ID do cliente e segredo do cliente — insira o ID do cliente e o segredo do cliente.

   e. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

   f. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

   g. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

h. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

8. a. Para Escopo de sincronização — Escolha quais OneDrive dados dos usuários serão indexados. Você pode adicionar no máximo 10 usuários manualmente.

   b. Para as Configurações adicionais, adicione padrões de expressão regular para incluir ou excluir determinados arquivos. Você pode adicionar até 100 padrões.

   c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

      - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
      - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
      - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

   d. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

   e. Escolha Próximo.

9. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Escolha Próximo.

10. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a OneDrive

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `ONEDRIVEV2` quando você usa o esquema TemplateConfigurationJSON. Além disso, especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- ID do inquilino: especifique o ID do inquilino do Microsoft 365. Encontre o ID de inquilino nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte

de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. OneDrive

  Para usar a autenticação OAuth 2.0, a senha é armazenada em uma estrutura JSON com as seguintes chaves:

  ```
  {
      "clientId": "client ID",
      "clientSecret": "client secret"
  }
  ```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o OneDrive conector e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes OneDrive de dados.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão — você pode especificar se deseja incluir ou excluir determinados arquivos, OneNote seções e OneNote páginas.

  > **Note**
  >
  > A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL)

dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

- Mapeamentos de campo — Você só pode mapear campos de índice incorporados ou comuns para o conector. Amazon Kendra OneDrive O mapeamento de campo personalizado não está disponível para o OneDrive conector devido às limitações da API. Para obter mais informações, consulte Mapear campos de fonte de dados.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o esquema OneDrive do modelo.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte OneDrive de dados, consulte:

- Anunciando o OneDrive conector Microsoft atualizado (V2) para o. Amazon Kendra

Microsoft SharePoint

SharePoint é um serviço colaborativo de criação de sites que você pode usar para personalizar o conteúdo da Web e criar páginas, sites, bibliotecas de documentos e listas. Você pode usar Amazon Kendra para indexar sua fonte SharePoint de dados.

Amazon Kendra atualmente oferece suporte ao SharePoint Online e ao SharePoint Server (versões 2013, 2016, 2019 e Subscription Edition).

Você pode se conectar Amazon Kendra à sua fonte de SharePoint dados usando o Amazon Kendra console, a TemplateConfigurationAPI ou a SharePointConfigurationAPI.

Amazon Kendra tem duas versões do SharePoint conector. Os recursos suportados de cada versão incluem:

SharePoint Conector V1.0/API SharePointConfiguration

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Log de alterações
- Nuvem privada virtual (VPC)

SharePoint Conector V2.0/API TemplateConfiguration

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

> **Note**
>
> O suporte para o SharePoint conector V1.0 SharePointConfiguration /API está programado para terminar em 2023. Recomendamos migrar para ou usar o SharePoint conector V2.0/ TemplateConfiguration API.

Para solucionar problemas do conector da fonte de Amazon Kendra SharePoint dados, consulteSolucionar problemas de origens de dados.

Tópicos

- SharePoint conector V1.0
- SharePoint conector V2.0

SharePoint conector V1.0

SharePoint é um serviço colaborativo de criação de sites que você pode usar para personalizar o conteúdo da Web e criar páginas, sites, bibliotecas de documentos e listas. Se você for um SharePoint usuário, poderá usar Amazon Kendra para indexar sua fonte SharePoint de dados.

> **Note**
>
> O suporte para o SharePoint conector V1.0 SharePointConfiguration /API está programado para terminar em 2023. Recomendamos migrar para ou usar o SharePoint conector V2.0/ TemplateConfiguration API.

Para solucionar problemas do conector da fonte de Amazon Kendra SharePoint dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Log de alterações
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de SharePoint dados, faça essas alterações em suas SharePoint AWS contas.

Você precisa fornecer credenciais de autenticação, que você armazena com segurança em um segredo. AWS Secrets Manager

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não

recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Em SharePoint, verifique se você tem:

- Anote o URL dos SharePoint sites que você deseja indexar.
- Para SharePoint online:
  - Anotou as credenciais de autenticação básicas, que incluem um nome de usuário e uma senha com no mínimo permissão de administrador
  - Opcional: gerou credenciais do OAuth 2.0 como nome de usuário, senha, ID do cliente e senha do cliente.
  - Padrões de segurança desativados no portal do Azure usando um usuário administrativo. Para obter mais informações sobre como gerenciar as configurações padrão de segurança no portal do Azure, consulte a documentação da Microsoft sobre como habilitar/desabilitar padrões de segurança.
- Para SharePoint servidor:
  - Anote o nome de domínio do seu SharePoint servidor (o nome NetBIOS no seu Active Directory). Você usa isso, junto com seu nome de usuário e senha de autenticação SharePoint básica, para conectar o SharePoint Servidor Amazon Kendra a.

> **Note**
>
> Se você usa o SharePoint Servidor e precisa converter sua Lista de Controle de Acesso (ACL) para o formato de e-mail para filtragem no contexto do usuário, forneça a URL do servidor LDAP e a base de pesquisa LDAP. Ou você pode usar a substituição de domínio do diretório. O URL do servidor LDAP é o nome completo do domínio e o número da porta (por exemplo, ldap://example.com:389). A base de pesquisa do LDAP são os controladores de domínio “example” e “com”. Com a substituição do domínio do diretório, você pode usar o domínio de e-mail em vez de usar o URL do servidor LDAP e a base de pesquisa LDAP. Por exemplo, o domínio do e-mail de username@example.com é “exemplo.com”. Você pode usar essa substituição se não estiver preocupado em validar o domínio e simplesmente quiser usar seu domínio de e-mail.

- As seguintes permissões foram adicionadas à sua SharePoint conta:

Para SharePoint listas

- Itens abertos: visualize a origem dos documentos com manipuladores de arquivos do lado do servidor.
- Exibir páginas do aplicativo: visualize formulários, visualizações e páginas do aplicativo. Enumere as listas.
- Exibir itens: visualize itens em listas e documentos na bibliotecas de documentos.
- Exibir versões: visualize versões anteriores de um item de lista ou documento.

Para SharePoint sites

- Procurar diretórios — Enumere arquivos e pastas em um site usando o Designer e a interface Web DAV. SharePoint
- Procurar informações do usuário: visualize informações sobre os usuários do site.
- Enumerar permissões: enumere as permissões no site, na lista, na pasta, no documento ou no item da lista.
- Abrir: abra um site, lista ou pasta para acessar itens dentro do contêiner.
- Use os recursos de integração do cliente — Use SOAP, WebDAV, o modelo de objeto do cliente ou SharePoint as interfaces do Designer para acessar o site.
- Use interfaces remotas: use recursos que iniciam aplicativos cliente.
- Exibir páginas: exibir páginas em um site.

- Verifique se cada documento é exclusivo em SharePoint e entre outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de SharePoint autenticação em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de SharePoint dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de SharePoint dados, você deve fornecer detalhes de suas SharePoint credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou SharePoint para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a SharePoint

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha SharePoint conector v1.0 e, em seguida, escolha Adicionar fonte de dados.
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. Para o método de hospedagem — Escolha entre SharePoint Online e SharePointServidor.
      i. Para SharePointon-line — insira os URLs do site específicos do seu SharePoint repositório.
      ii. Para SharePointServidor — Escolha sua SharePoint versão, insira URLs de sites específicos para seu SharePoint repositório e insira o Amazon S3 caminho para a localização do seu certificado SSL.
   b. (Somente SharePoint servidor) Para proxy da Web — insira o nome do host e o número da porta da sua SharePoint instância interna. O número da porta deve ser um valor numérico entre 0 e 65535.
   c. Para autenticação, escolha entre as seguintes opções com base no caso de uso:

i. Para SharePoint Online—Escolha entre a autenticação básica e a autenticação OAuth 2.0.

ii. Para SharePoint Servidor — Escolha entre Nenhum, LDAP e Manual.

d. Por AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de SharePoint autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta. Você deve inserir uma Senha. O prefixo 'AmazonKendra- SharePoint -' é adicionado automaticamente ao seu nome secreto.

e. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

i. Escolha entre as seguintes opções de autenticação SharePoint na nuvem, com base no seu caso de uso:

A. Autenticação básica — insira o nome de usuário SharePoint da sua conta como Nome de usuário e a senha SharePoint da conta como Senha.

B. Autenticação OAuth 2.0 — Insira o nome de usuário da sua SharePoint conta como Nome de usuário, a senha da SharePoint conta como Senha, seu SharePoint ID exclusivo gerado automaticamente como ID do cliente e a string secreta compartilhada usada por ambos SharePoint e Amazon Kendra como segredo do cliente.

ii. Escolha entre as seguintes opções de autenticação SharePoint do servidor, com base no seu caso de uso:

A. Nenhuma — Insira o nome de usuário SharePoint da conta como Nome de usuário, a senha SharePoint da conta como Senha e o nome de domínio do servidor.

B. LDAP — ***Insira o nome de usuário da SharePoint conta como Nome de usuário, a senha da SharePoint conta como Senha, o endpoint do servidor LDAP (incluindo o protocolo e o número da porta, por exemplo, ldap: //example.com:389) e sua base de pesquisa LDAP (por exemplo, dc=example, dc=com).***

C. Manual — Insira o nome de usuário da SharePoint conta como nome de usuário, a senha SharePoint da conta como senha e a substituição do domínio de e-mail (domínio de e-mail do usuário ou grupo do diretório).

iii. Escolha Salvar.

f. nuvem privada virtual (VPC): adicione também sub-redes e grupos de segurança de VPC.

> **Note**
>
> Você deve usar uma VPC se usar SharePoint o Server. Amazon VPC é opcional para outras SharePoint versões.

g. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

h. Escolha Próximo.

7. Na página Configurações de sincronização, insira as seguintes informações:

   a. Usar log de alterações: selecione para atualizar o índice em vez de sincronizar todos os arquivos.

   b. Rastrear anexos: selecione para rastrear anexos.

   c. Usar mapeamentos de grupos locais: selecione para garantir que os documentos sejam filtrados adequadamente.

   d. Configuração adicional: adicione padrões de expressão regular para incluir ou excluir determinados arquivos. Você pode adicionar até 100 padrões.

   e. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

   f. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Amazon Kendra mapeamentos de campo padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Para Mapeamentos de campo personalizados, adicione campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a SharePoint

Você deve especificar o seguinte usando a SharePointConfigurationAPI:

- SharePointVersão — especifique a SharePoint versão que você usa ao configurar SharePoint. Esse é o caso, não importa se você usa o SharePoint Server 2013, o SharePoint Server 2016, o SharePoint Server 2019 ou o SharePoint Online.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contém as credenciais de autenticação que você criou em sua SharePoint conta. O segredo é armazenado em uma estrutura JSON.

  Para a autenticação básica SharePoint on-line, a seguinte é a estrutura JSON mínima que deve estar em seu segredo:

```
{
    "userName": "user name",
    "password": "password"
}
```

  Para a autenticação SharePoint online do OAuth 2.0, a seguinte é a estrutura JSON mínima que deve estar em seu segredo:

```
{
    "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 a autenticação básica do SharePoint servidor, a seguinte é a estrutura JSON mínima que deve estar em seu segredo:

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name"
}
```

Para a autenticação LDAP SharePoint do servidor (se você precisar converter sua lista de controle de acesso (ACL) para o formato de e-mail para filtrar no contexto do usuário, você pode incluir a URL do servidor LDAP e a base de pesquisa LDAP em seu segredo), a seguir está a estrutura JSON mínima que deve estar em seu segredo:

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name"
    "ldapServerUrl": "ldap://example.com:389",
    "ldapSearchBase": "dc=example,dc=com"
}
```

Para a autenticação manual do SharePoint servidor, a seguinte é a estrutura JSON mínima que deve estar em seu segredo:

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name",
    "emailDomainOverride": "example.com"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o SharePoint conector e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes SharePoint de dados.

- Amazon VPC—Se você usa o SharePoint Server, especifique `VpcConfiguration` como parte da configuração da fonte de dados. Consulte Configuração Amazon Kendra para usar uma VPC.

Você também pode adicionar os seguintes recursos opcionais:

- Proxy da Web — Se você deve se conectar aos URLs SharePoint do seu site por meio de um proxy da Web. Você pode usar essa opção somente para SharePoint Servidor.
- Listas de indexação — Se o conteúdo dos anexos Amazon Kendra deve ser indexado aos itens da lista. SharePoint
- Registro de alterações — Se Amazon Kendra deve usar o mecanismo de registro de alterações da fonte de SharePoint dados para determinar se um documento deve ser atualizado no índice.

> **Note**
>
> Use o log de alterações se o Amazon Kendra não quiser digitalizar todos os documentos. Se o registro de alterações for grande, talvez leve Amazon Kendra menos tempo para digitalizar os documentos na fonte de SharePoint dados do que para processar o registro de alterações. Se você estiver sincronizando sua fonte de SharePoint dados com seu índice pela primeira vez, todos os documentos serão digitalizados.

- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados tipos de conteúdo.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Mapeamentos de campo — Escolha mapear os campos da fonte de SharePoint dados para os Amazon Kendra campos de índice. Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice `_document_body`. Todos os demais campos são opcionais.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte SharePoint de dados, consulte:

- Introdução ao conector Amazon Kendra SharePoint online

SharePoint conector V2.0

SharePoint é um serviço colaborativo de criação de sites que você pode usar para personalizar o conteúdo da Web e criar páginas, sites, bibliotecas de documentos e listas. Você pode usar Amazon Kendra para indexar sua fonte SharePoint de dados.

Amazon Kendra atualmente suporta SharePoint Online e SharePoint Server (2013, 2016, 2019 e Subscription Edition).

> **Note**
>
> O suporte para o SharePoint conector V1.0 SharePointConfiguration /API está programado para terminar em 2023. Recomendamos migrar para ou usar o SharePoint conector V2.0/ TemplateConfiguration API.

Para solucionar problemas do conector da fonte de Amazon Kendra SharePoint dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

Amazon Kendra SharePoint o conector de fonte de dados oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de SharePoint dados, faça essas alterações em suas SharePoint AWS contas.

Você precisa fornecer credenciais de autenticação, que você armazena com segurança em um segredo. AWS Secrets Manager

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

No SharePoint Online, verifique se você tem:

- Copiou os URLs da sua SharePoint instância. O formato do URL do host que você insere é *`https://yourdomain.sharepoint.com/sites/mysite`*. O URL deve começar com `https` e conter `sharepoint.com`.

- Copiou o nome de domínio do URL da sua SharePoint instância.
- Anote suas credenciais básicas de autenticação contendo o nome de usuário e a senha com permissões de administrador do site para se conectar ao SharePoint Online.
- Padrões de segurança desativados no portal do Azure usando um usuário administrativo. Para obter mais informações sobre como gerenciar as configurações padrão de segurança no portal do Azure, consulte a documentação da Microsoft sobre como habilitar/desabilitar padrões de segurança.
- Autenticação multifator (MFA) desativada em sua SharePoint conta, para que ela não Amazon Kendra seja impedida de rastrear seu conteúdo. SharePoint
- Se estiver usando um tipo de autenticação diferente da autenticação básica: copiou o ID do locatário da sua SharePoint instância. Para obter detalhes sobre como encontrar o ID de inquilino, consulte Encontre o ID de inquilino do Microsoft 365.
- Se você precisar migrar para a autenticação de usuário na nuvem com o Microsoft Entra, consulte a documentação da Microsoft sobre autenticação na nuvem.
- Para autenticação OAuth 2.0 e autenticação de token de atualização do OAuth 2.0: anote suas credenciais de autenticação básica contendo o nome de usuário e a senha que você usa para se conectar ao SharePoint Online e o ID do cliente e o segredo do cliente gerados após o registro no Azure AD. SharePoint
  - Se não estiver usando a ACL, adicione as seguintes permissões:

Microsoft Graph	SharePoint
• Notes.Read.All (Aplicativo) — Leia todos os cadernos OneNote • Sites.Read.All (Application): leia itens em todos os conjuntos de sites (Site.Read.All)	• AllSites.Ler (delegado) — Leia itens em todos os conjuntos de sites

> **Note**
>
> Note.Read.All e Sites.Read.All são necessários somente se você quiser rastrear documentos. OneNote
> Se você quiser rastrear sites específicos, a permissão pode ser restrita a sites específicos, em vez de a todos os sites disponíveis no domínio. Você configura a permissão Sites.Selected (Aplicativo). Com essa permissão de API, você precisa definir

explicitamente a permissão de acesso em cada site por meio da API do Microsoft Graph. Para obter mais informações, consulte o blog da Microsoft em Sites.Selected permissions.

- Se não estiver usando a ACL, adicione as seguintes permissões:

Microsoft Graph	SharePoint
• Group.Member.Read.All (Aplicativo): leia todas as associações do grupo • Notes.Read.All (Aplicativo) — Leia todos os cadernos OneNote • Sites. FullControl.Tudo (delegado) — Necessário para recuperar ACLs dos documentos • Sites.Read.All (Application): leia itens em todos os conjuntos de sites (Site.Read.All) • User.Read.All (Application): leia o perfil completo de todos os usuários (User.Read.All)	• AllSites.Ler (delegado) — Leia itens em todos os conjuntos de sites

Note

GroupMember.Read.All e User.Read.All são necessários somente se o Identity Crawler estiver ativado.
Se você quiser rastrear sites específicos, a permissão pode ser restrita a sites específicos, em vez de a todos os sites disponíveis no domínio. Você configura a permissão Sites.Selected (Aplicativo). Com essa permissão de API, você precisa definir explicitamente a permissão de acesso em cada site por meio da API do Microsoft Graph. Para obter mais informações, consulte o blog da Microsoft em Sites.Selected permissions.

- Para autenticação somente do aplicativo Azure AD: chave privada e a ID do cliente que você gerou após se registrar SharePoint no Azure AD. Observe também o certificado X.509.
  - Se não estiver usando a ACL, adicione as seguintes permissões:

SharePoint
• Sites.Read.All (Aplicativo) — Obrigatório para acessar itens e listas em todos os conjuntos de sites

> **Note**
>
> Se você quiser rastrear sites específicos, a permissão pode ser restrita a sites específicos, em vez de a todos os sites disponíveis no domínio. Você configura a permissão Sites.Selected (Aplicativo). Com essa permissão de API, você precisa definir explicitamente a permissão de acesso em cada site por meio da API do Microsoft Graph. Para obter mais informações, consulte o blog da Microsoft em Sites.Selected permissions.

- Se não estiver usando a ACL, adicione as seguintes permissões:

SharePoint
• Sites. FullControl.All (Aplicativo) — Necessário para recuperar ACLs dos documentos

> **Note**
>
> Se você quiser rastrear sites específicos, a permissão pode ser restrita a sites específicos, em vez de a todos os sites disponíveis no domínio. Você configura a permissão Sites.Selected (Aplicativo). Com essa permissão de API, você precisa definir explicitamente a permissão de acesso em cada site por meio da API do Microsoft Graph. Para obter mais informações, consulte o blog da Microsoft em Sites.Selected permissions.

- Para autenticação SharePoint somente do aplicativo: anote sua ID SharePoint do cliente e segredo do cliente gerados ao conceder permissão somente ao SharePoint aplicativo, e sua ID do cliente e segredo do cliente gerados quando você registrou seu SharePoint aplicativo no Azure AD.

> **Note**
>
> SharePoint A autenticação somente de aplicativos não é compatível com a versão SharePoint 2013.

- (Opcional) Se você estiver rastreando OneNote documentos e usando o Identity Crawler, adicione as seguintes permissões:

Microsoft Graph
• GroupMember.Read.All (Aplicativo) — Leia todas as associações do grupo • Notes.Read.All (Aplicativo) — Leia todos os cadernos OneNote • Sites.Read.All (Application): leia itens em todos os conjuntos de sites (Site.Read.All) • User.Read.All (Application): leia o perfil completo de todos os usuários (User.Read.All)

> **Note**
>
> Nenhuma permissão de API é necessária para rastrear entidades usando a autenticação básica e a autenticação somente do SharePoint aplicativo.

No SharePoint Servidor, verifique se você tem:

- Copiou seus URLs de SharePoint instância e o nome de domínio de seus SharePoint URLs. O formato do URL do host que você insere é *`https://yourcompany/sites/mysite`*. O URL deve começar com `https`.

> **Note**
>
> (Local/servidor) Amazon Kendra verifica se as informações do endpoint incluídas são iguais às informações do endpoint especificadas nos AWS Secrets Manager detalhes de configuração da fonte de dados. Isso ajuda a proteger contra o problema de assistência confusa, que é um problema de segurança em que um usuário não tem permissão para realizar uma ação, mas usa o Amazon Kendra como proxy para acessar a senha configurada e realizar a ação. Se você alterar posteriormente as informações do endpoint, crie uma nova senha para sincronizar essas informações.

- Autenticação multifator (MFA) desativada em sua SharePoint conta, para que ela não Amazon Kendra seja impedida de rastrear seu conteúdo. SharePoint
- Se estiver usando a autenticação SharePoint somente de aplicativos para controle de acesso:
  - Copiou o ID SharePoint do cliente gerado quando você registrou o aplicativo somente no nível do site. O formato da ID do cliente é ClientId @TenantId. Por exemplo, *ffa956f3-8f89-44e7-b0e4-49670756342c @888d0b57 -69f1-4fb8-957f-e1f0bedf82fe*.
  - Copiou o segredo SharePoint do cliente gerado quando você registrou o aplicativo somente no nível do site.

  Observação: como os IDs e segredos do cliente são gerados para sites únicos somente quando você registra a autenticação SharePoint Server for App Only, somente um URL do site é suportado para a autenticação SharePoint App Only.

  > **Note**
  >
  > SharePoint A autenticação somente de aplicativos não é compatível com a versão SharePoint 2013.

- Se estiver usando o ID de e-mail com domínio personalizado para controle de acesso:
  - Anote o valor do domínio de e-mail personalizado, por exemplo: *"amazon.com"*.
- Se estiver usando o ID de e-mail com domínio a partir da autorização do IDP, faça uma cópia:
  - Endpoint do servidor LDAP (endpoint do servidor LDAP, incluindo protocolo e número da porta). Por exemplo: *ldap://example.com:389*.

  - Base de pesquisa LDAP (base de pesquisa do usuário LDAP). Por exemplo: *CN=Users,DC=sharepoint,DC=com*.
  - Nome de usuário e senha LDAP.
- Credenciais de autenticação NTLM configuradas ou credenciais de autenticação Kerberos configuradas contendo um nome de usuário (nome de usuário da SharePoint conta) e senha (senha da conta). SharePoint

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de SharePoint autenticação em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de SharePoint dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de SharePoint dados, você deve fornecer detalhes de suas SharePoint credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou SharePoint para Amazon Kendra verPré-requisitos.

Console: SharePoint Online

Para se conectar Amazon Kendra ao SharePoint Online

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha SharePoint conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o SharePoint conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
    a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
    b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
    c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
    d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
    e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
    a. Método de hospedagem — Escolha SharePoint online.

b. URLs do site específicos do seu SharePoint repositório — insira os URLs do SharePoint host. O formato do URL do host que você insere é *https://yourdomain.sharepoint.com/sites/mysite*. O URL deve começar com o protocolo `https`. Separe os URLs com uma nova linha. Você pode adicionar até 100 URLs.

c. Domínio — insira o SharePoint domínio. Por exemplo, o domínio no URL *https://yourdomain.sharepoint.com/sites/mysite* é *yourdomain*.

d. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   Você também pode escolher o tipo de ID do usuário, seja o nome principal do usuário ou o e-mail do usuário obtido no Portal do Azure. Se você não especificar, o e-mail será usado por padrão.

e. Autenticação — escolha entre básica, OAuth 2.0, autenticação somente de aplicativo do Azure AD, autenticação somente de aplicativo ou autenticação de token de SharePoint atualização do OAuth 2.0. Você escolhe um AWS Secrets Manager segredo existente para armazenar suas credenciais de autenticação ou cria um segredo.

   i. Se estiver usando a Autenticação Básica, seu segredo deve incluir um nome secreto, nome de SharePoint usuário e senha.

   ii. Se estiver usando a autenticação OAuth 2.0, seu segredo deverá incluir o ID do SharePoint locatário, o nome do segredo, o nome SharePoint do usuário, a senha, o ID do cliente do Azure AD gerado quando você se registra SharePoint no Azure AD e o segredo do cliente do Azure AD gerado quando você se registra SharePoint no Azure AD.

   iii. Se estiver usando a autenticação somente de aplicativos do Azure AD, seu segredo deverá incluir a ID do SharePoint locatário, o certificado X.509 autoassinado do Azure AD, o nome secreto, a ID do cliente do Azure AD gerada quando você se registra SharePoint no Azure AD e a chave privada para autenticar o conector para o Azure AD.

   iv. Se estiver usando a autenticação SharePoint Somente Aplicativo, seu segredo deverá incluir a ID do SharePoint locatário, o nome do segredo, a ID SharePoint

do cliente que você gerou ao registrar o Aplicativo Somente no Nível do Locatário, o segredo SharePoint do cliente gerado ao se registrar no Aplicativo Somente no Nível do Locatário, o ID do cliente do Azure AD gerado ao se registrar SharePoint no Azure AD e o segredo do cliente do Azure AD gerado ao se registrar no Azure AD. SharePoint

O formato da ID do SharePoint cliente é *TenantIdClientID@*. Por exemplo, *ffa956f3-8f89-44e7-b0e4-49670756342c @888d0b57 -69f1-4fb8-957f-e1f0bedf82fe*.

v. Se estiver usando a autenticação de token de atualização do OAuth 2.0, seu segredo deverá incluir a ID do SharePoint locatário, o nome secreto, a ID exclusiva do cliente do Azure AD gerada quando você se registra SharePoint no Azure AD, o segredo do cliente do Azure AD gerado quando você se SharePoint registra no Azure AD e o token de atualização gerado para se conectar. Amazon Kendra SharePoint

f. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

g. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

Você também pode optar por rastrear o mapeamento de grupos locais ou o mapeamento de grupos do Azure Active Directory.

> **Note**
>
> O rastreamento de mapeamento de grupos do AD está disponível somente para OAuth 2.0, token de atualização do OAuth 2.0 e autenticação somente de aplicativo. SharePoint

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Em Sincronizar escopo, escolha uma das opções a seguir:

      i. Selecionar entidades: escolha as entidades que você deseja rastrear. Você pode optar por rastrear todas as entidades ou qualquer combinação de arquivos, anexos, páginas de links, eventos, comentários e dados da lista.

      ii. Em Configuração adicional, para padrões de Regex de entidades, adicione padrões de expressão regular para links, páginas e eventos para incluir entidades específicas em vez de sincronizar todos os documentos.

      iii. Padrões Regex — Adicione padrões de expressão regular para incluir ou excluir arquivos por caminho do arquivo, nome do arquivo, tipo de arquivo, nome da OneNote seção e nome da OneNote página, em vez de sincronizar todos os seus documentos. Você pode adicionar até 100.

      > **Note**
      >
      > OneNote o rastreamento está disponível somente para OAuth 2.0, token de atualização OAuth 2.0 e autenticação somente de aplicativo. SharePoint

b. No Modo de sincronização, escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Ao sincronizar a fonte de dados do Amazon Kendra pela primeira vez, todo o conteúdo é sincronizado por padrão.

- Sincronização completa: sincronize todo o conteúdo, independentemente do status de sincronização anterior.
- Sincronização de documentos novos ou modificados: sincronize somente documentos novos e modificados.
- Sincronização de documentos novos, modificados ou excluídos: sincronize somente documentos novos, modificados e excluídos.

c. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

d. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

Console: SharePoint Server

Para se conectar Amazon Kendra a SharePoint

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha SharePoint conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o SharePoint conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. Método de hospedagem — Escolha o SharePointservidor.
   b. Escolha a SharePoint versão — escolha SharePoint 2013, SharePoint 2016, SharePoint 2019 e SharePoint (Edição por assinatura).
   c. URLs do site específicos do seu SharePoint repositório — insira os URLs do SharePoint host. O formato do URL do host que você insere é *`https://yourcompany/sites/mysite`*. O URL deve começar com o protocolo `https`. Separe os URLs com uma nova linha. Você pode adicionar até 100 URLs.
   d. Domínio — insira o SharePoint domínio. Por exemplo, o domínio no URL *`https://yourcompany/sites/mysite`* é *`yourcompany`*
   e. Local do certificado SSL — Insira o Amazon S3 caminho para o arquivo do certificado SSL.

f. (Opcional) Proxy Web: insira o nome do host (sem o protocolo `http://` ou `https://`) e o número da porta usada pelo protocolo de transporte de URL do host. O número da porta deve ser um valor numérico entre 0 e 65535.

g. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   Para SharePoint Servidor, você pode escolher entre as seguintes opções de ACL:

   i. ID de e-mail com domínio do IDP — O ID de usuário é baseado em IDs de e-mail com seus domínios obtidos do provedor de identidade subjacente (IDP). Você fornece os detalhes da conexão do IDP em seu Secrets Manager segredo como parte da Autenticação.

   ii. ID de e-mail com domínio personalizado — O ID de usuário é baseado no valor do domínio de e-mail personalizado. Por exemplo, *"amazon.com"*. O domínio de e-mail será usado para criar o ID de e-mail para controle de acesso. Você deve inserir seu domínio de e-mail personalizado.

   iii. Domínio\ Usuário com domínio — O ID do usuário é construído usando o formato Domínio\ ID do usuário. Você precisa fornecer um nome de domínio válido. Por exemplo: *“sharepoint2019"* para construir o controle de acesso.

h. Para Autenticação, escolha SharePoint Autenticação somente de aplicativo, autenticação NTLM ou autenticação Kerberos. Você escolhe um AWS Secrets Manager segredo existente para armazenar suas credenciais de autenticação ou cria um segredo.

   i. Se estiver usando a autenticação NTLM ou a autenticação Kerberos, seu segredo deverá incluir um nome secreto, nome de usuário e senha.

      Se estiver usando o ID de e-mail com o domínio do IDP, insira também:

      - Endpoint do servidor LDAP: endpoint do servidor LDAP, incluindo protocolo e número da porta. Por exemplo: *ldap://example.com:389*.
      - Base de pesquisa LDAP — Base de pesquisa do usuário LDAP. Por exemplo: *CN=Users,DC=sharepoint,DC=com*.
      - Nome de usuário do LDAP: o nome de usuário do LDAP.

- Senha LDAP: a senha LDAP.

ii. Se estiver usando a autenticação SharePoint somente de aplicativo, seu segredo deverá incluir um nome secreto, ID SharePoint do cliente que você gerou ao registrar o aplicativo somente no nível do site, o segredo SharePoint do cliente gerado quando você se registrou no aplicativo somente no nível do site.

O formato da ID do SharePoint cliente é *TenantIdClientID@*. Por exemplo, *ffa956f3-8f89-44e7-b0e4-49670756342c @888d0b57 -69f1-4fb8-957f-e1f0bedf82fe*.

Observação: como os IDs e segredos do cliente são gerados para sites únicos somente quando você registra a autenticação SharePoint Server for App Only, somente um URL do site é suportado para a autenticação SharePoint App Only.

Se estiver usando o ID de e-mail com o domínio do IDP, insira também:

- Endpoint do servidor LDAP: endpoint do servidor LDAP, incluindo protocolo e número da porta. Por exemplo: *ldap://example.com:389*.
- Base de pesquisa LDAP — Base de pesquisa do usuário LDAP. Por exemplo: *CN=Users,DC=sharepoint,DC=com*.
- Nome de usuário do LDAP: o nome de usuário do LDAP.
- Senha LDAP: a senha LDAP.

i. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

j. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

Você também pode optar por rastrear o mapeamento de grupos locais ou o mapeamento de grupos do Azure Active Directory.

> **Note**
>
> O rastreamento de mapeamento de grupos do AD está disponível somente na autenticação SharePoint App Only.

k. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

l. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Em Sincronizar escopo, escolha uma das opções a seguir:

      i. Selecionar entidades: escolha as entidades que você deseja rastrear. Você pode optar por rastrear todas as entidades ou qualquer combinação de arquivos, anexos, páginas de links, eventos e dados da lista.

      ii. Em Configuração adicional, para padrões de Regex de entidades, adicione padrões de expressão regular para links, páginas e eventos para incluir entidades específicas em vez de sincronizar todos os documentos.

      iii. Padrões Regex — Adicione padrões de expressão regular para incluir ou excluir arquivos por caminho do arquivo Nome do arquivo Tipo de arquivo, nome da OneNoteseção e nome da OneNotepágina em vez de sincronizar todos os seus documentos. Você pode adicionar até 100.

> **Note**
>
> OneNote o rastreamento está disponível somente para autenticação somente de SharePoint aplicativos.

b. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

   - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
   - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
   - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

c. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

d. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta

página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a SharePoint

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `SHAREPOINTV2` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Metadados do endpoint do repositório — especifique o fim da `tenantID` `domain` sua instância. `siteUrls` SharePoint
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem

de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

> **Note**
>
> O rastreador de identidade está disponível somente quando você define como`crawlAcl`. `true`

- Propriedades adicionais do repositório, especifique:
  - (Para Azure AD) `s3bucketName` e `s3certificateName` você usa para armazenar seu certificado X.509 autoassinado do Azure AD.
  - Tipo de autenticação (`auth_Type`) que você usa `OAuth2`OAuth2App, se `OAuth2CertificateBasic`,`OAuth2_RefreshToken`,`NTLM`,, `Kerberos` e.
  - Versão (`version`) que você usa, seja `Server` ou`Online`. Se você usar `Server`, poderá especificar ainda mais on`PremVersion` como `2013`, `2016`, `2019` ou `SubscriptionEdition`.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. SharePoint

  Se você usa SharePoint Online, pode escolher entre a autenticação Básica, OAuth 2.0, Somente Aplicativo do Azure AD e SharePoint Somente Aplicativo. Veja a seguir a estrutura JSON mínima que deve estar presente em sua senhao para cada opção de autenticação:
  - Autenticação básica

```
{
    "userName": "SharePoint account user name",
    "password": "SharePoint account password"
}
```

  - Autenticação 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"
}
```

- Autenticação somente para aplicativo do Azure AD

```
{
    "clientId": "client id generated when registering SharePoint with Azure AD",
    "privateKey": "private key to authorize connection with Azure AD"
}
```

- SharePoint Autenticação somente por aplicativo

```
{
    "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"
}
```

- Autenticação de token de atualização 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"
}
```

Se você usa o SharePoint Server, pode escolher entre autenticação SharePoint somente de aplicativo, autenticação NTLM e autenticação Kerberos. Veja a seguir a estrutura JSON mínima que deve estar presente em sua senhao para cada opção de autenticação:

- SharePoint Autenticação somente por aplicativo

```
{
    "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 Autenticação somente de aplicativo com domínio a partir da autorização do 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"
}
```

- (Somente servidor) Autenticação NTLM ou Kerberos

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "userName": "SharePoint account user name",
    "password": "SharePoint account password"
}
```

- (Somente servidor) Autenticação NTLM ou Kerberos com domínio a partir da autorização do 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 role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as

APIs públicas necessárias para o SharePoint conector e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes SharePoint de dados.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão — você pode especificar se deseja incluir ou excluir determinados arquivos e outros conteúdos. OneNotes

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Mapeamentos de campo — Escolha mapear os campos da fonte de SharePoint dados para os Amazon Kendra campos de índice. Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o esquema SharePoint do modelo.

Observações

- O conector oferece suporte a mapeamentos de campo personalizados somente para a entidade Arquivos.
- Para todas as versões SharePoint do servidor, o token ACL deve estar em minúsculas. *Para **eE-mail com domínio do IDP** e **ID de e-mail com ACL de domínio personalizado** como, por exemplo: user@sharepoint2019.com.* Para Domínio\Usuário com ACL de domínio como , por exemplo: *sharepoint2013\user*.
- O conector não suporta o modo de registro de alterações/sincronização de conteúdo novo ou modificado para SharePoint 2013.
- Se o nome de uma entidade tiver um caractere % em seu nome, o conector ignorará esses arquivos devido às limitações da API.
- OneNote só pode ser rastreado pelo conector usando uma ID de locatário e com o OAuth 2.0, o token de atualização do OAuth 2.0 ou a autenticação somente do aplicativo ativada para o Online. SharePoint SharePoint
- O conector rastreia a primeira seção de um OneNote documento usando somente o nome padrão, mesmo que o documento seja renomeado.
- O conector rastreia links na edição SharePoint 2019, SharePoint on-line e por assinatura, somente se as páginas e os arquivos forem selecionados como entidades a serem rastreadas, além dos links.
- O conector rastreia links em SharePoint 2013 e SharePoint 2016 se Links for selecionado como uma entidade a ser rastreada.
- O conector rastreia os anexos e os comentários da lista somente quando os Dados da Lista também são selecionados como uma entidade a ser rastreada.
- O conector rastreia os anexos de eventos somente quando os Eventos também são selecionados como uma entidade a ser rastreada.
- Para a versão SharePoint online, o token ACL estará em minúsculas. Por exemplo, se o nome principal do usuário for *MaryMajor@domain .com* no portal do Azure, o token ACL no SharePoint Conector será *marymajor@domain.com.*
- No Identity Crawler for SharePoint Online and Server, se você quiser rastrear grupos aninhados, precisará ativar o rastreamento local e o rastreamento de grupos do AD.
- Se você estiver usando SharePoint Online e o Nome Principal do Usuário em seu Portal do Azure for uma combinação de maiúsculas e minúsculas, a SharePoint API o converterá internamente

em minúsculas. Por esse motivo, o Amazon Kendra SharePoint conector define a ACL em letras minúsculas.

Microsoft SQL Server

O Microsoft SQL Server é um sistema de gerenciamento de banco de dados relacional (RDBMS) desenvolvido pela Microsoft. Se você for um Microsoft SQL Server usuário, poderá usar Amazon Kendra para indexar sua fonte Microsoft SQL Server de dados. O conector da fonte de Amazon Kendra Microsoft SQL Server dados é compatível com o MS SQL Server 2019.

Você pode se conectar Amazon Kendra à sua fonte de Microsoft SQL Server dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra Microsoft SQL Server dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de Microsoft SQL Server dados, faça essas alterações em suas Microsoft SQL Server AWS contas.

Em Microsoft SQL Server, verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados.
- Verifique se cada documento é exclusivo em Microsoft SQL Server e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de Microsoft SQL Server em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de Microsoft SQL Server

dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de Microsoft SQL Server dados, você deve fornecer detalhes de suas Microsoft SQL Server credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Microsoft SQL Server para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a Microsoft SQL Server

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Microsoft SQL Serverconector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o Microsoft SQL Serverconector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:

a. Em Fonte, insira o seguinte:

b. Host: insira o nome do host do banco de dados.

c. Port: insira a porta do banco de dados.

d. Instância: insira a instância do banco de dados.

e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

f. Em Autenticação: insira as seguintes informações:

- AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de Microsoft SQL Server autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

  A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

     I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- Microsoft SQL Server -' é adicionado automaticamente ao seu nome secreto.

     II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

  B. Escolha Salvar.

g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Em Sincronizar escopo, escolha uma das opções a seguir:

- Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.

  > **Note**
  >
  > Se o nome de uma tabela incluir caracteres especiais (não alfanuméricos) no nome, você deverá usar colchetes ao redor do nome da tabela. Por exemplo, *`selecione * em [my-database-table]`*

- Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
- Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
- Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

- Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
- Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.
- Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
- Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.

- Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta

página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Microsoft SQL Server

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `sqlserver`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.

  > **Note**
  >
  > Se o nome de uma tabela incluir caracteres especiais (não alfanuméricos) no nome, você deverá usar colchetes ao redor do nome da tabela. Por exemplo, *`selecione * em [my-database-table]`*

- Modo de sincronização — Amazon Kendra especifique como atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte

de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. Microsoft SQL Server A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o Microsoft SQL Server conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do Microsoft SQL Server.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Microsoft SQL Server para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

  > **Note**
  >
  > O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo do Microsoft SQL Server.

Observações

- As linhas excluídas do banco de dados não serão rastreadas ao Amazon Kendra verificar o conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.
- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Microsoft Teams

O Microsoft Teams é uma ferramenta de colaboração corporativa para mensagens, reuniões e compartilhamento de arquivos. Se você for um usuário do Microsoft Teams, poderá usar Amazon Kendra para indexar sua fonte de dados do Microsoft Teams.

Você pode se conectar Amazon Kendra à sua fonte de dados do Microsoft Teams usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Microsoft Teams, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de usar Amazon Kendra para indexar sua fonte de dados do Microsoft Teams, faça essas alterações nas suas AWS contas e no Microsoft Teams.

No Microsoft Exchange, verifique se você:

- Criou uma conta do Microsoft Teams no Office 365.
- Anotou o ID de inquilino do Microsoft 365. Encontre o ID de inquilino nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
- Configurei um aplicativo OAuth no portal do Azure e anotei a ID e o segredo do cliente ou as credenciais do cliente. Consulte o tutorial da Microsoft e o exemplo de aplicativo registrado para obter mais informações.

> **Note**
>
> Quando você cria ou registra um aplicativo no portal do Azure, a ID secreta representa o valor secreto real. Você deve anotar ou salvar o valor real do segredo imediatamente ao criar o segredo e o aplicativo. Você pode acessar seu segredo selecionando o nome do seu aplicativo no portal do Azure e navegando até a opção de menu sobre certificados e segredos.
> Você pode acessar sua ID de cliente selecionando o nome do seu aplicativo no portal do Azure e navegando até a página de visão geral. O ID do aplicativo (cliente) é o ID do cliente.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Permissões necessárias adicionadas. Você pode optar por adicionar todas as permissões ou limitar o escopo selecionando menos permissões com base nas entidades que deseja rastrear. A tabela a seguir lista as permissões no nível do aplicativo por entidade correspondente:

Entidade	Permissões necessárias para sincronização de dados	Permissões necessárias para o Identity Sync
Publicação do canal	• ChannelMessage.Ler. Tudo • Group.Read.All • User.Read • User.Read.All	TeamMember.Ler. Tudo
Channel Attachment	• ChannelMessage.Ler. Tudo • Group.Read.All • User.Read • User.Read.All	TeamMember.Ler. Tudo

Entidade	Permissões necessárias para sincronização de dados	Permissões necessárias para o Identity Sync
Channel Wiki	• Group.Read.All • User.Read • User.Read.All	TeamMember.Ler. Tudo
Chat Message	• Chat.Read.All • ChatMessage.Ler. Tudo • ChatMember.Ler. Tudo • User.Read • User.Read.All • Group.Read.All	TeamMember.Ler. Tudo
Meeting Chat	• Chat.Read.All • ChatMessage.Leia • ChatMember.Ler. Tudo • User.Read • User.Read.All • Group.Read.All	TeamMember.Ler. Tudo
Chat Attachment	• Chat.Read.All • ChatMessage.Leia • ChatMember.Ler. Tudo • User.Read • User.Read.All • Group.Read.All	TeamMember.Ler. Tudo

Entidade	Permissões necessárias para sincronização de dados	Permissões necessárias para o Identity Sync
Meeting File	• Chat.Read.All • ChatMessage.Ler. Tudo • ChatMember.Ler. Tudo • User.Read • User.Read.All • Group.Read.All • Files.Read.All	TeamMember.Ler. Tudo
Calendar Meeting	• Chat.Read.All • ChatMessage.Ler. Tudo • ChatMember.Ler. Tudo • User.Read • User.Read.All • Group.Read.All • Files.Read.All	TeamMember.Ler. Tudo
Meeting Notes	• User.Read • User.Read.All • Group.Read.All • Files.Read.All	TeamMember.Ler. Tudo

- Verifique se cada documento é exclusivo no Microsoft Teams e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Microsoft Teams em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Microsoft Teams Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Microsoft Teams, você deve fornecer os detalhes necessários da sua fonte de dados do Microsoft Teams para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Microsoft Teams para Amazon Kendra, consulte Pré-requisitos.

Console

Para se conectar Amazon Kendra ao Microsoft Teams

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector Microsoft Teams e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Microsoft Teams com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. ID do inquilino — insira seu ID de inquilino do Microsoft 365. Encontre o ID de inquilino nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
   b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
   c. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Microsoft Teams. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Microsoft Teams-' é adicionado automaticamente ao seu nome secreto.

B. Para ID do cliente e segredo do cliente — insira as credenciais de autenticação configuradas no Microsoft Teams no portal do Azure.

ii. Salve e adicione seu segredo.

d. Modelo de pagamento: você pode escolher um modelo de licenciamento e pagamento para sua conta do Microsoft Teams. Os modelos de pagamento do modelo A são restritos aos modelos de licenciamento e pagamento que exigem conformidade de segurança. Os modelos de pagamento do modelo A são restritos aos modelos de licenciamento e pagamento que exigem conformidade de segurança.

e. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

f. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

g. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

h. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Sincronizar conteúdo — Selecione os tipos de conteúdo a serem rastreados. Você pode escolher rastrear o conteúdo do chat, das equipes e do calendário.

   b. Configuração adicional — especifique determinadas datas de início e término do calendário, e-mails de usuários, nomes de equipes e nomes de canais, anexos e. OneNotes

   c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

      - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
      - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
      - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

   d. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

   e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Microsoft Teams

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `MSTEAMS` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- ID do inquilino: encontre o ID de inquilino nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Microsoft Teams. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector do Microsoft Teams e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Microsoft Teams.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Tipos de documento/conteúdo — especifique se deseja rastrear mensagens e anexos de bate-papo, publicações e anexos de canais, wikis de canais, conteúdo do calendário, bate-papos de reuniões, arquivos e notas.
- Conteúdo do calendário — especifique uma data e hora de início e término para rastrear o conteúdo do calendário.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados tipos de conteúdo no Microsoft Teams. Você pode incluir ou excluir nomes de equipes, nomes de canais, nomes e tipos de arquivos, e-mails de usuários, OneNote seções e OneNote páginas.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Microsoft Teams para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para obter uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo do Microsoft Teams.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Microsoft Teams, consulte:

- Pesquise de forma inteligente a fonte de dados do Microsoft Teams da sua organização com o Amazon Kendra conector do Microsoft Teams

Microsoft Yammer

O Microsoft Yammer é uma ferramenta de colaboração corporativa para mensagens, reuniões e compartilhamento de arquivos. Se você for um usuário do Microsoft Yammer, poderá usar Amazon Kendra para indexar sua fonte de dados do Microsoft Yammer.

Você pode se conectar Amazon Kendra à sua fonte de dados do Microsoft Yammer usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Microsoft Yammer, consulteSolucionar problemas de origens de dados.

Atributos compatíveis

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de usar Amazon Kendra para indexar sua fonte de dados do Microsoft Yammer, faça essas alterações no Microsoft Yammer e AWS nas contas.

No Microsoft Yammer, verifique se você:

- Criou uma conta administrativa do Microsoft Yammer no Office 365.
- Anotou o nome de usuário e a respectiva senha do Microsoft Yammer.
- Anotou o ID de inquilino do Microsoft 365. Encontre o ID de inquilino nas propriedades do portal do Azure Active Directory ou no aplicativo OAuth.
- Configurei um aplicativo OAuth no portal do Azure e anotei a ID e o segredo do cliente ou as credenciais do cliente. Consulte o tutorial da Microsoft e o exemplo de aplicativo registrado para obter mais informações.

> **Note**
>
> Quando você cria ou registra um aplicativo no portal do Azure, a ID secreta representa o valor secreto real. Você deve anotar ou salvar o valor real do segredo imediatamente ao criar o segredo e o aplicativo. Você pode acessar seu segredo selecionando o nome do seu aplicativo no portal do Azure e navegando até a opção de menu sobre certificados e segredos.
> Você pode acessar sua ID de cliente selecionando o nome do seu aplicativo no portal do Azure e navegando até a página de visão geral. O ID do aplicativo (cliente) é o ID do cliente.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Verifique se cada documento é exclusivo no Microsoft Yammer e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Microsoft Yammer em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Microsoft Yammer a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Microsoft Yammer, você deve fornecer os detalhes necessários da sua fonte de dados do Microsoft Yammer para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Microsoft Yammer para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao Microsoft Yammer

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.

4. Na página Adicionar fonte de dados, escolha Conector Microsoft Yammer e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Microsoft Yammer com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
    a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
    b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
    c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
    d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
    e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
    a. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
    b. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Microsoft Yammer. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.
        i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :
            A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Microsoft Yammer-' é adicionado automaticamente ao seu nome secreto.
            B. Em Nome de usuário, Senha: digite o nome de usuário e senha do Microsoft Yammer.
            C. Para ID do cliente, segredo do cliente — insira as credenciais de autenticação configuradas no Microsoft Yammer no portal do Azure.

ii. Salve e adicione seu segredo.

c. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

d. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

e. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

f. Escolha Próximo.

7. Na página Configurações de sincronização, insira as seguintes informações:

a. Desde a data: especifique a data para começar a rastrear os dados no Microsoft Yammer.

b. Sincronizar conteúdo — Selecione o tipo de conteúdo a ser rastreado. Por exemplo, mensagem pública, mensagens privadas e anexos.

c. Configuração adicional — especifique determinados nomes de comunidades que você deseja rastrear e também use padrões de expressão regular para incluir ou excluir determinados conteúdos.

d. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você

deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

e. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

f. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Microsoft Yammer

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `YAMMER` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Modo de sincronização — Amazon Kendra especifique como atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Microsoft Yammer. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "username": "user name",
    "password": "password",
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM role — Especifique `RoleArn` quando você chama `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Microsoft Yammer e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Microsoft Yammer.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Tipos de documento/conteúdo — especifique se deseja rastrear conteúdo, mensagens e anexos da comunidade e mensagens privadas.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados tipos de conteúdo.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Microsoft Yammer para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você

deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice_`document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo do Microsoft Yammer.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Microsoft Yammer, consulte:

- Anunciando o conector Yammer para Amazon Kendra

MySQL

O MySQL é um sistema de gerenciamento de banco de dados relacional de código aberto. Se você for um MySQL usuário, poderá usar Amazon Kendra para indexar sua fonte MySQL de dados. O conector da fonte de Amazon Kendra MySQL dados é compatível com o MySQL 8.0. 21.

Você pode se conectar Amazon Kendra à sua fonte de MySQL dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra MySQL dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário

- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de MySQL dados, faça essas alterações em suas MySQL AWS contas.

Em MySQL, verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados.
- Verifique se cada documento é exclusivo em MySQL e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de MySQL em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de MySQL dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de MySQL dados, você deve fornecer detalhes de suas MySQL credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou MySQL para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a MySQL

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

   > **Note**
   >
   > Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha MySQLconector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o MySQLconector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. Em Fonte, insira o seguinte:

b. Host: insira o nome do host do banco de dados.

c. Port: insira a porta do banco de dados.

d. Instância: insira a instância do banco de dados.

e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

f. Em Autenticação: insira as seguintes informações:

- AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de MySQL autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

  A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

     I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- MySQL -' é adicionado automaticamente ao seu nome secreto.

     II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

  B. Escolha Salvar.

g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Em Sincronizar escopo, escolha uma das opções a seguir:

   - Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
   - Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
   - Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
   - Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

   b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

   - Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
   - Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
   - Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
   - Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.

- Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
- Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
- Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a MySQL

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `mySql`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte

de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. MySQL A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o MySQL conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do MySQL.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do MySQL para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Observações

- As linhas excluídas do banco de dados não serão rastreadas durante a Amazon Kendra verificação do conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.
- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Oracle Database

O Oracle Database é um sistema de gerenciamento de banco de dados. Se você for um Oracle Database usuário, poderá usar Amazon Kendra para indexar sua fonte Oracle Database de dados. O conector da fonte de Amazon Kendra Oracle Database dados é compatível com o Oracle Database 18c, 19c e 21c.

Você pode se conectar Amazon Kendra à sua fonte de Oracle Database dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra Oracle Database dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de Oracle Database dados, faça essas alterações em suas Oracle Database AWS contas.

Em Oracle Database, verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados.
- Verifique se cada documento é exclusivo em Oracle Database e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de Oracle Database em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de Oracle Database dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de Oracle Database dados, você deve fornecer detalhes de suas Oracle Database credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou Oracle Database para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a Oracle Database

1. Faça login no AWS Management Console e abra o Amazon Kendra console.

2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Oracle Databaseconector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o Oracle Databaseconector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Em Fonte, insira o seguinte:

   b. Host: insira o nome do host do banco de dados.

   c. Port: insira a porta do banco de dados.

   d. Instância: insira a instância do banco de dados.

   e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

   f. Em Autenticação: insira as seguintes informações:

      - AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de Oracle Database

autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- Oracle Database -' é adicionado automaticamente ao seu nome secreto.

II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

B. Escolha Salvar.

g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Em Sincronizar escopo, escolha uma das opções a seguir:

- Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
- Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
- Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

- Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
- Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
- Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.
- Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
- Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
- Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a Oracle Database

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `oracle`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de

dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:

- `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
- `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. Oracle Database A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o Oracle Database conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do Oracle Database.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Oracle Database para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelos do Oracle Database.

Observações

- As linhas excluídas do banco de dados não serão rastreadas durante a Amazon Kendra verificação do conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.

- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

PostgreSQL

O PostgreSQL é um sistema de gerenciamento de banco de dados relacional de código aberto. Se você for um PostgreSQL usuário, poderá usar Amazon Kendra para indexar sua fonte PostgreSQL de dados. O conector da fonte de Amazon Kendra PostgreSQL dados é compatível com o PostgreSQL 9.6.

Você pode se conectar Amazon Kendra à sua fonte de PostgreSQL dados usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de Amazon Kendra PostgreSQL dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Observações

Atributos compatíveis

- Mapeamentos de campos
- Filtragem de contexto do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de PostgreSQL dados, faça essas alterações em suas PostgreSQL AWS contas.

Em PostgreSQL, verifique se você:

- Anotou o nome de usuário e senha do banco de dados

> **Important**
>
> Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.

- Copiou a URL, a porta e a instância do host do banco de dados.
- Verifique se cada documento é exclusivo em PostgreSQL e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de PostgreSQL em um AWS Secrets Manager senha e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de PostgreSQL dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de PostgreSQL dados, você deve fornecer detalhes de suas PostgreSQL credenciais para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou PostgreSQL para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a PostgreSQL

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha PostgreSQLconector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o PostgreSQLconector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. Em Fonte, insira o seguinte:

   b. Host: insira o nome do host do banco de dados.

   c. Port: insira a porta do banco de dados.

   d. Instância: insira a instância do banco de dados.

   e. Ativar localização do certificado SSL — Escolha inserir o Amazon S3 caminho para seu arquivo de certificado SSL.

   f. Em Autenticação: insira as seguintes informações:

      - AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de PostgreSQL autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

        A. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

           I. Senha: um nome para sua senha. O prefixo 'AmazonKendra- PostgreSQL -' é adicionado automaticamente ao seu nome secreto.

           II. Em Nome de usuário do banco de dados e Senha, insira os valores da credencial de autenticação que você copiou do banco de dados.

        B. Escolha Salvar.

   g. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

   h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

   > **Note**
   >
   > IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

   i. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

   a. Em Sincronizar escopo, escolha uma das opções a seguir:

      - Consulta SQL: insira instruções de consulta SQL, como as operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
      - Coluna da chave primária: forneça a chave primária da tabela do banco de dados. Isso identifica uma tabela no banco de dados.
      - Coluna de título: forneça o nome da coluna do título do documento na tabela do banco de dados.
      - Coluna do corpo — Forneça o nome da coluna do corpo do documento na tabela do banco de dados.

   b. Em Configuração adicional: opcional, escolha entre as seguintes opções para sincronizar um conteúdo específico em vez de sincronizar todos os arquivos:

      - Colunas de detecção de alterações — insira os nomes das colunas que Amazon Kendra serão usadas para detectar alterações no conteúdo. Amazon Kendra reindexará o conteúdo quando houver uma alteração em qualquer uma dessas colunas.
      - Coluna de IDs dos usuários: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
      - Coluna de grupos: insira o nome da coluna que contém os IDs de usuário para ter acesso ao conteúdo.
      - Coluna de URLs de origem: insira o nome da coluna que contém os URLs de origem a serem indexados.
      - Coluna de carimbos de data e hora — Insira o nome da coluna que contém carimbos de data e hora. Amazon Kendra usa informações de data e hora para detectar alterações em seu conteúdo e sincronizar somente o conteúdo alterado.
      - Coluna de fusos horários: insira o nome da coluna que contém os fusos horários para o conteúdo a ser rastreado.
      - Formato de carimbos de data/hora: insira o nome da coluna que contém carimbos de data e hora para usar para detectar alterações de conteúdo e sincronizar novamente o conteúdo.

c. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

   - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
   - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
   - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

d. Em Cronograma de execução da sincronização, em Frequência, escolha com que frequência o Amazon Kendra será sincronizado com a fonte de dados.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Selecione entre os campos de fonte de dados padrão gerados — IDs de documentos, títulos de documentos e URLs de origem — que você deseja mapear para indexar Amazon Kendra .

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a PostgreSQL

Você deve especificar o seguinte usando a TemplateConfigurationAPI:

- Fonte de dados — especifique o tipo de fonte de dados como `JDBC` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- Tipo de banco de dados: especifique o tipo de banco de dados como `postgresql`.
- Consulta SQL — especifique instruções de consulta SQL, como operações SELECT e JOIN. As consultas SQL devem ser inferiores a 32 KB. O Amazon Kendra rastreará todo o conteúdo do banco de dados correspondente à sua consulta.
- Modo de sincronização — Amazon Kendra especifique como atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. PostgreSQL A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "user name": "database user name",
    "password": "password"
```

```
}
```

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o PostgreSQL conector e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do PostgreSQL.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir conteúdo específico usando IDs de usuário, grupos, URLs de origem, carimbos de data e hora e fusos horários.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do PostgreSQL para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice_`document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo do (PostgreSQL.

Observações

- As linhas excluídas do banco de dados não serão rastreadas ao Amazon Kendra verificar o conteúdo atualizado.
- O tamanho dos nomes e valores dos campos em uma linha do banco de dados não pode exceder 400 KB.
- Se você tiver uma grande quantidade de dados na fonte de dados do banco de dados e não quiser Amazon Kendra indexar todo o conteúdo do banco de dados após a primeira sincronização, poderá optar por sincronizar somente documentos novos, modificados ou excluídos.
- Como prática recomendada, forneça credenciais de banco Amazon Kendra de dados somente para leitura.
- Como prática recomendada, evite adicionar tabelas com dados confidenciais ou informações pessoais identificáveis (PII).

Quip

O Quip é um software colaborativo de produtividade que oferece recursos de criação de documentos em tempo real. Você pode usar Amazon Kendra para indexar suas pastas, arquivos, comentários de arquivos, salas de bate-papo e anexos do Quip.

Você pode se conectar Amazon Kendra à sua fonte de dados do Quip usando o Amazon Kendra console e a QuipConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Quip, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra O conector de fonte de dados do Quip é compatível com os seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Quip, faça essas alterações no Quip e AWS nas contas.

No Quip, verifique se você:

- Tem uma conta do Quip com permissões administrativas.
- Criou credenciais de autenticação do Quip que incluem um token de acesso pessoal. O token é usado como sua credencial de autenticação armazenada em AWS Secrets Manager segredo. Consulte a documentação do Quip sobre autenticação para obter mais informações.

  > **Note**
  >
  > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Copiou o domínio do site do Quip. Por exemplo, *`https://quip-company.quipdomain.com/browse`* em que *`quipdomain`* é o domínio.
- Verifique se cada documento é exclusivo no Quip e em outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Quip em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e Secrets Manager segredo ao conectar sua fonte de dados do Quip a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Quip, você deve fornecer os detalhes necessários da sua fonte de dados do Quip para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Quip para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao Quip

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector Quip e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Quip com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:
   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.
   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.
   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.
   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.
   e. Escolha Próximo.
6. Na página Definir seção e segurança, insira as informações a seguir:
   a. Nome de domínio do Quip: insira o Quip que você copiou da conta do Quip.
   b. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Quip. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.
      i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :
         A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Quip-' é adicionado automaticamente ao seu nome secreto.
         B. Token do Quip — Insira o acesso pessoal do Quip configurado no Quip.
      ii. Adicione e salve seu segredo.

c. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

d. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

e. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Adicionar IDs de pasta do Quip para rastrear: os IDs da pasta do Quip que você deseja rastrear.

> **Note**
>
> Para rastrear uma pasta raiz, incluindo todas as subpastas e documentos dentro dela, adicione o ID da pasta raiz. Para rastrear subpastas específicas, adicione as IDs de subpastas específicas.

b. Configuração adicional (tipos de conteúdo): insira os tipos de conteúdo que você deseja rastrear.

c. Padrões Regex: os padrões de expressão regular para incluir ou excluir determinados arquivos. Você pode adicionar até 100 padrões.

d. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

e. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Selecione entre os campos da fonte de dados padrão gerados que você deseja mapear para Amazon Kendra indexar.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Quip

Você deve especificar o seguinte usando a QuipConfiguration API:

- Domínio do site do Quip: por exemplo, *https://quip-company.quipdomain.com/browse* em que *quipdomain* é o domínio.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Quip. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "accessToken": "token"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Quip e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Quip.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique `VpcConfiguration` como parte da configuração da fonte de dados. Consulte Configuração do Amazon Kendra para usar uma VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinadas arquivos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer

documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Pastas — Especifique as pastas e subpastas do Quip que você deseja indexar

> **Note**
>
> Para rastrear uma pasta raiz, incluindo todas as subpastas e documentos dentro dela, insira o ID da pasta raiz. Para rastrear subpastas específicas, adicione as IDs de subpastas específicas.

- Anexos, salas de bate-papo, comentários de arquivos — Escolha se deseja incluir o rastreamento de anexos, conteúdo de salas de bate-papo e comentários de arquivos.
- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Quip para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Quip, consulte:

- Pesquise conhecimento em documentos do Quip com a pesquisa inteligente usando o conector do Quip para Amazon Kendra

Salesforce

O Salesforce é uma ferramenta de gerenciamento de relacionamento com o cliente (CRM) para gerenciar equipes de suporte, vendas e marketing. Você pode usar Amazon Kendra para indexar seus objetos padrão do Salesforce e até mesmo objetos personalizados.

Você pode se conectar Amazon Kendra à sua fonte de dados do Salesforce usando o Amazon Kendra console, a TemplateConfigurationAPI ou a SalesforceConfigurationAPI.

Amazon Kendra tem duas versões do conector Salesforce. Os recursos suportados de cada versão incluem:

Conector Salesforce V1.0/API SalesforceConfiguration

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão

Conector Salesforce V2.0/API TemplateConfiguration

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

> **Note**
>
> O suporte para o conector Salesforce V1.0 SalesforceConfiguration /API está programado para terminar em 2023. Recomendamos migrar para ou usar o conector V2.0/API do Salesforce. TemplateConfiguration

Para solucionar problemas do conector da fonte de dados Amazon Kendra do Salesforce, consulte. Solucionar problemas de origens de dados

Tópicos

- Connector V1.0 do Salesforce
- Connector V2.0 do Salesforce

Connector V1.0 do Salesforce

O Salesforce é uma ferramenta de gerenciamento de relacionamento com o cliente (CRM) para gerenciar equipes de suporte, vendas e marketing. Você pode usar Amazon Kendra para indexar seus objetos padrão do Salesforce e até mesmo objetos personalizados.

> **Important**
>
> Amazon Kendra usa a API Salesforce versão 48. A API do Salesforce limita o número de solicitações que podem ser feitas por dia. Se o Salesforce exceder essas solicitações, ele tentará novamente até conseguir continuar.

> **Note**
>
> O suporte para o conector Salesforce V1.0 SalesforceConfiguration /API está programado para terminar em 2023. Recomendamos migrar para ou usar o conector V2.0/API do Salesforce. TemplateConfiguration

Para solucionar problemas do conector da fonte de dados Amazon Kendra do Salesforce, consulte. Solucionar problemas de origens de dados

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão

Atributos compatíveis

Amazon Kendra O conector de fonte de dados do Salesforce oferece suporte aos seguintes recursos:

- Mapeamentos de campos

- Controle de acesso do usuário
- Filtros de inclusão/exclusão

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Salesforce, faça essas alterações em seu Salesforce e em suas contas. AWS

No Salesforce, verifique se você:

- Criou uma conta do Salesforce e anotou o nome de usuário e a senha que você usa para se conectar ao Salesforce.
- Criou uma conta do Salesforce Connected App com o OAuth ativado e copiou a chave do consumidor (ID do cliente) e a senha do consumidor (senha do cliente) atribuídos ao Salesforce Connected App. O ID do cliente e o segredo do cliente são usados como suas credenciais de autenticação armazenadas em um AWS Secrets Manager segredo. Consulte a Documentação do Salesforce sobre aplicativos conectados para obter mais informações.

  > **Note**
  >
  > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Copiou o token de segurança do Salesforce associado à conta usada para se conectar ao Salesforce.
- Copiou o URL da instância do Salesforce que você deseja indexar. Normalmente, ele é *https://.salesforce.com/*<company>. O servidor deve estar executando um aplicativo conectado ao Salesforce.
- Adicionou credenciais ao seu servidor Salesforce para um usuário com acesso somente de leitura ao Salesforce clonando o ReadOnly perfil e adicionando as permissões Exibir todos os dados e Gerenciar artigos. Essas credenciais identificam o usuário que está fazendo a conexão e o aplicativo conectado ao Salesforce ao qual Amazon Kendra se conecta.
- Verifique se cada documento é exclusivo no Salesforce e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve

conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Salesforce em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados do Salesforce a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Salesforce, você deve fornecer os detalhes necessários da sua fonte de dados do Salesforce para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Salesforce para Amazon Kendra ver. Pré-requisitos

Console

Para se conectar Amazon Kendra ao Salesforce

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha o Conector do Salesforce V1.0 e, em seguida, escolha Adicionar conector.
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Nome da fonte de dados: digite um nome para sua fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Idioma padrão: um idioma para filtrar os documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Adicionar nova tag: tags para pesquisar e filtrar os recursos ou monitorar os custos compartilhados.

   e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. URL do Salesforce: insira o URL da instância do site do Salesforce que você deseja indexar.

   b. Em Tipo de autenticação, escolha entre Existente e Novo para armazenar as credenciais de autenticação do Salesforce. Se você optar por criar um novo segredo, uma janela AWS Secrets Manager secreta será aberta.

      - Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

        A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Salesforce-' é adicionado automaticamente ao seu nome secreto.

B. Em Nome de usuário, senha, token de segurança, chave do consumidor, senha do consumidor e URL de autenticação, insira os valores da credencial de autenticação criados na conta do Salesforce.

C. Escolha Salvar autenticação.

c. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

d. Escolha Próximo.

7. Na página Configurações de sincronização, insira as seguintes informações:

a. Para Rastrear anexos: selecione para rastrear todos os objetos, os artigos e os feeds anexados.

b. Para Objetos padrão, Artigos do conhecimento e Feeds do Chatter, selecione as entidades do Salesforce ou tipos de conteúdo que deseja rastrear.

> **Note**
>
> Você deve fornecer informações de configuração para indexar pelo menos um dos objetos padrão, artigos de conhecimento ou feeds do Chatter. Se optar por rastrear Artigos de conhecimento, especifique os tipos de artigos de conhecimento a serem indexados, o nome dos artigos e se deseja indexar os campos padrão de todos os artigos de conhecimento ou somente os campos de um tipo de artigo personalizado. Se optar por indexar artigos personalizados, deverá especificar o nome interno do tipo de artigo. Você pode especificar até 10 tipos de artigos.

c. Frequência — Com que frequência Amazon Kendra será sincronizada com sua fonte de dados.

d. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Para artigos de conhecimento padrão, anexos de objetos padrão e mapeamentos de campo adicionais sugeridos, selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para seu índice.

> **Note**
>
> É necessário um mapeamento de `_document_body`. Você não pode alterar o mapeamento entre o campo `Salesforce ID` e o campo Amazon Kendra `_document_id` .

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Salesforce

Você deve especificar o seguinte na SalesforceConfigurationAPI:

- URL do servidor: o URL da instância do site do Salesforce que você deseja indexar.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Salesforce. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "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 role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Salesforce e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Salesforce.
- Você deve fornecer informações de configuração para indexar pelo menos um dos objetos padrão, artigos de conhecimento ou feeds do Chatter.
  - Objetos padrão: se você optar por rastrear Objetos padrão, deverá especificar o nome do objeto padrão e o nome do campo na tabela de objetos padrão que contém o conteúdo do documento.
  - Artigos de conhecimento: se optar por rastrear Artigos de conhecimento, especifique os tipos de artigos de conhecimento a serem indexados, os estados dos artigos de conhecimento a serem indexados e se deseja indexar os campos padrão de todos os artigos de conhecimento ou somente os campos de um tipo de artigo personalizado.
  - Feeds do Chatter — Se você optar por rastrear os feeds do Chatter, deverá especificar o nome da coluna na tabela do Salesforce FeedItem que contém o conteúdo a ser indexado.

Você também pode adicionar os seguintes recursos opcionais:

- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados anexos de arquivos.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Salesforce para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

Connector V2.0 do Salesforce

O Salesforce é uma ferramenta de gerenciamento de relacionamento com o cliente (CRM) para gerenciar equipes de suporte, vendas e marketing. Você pode usar Amazon Kendra para indexar seus objetos padrão do Salesforce e até mesmo objetos personalizados.

O conector da fonte de dados Amazon Kendra do Salesforce é compatível com as seguintes edições do Salesforce: Developer Edition e Enterprise Edition.

> **Note**
>
> O suporte para o conector Salesforce V1.0 SalesforceConfiguration /API está programado para terminar em 2023. Recomendamos migrar para ou usar o conector V2.0/API do Salesforce. TemplateConfiguration

Para solucionar problemas do conector da fonte de dados Amazon Kendra do Salesforce, consulte. Solucionar problemas de origens de dados

Tópicos

- Atributos compatíveis

- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra O conector de fonte de dados do Salesforce oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Salesforce, faça essas alterações em seu Salesforce e em suas contas. AWS

No Salesforce, verifique se você:

- Criou uma conta administrativa do Salesforce e anotou o nome de usuário e a senha que você usa para se conectar ao Salesforce.
- Copiou o token de segurança do Salesforce associado à conta usada para se conectar ao Salesforce.
- Criou uma conta do Salesforce Connected App com o OAuth ativado e copiou a chave do consumidor (ID do cliente) e a senha do consumidor (senha do cliente) atribuídos ao Salesforce Connected App. O ID do cliente e o segredo do cliente são usados como suas credenciais de autenticação armazenadas em um AWS Secrets Manager segredo. Consulte a Documentação do Salesforce sobre aplicativos conectados para obter mais informações.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não

recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Copiou o URL da instância do Salesforce que você deseja indexar. Normalmente, ele é *https://.salesforce.com/*<company>. O servidor deve estar executando um aplicativo conectado ao Salesforce.
- Adicionou credenciais ao seu servidor Salesforce para um usuário com acesso somente de leitura ao Salesforce clonando o ReadOnly perfil e adicionando as permissões Exibir todos os dados e Gerenciar artigos. Essas credenciais identificam o usuário que está fazendo a conexão e o aplicativo conectado ao Salesforce ao qual Amazon Kendra se conecta.
- Verifique se cada documento é exclusivo no Salesforce e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Salesforce em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados do Salesforce a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Salesforce, você deve fornecer os detalhes necessários da sua fonte de dados do Salesforce para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Salesforce para Amazon Kendra ver. Pré-requisitos

Console

Para se conectar Amazon Kendra ao Salesforce:

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector Salesforce e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Salesforce com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. URL do Salesforce: insira o URL da instância do site do Salesforce que você deseja indexar.

   b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   c. Insira uma senha existente ou, se você criar uma nova senha, uma janela secreta da AWS Secrets Manager será aberta.

      - Autenticação — insira as seguintes informações na janela Criar um AWS Secrets Manager segredo:

        A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Salesforce-' é adicionado automaticamente ao seu nome secreto.

        B. Em Nome de usuário, senha, token de segurança, chave do consumidor, senha do consumidor e URL de autenticação, insira os valores das credenciais de autenticação que você gerou e baixou da conta do Salesforce.

        > **Note**
        >
        > Se você usa o Salesforce Developer Edition, use **`https://login.salesforce.com/services/oauth2/token`** o URL de login do Meu domínio (por exemplo, ***`https://MyCompany.my.salesforce.com`***) como o URL de autenticação. Se você usa o Salesforce Sandbox Edition, use **`https://test.salesforce.com/services/oauth2/token`** o URL de login do Meu domínio (por exemplo, ***`MyDomainName--SandboxName .sandbox.my.salesforce.com) como o URL de autenticação.`***

        C. Escolha Salvar autenticação.

d. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

e. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

f. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

g. Escolha Próximo.

7. Na página Configurações de sincronização, insira as seguintes informações:

a. Para Rastrear anexos: selecione para rastrear todos os objetos do Salesforce.

b. Para Objetos padrão, Objetos padrão com anexo e Objeto padrão sem anexo e Artigos de conhecimento, selecione entidades do Salesforce ou tipos de conteúdo que você deseja rastrear.

c. Você deve fornecer informações de configuração para indexar pelo menos um dos objetos padrão, artigos de conhecimento ou feeds do Chatter. Se você optar por rastrear Artigos de conhecimento, deverá especificar os tipos de artigos de conhecimento a serem indexados. Você pode escolher textos publicados, arquivados, rascunhos e anexos.

Filtro Regex: especifique um padrão regex para incluir itens específicos do catálogo.

8. Para Configuração adicional:

   - Informações sobre ACL: todas as listas de controle de acesso são incluídas por padrão. Desmarcar uma lista de controle de acesso tornará públicos todos os arquivos dessa categoria.
   - Padrões Regex: adicionar padrões de expressão regular para incluir ou excluir determinados arquivos. Você pode adicionar até 100 padrões.

   Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

   - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
   - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
   - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

9. Escolha Próximo.
10. Na página Definir mapeamentos de campo, insira as seguintes informações:

    a. Para artigos de conhecimento padrão, anexos de objetos padrão e mapeamentos de campo adicionais sugeridos, selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para seu índice.

    > **Note**
    >
    > É necessário um mapeamento de `_document_body`. Você não pode alterar o mapeamento entre o campo `Salesforce ID` e o campo Amazon Kendra `_document_id` . Você pode mapear qualquer campo do Salesforce para os

campos de índice reservados/padrão do título ou corpo do documento Amazon Kendra.
Se você mapear qualquer campo do Salesforce para os campos de título e corpo do documento do Amazon Kendra, o Amazon Kendra usará dados dos campos de título e corpo do documento nas respostas de pesquisa.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

11. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Salesforce

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `SALESFORCEV2` quando você usa o esquema TemplateConfigurationJSON. Além disso, especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- URL do host: especifique o URL do host da instância do Salesforce.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo

da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Salesforce. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "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 role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Salesforce e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Salesforce.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão — você pode especificar se deseja incluir ou excluir determinados documentos, contas, campanhas, casos, contatos, leads, oportunidades, soluções, tarefas, grupos, conversas e arquivos de entidades personalizadas.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Salesforce para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

> **Note**
>
> É necessário um mapeamento de `_document_body`. Você não pode alterar o mapeamento entre o campo `Salesforce ID` e o campo Amazon Kendra

`_document_id` . Você pode mapear qualquer campo do Salesforce para os campos de índice reservados/padrão do título ou corpo do documento Amazon Kendra.
Se você mapear qualquer campo do Salesforce para os campos de título e corpo do documento do Amazon Kendra, o Amazon Kendra usará dados dos campos de título e corpo do documento nas respostas de pesquisa.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo do Salesforce.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Salesforce, consulte:

- Anunciando o conector Salesforce atualizado (V2) para Amazon Kendra

ServiceNow

ServiceNow fornece um sistema de gerenciamento de serviços baseado em nuvem para criar e gerenciar fluxos de trabalho em nível organizacional, como serviços de TI, sistemas de emissão de bilhetes e suporte. Você pode usar Amazon Kendra para indexar seus ServiceNow catálogos, artigos de conhecimento, incidentes e seus anexos.

Você pode se conectar Amazon Kendra à sua fonte de ServiceNow dados usando o Amazon Kendra console, a TemplateConfigurationAPI ou a ServiceNowConfigurationAPI.

Amazon Kendra tem duas versões do ServiceNow conector. Os recursos suportados de cada versão incluem:

ServiceNow conector V1.0/API ServiceNowConfiguration

- Mapeamentos de campos
- ServiceNow versões de instância: Londres, outras
- Filtros de inclusão/exclusão

ServiceNow conector V2.0/API TemplateConfiguration

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- ServiceNow versões de instância: Roma, Sandiego, Tóquio, outras
- Nuvem privada virtual (VPC)

> **Note**
>
> O suporte para o ServiceNow conector V1.0 ServiceNowConfiguration /API está programado para terminar em 2023. Recomendamos migrar para ou usar o ServiceNow conector V2.0/ TemplateConfiguration API.

Para solucionar problemas do conector da fonte de Amazon Kendra ServiceNow dados, consulteSolucionar problemas de origens de dados.

Tópicos

- ServiceNow conector V1.0
- ServiceNow conector V2.0
- Especificando documentos para indexar com uma consulta

ServiceNow conector V1.0

ServiceNow fornece um sistema de gerenciamento de serviços baseado em nuvem para criar e gerenciar fluxos de trabalho em nível organizacional, como serviços de TI, sistemas de emissão de bilhetes e suporte. Você pode usar Amazon Kendra para indexar seus ServiceNow catálogos, artigos de conhecimento e seus anexos.

> **Note**
>
> O suporte para o ServiceNow conector V1.0 ServiceNowConfiguration /API está programado para terminar em 2023. Recomendamos migrar para ou usar o ServiceNow conector V2.0/ TemplateConfiguration API.

Para solucionar problemas do conector da fonte de Amazon Kendra ServiceNow dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra ServiceNow o conector de fonte de dados oferece suporte aos seguintes recursos:

- ServiceNow versões de instância: Londres, outras
- Padrões de inclusão/exclusão: catálogos de serviços, artigos de conhecimento e anexos

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de ServiceNow dados, faça essas alterações em suas ServiceNow AWS contas.

Em ServiceNow, verifique se você tem:

- Criei uma conta de ServiceNow administrador e criei uma ServiceNow instância.
- Copiou o host do URL da sua ServiceNow instância. Por exemplo, se o URL da instância for *`https://your-domain.service-now.com`*, o formato do URL do host inserido será *`your-domain.service-now.com`*.
- Anote suas credenciais básicas de autenticação contendo um nome de usuário e uma senha para permitir Amazon Kendra a conexão com sua ServiceNow instância.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Opcional: configurou um token de credencial do OAuth 2.0 que pode identificar Amazon Kendra e gerar um nome de usuário, senha, ID do cliente e segredo do cliente. O nome de usuário e a senha devem fornecer acesso à base de ServiceNow conhecimento e ao catálogo de serviços. Consulte a ServiceNow documentação sobre a autenticação OAuth 2.0 para obter mais informações.
- Adicionou as seguintes permissões:
  - 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
- Verifique se cada documento é exclusivo em ServiceNow e entre outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de ServiceNow autenticação em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de ServiceNow dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de ServiceNow dados, você deve fornecer os detalhes necessários da sua fonte de ServiceNow dados para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou ServiceNow para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a ServiceNow

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha ServiceNowconector V1.0 e, em seguida, escolha Adicionar fonte de dados.
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. ServiceNow host — Insira a URL do ServiceNow host.

b. ServiceNow versão — Selecione sua ServiceNow versão.

c. Escolha entre Autenticação básica e Autenticação Oauth 2.0 com base no seu caso de uso.

d. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de ServiceNow autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

i. Senha: um nome para sua senha. O prefixo 'AmazonKendra- ServiceNow -' é adicionado automaticamente ao seu nome secreto.

ii. Se estiver usando a autenticação básica, insira o nome secreto, o nome de usuário e a senha da sua conta. ServiceNow

Se estiver usando a autenticação OAuth2, insira o nome secreto, nome de usuário, senha, ID do cliente e segredo do cliente que você criou na sua conta. ServiceNow

iii. Selecione Salvar e adicionar senha.

e. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

f. Escolha Próximo.

7. Na página Configurações de sincronização, insira as seguintes informações:

   a. Inclua artigos de conhecimento: escolha indexar artigos de conhecimento.

   b. Tipo de artigos de conhecimento — Escolha entre Incluir somente artigos públicos e Incluir artigos com base na consulta de ServiceNow filtro com base no seu caso de uso. Se você selecionar Incluir artigos com base na consulta de ServiceNow filtro, deverá inserir uma consulta de filtro copiada da sua ServiceNow conta.

   c. Inclua anexos de artigos de conhecimento: escolha indexar anexos de artigos de conhecimento. Você também pode selecionar tipos de arquivo específicos para indexar.

   d. Incluir itens do catálogo: escolha indexar os itens do catálogo.

   e. Incluir anexos de itens de catálogo: escolha indexar anexos de itens de catálogo. Você também pode selecionar tipos de arquivo específicos para indexar.

   f. Frequência — Com que frequência Amazon Kendra será sincronizada com sua fonte de dados.

   g. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Artigos do Knowledge e catálogo de serviços — Selecione entre os campos da fonte de dados padrão Amazon Kendra gerados e os mapeamentos de campo adicionais sugeridos que você deseja mapear para o seu índice.

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a ServiceNow

Você deve especificar o seguinte usando a ServiceNowConfiguration API:

- URL da fonte de dados — especifique o ServiceNow URL. O endpoint do host deve ser semelhante a: *your-domain.service-now.com*.
- Instância do host da fonte de dados — especifique a versão da instância do ServiceNow host como `LONDON` ou`OTHERS`.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. ServiceNow

  Se você usar uma autenticação básica, a senha deverá conter uma estrutura JSON com as seguintes chaves:

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  Se você usar uma autenticação OAuth2, a senha deverá conter uma estrutura JSON com as seguintes chaves:

  ```
  {
      "username": "user name",
      "password": "password",
      "clientId": "client id",
      "clientSecret": "client secret"
  }
  ```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o ServiceNow conector e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes ServiceNow de dados.

Você também pode adicionar os seguintes recursos opcionais:

- Mapeamentos de campo — Escolha mapear os campos da fonte de ServiceNow dados para os Amazon Kendra campos de índice. Para obter mais informações, consulte Mapear campos de fonte de dados.

  > **Note**
  >
  > O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados anexos de arquivos de catálogos e artigos de conhecimento.

  > **Note**
  >
  > A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Parâmetros de indexação: você também pode optar por especificar se deseja:
  - Indexar artigos de conhecimento, catálogos de serviços ou os dois Se você optar por indexar artigos de conhecimento e itens do catálogo de serviços, deverá fornecer o nome do ServiceNow campo mapeado para o campo de conteúdo do documento de Amazon Kendra índice no índice.
  - Indexar anexos a artigos de conhecimento e itens de catálogo.
  - Use uma ServiceNow consulta que selecione documentos de uma ou mais bases de conhecimento. As bases de conhecimento podem ser públicas ou privadas. Para obter mais informações, consulte Como especificar documentos para indexar com uma consulta.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte ServiceNow de dados, consulte:

- Começando com o conector Amazon Kendra ServiceNow online

ServiceNow conector V2.0

ServiceNow fornece um sistema de gerenciamento de serviços baseado em nuvem para criar e gerenciar fluxos de trabalho em nível organizacional, como serviços de TI, sistemas de emissão de bilhetes e suporte. Você pode usar Amazon Kendra para indexar seus ServiceNow catálogos, artigos de conhecimento, incidentes e seus anexos.

Para solucionar problemas do conector da fonte de Amazon Kendra ServiceNow dados, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra ServiceNow o conector de fonte de dados oferece suporte aos seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- ServiceNow versões de instância: Roma, Sandiego, Tóquio, outras
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de ServiceNow dados, faça essas alterações em suas ServiceNow AWS contas.

Em ServiceNow, verifique se você tem:

- Crie uma instância de desenvolvedor pessoal ou empresarial e tenha uma ServiceNow instância com uma função administrativa.
- Copiou o host do URL da sua ServiceNow instância. O formato do URL do host que você insere é *your-domain.service-now.com*. Você precisa do URL da sua ServiceNow instância para se conectar Amazon Kendra.
- Anote suas credenciais básicas de autenticação de um nome de usuário e senha para permitir Amazon Kendra a conexão com sua ServiceNow instância.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Opcional: credenciais de cliente OAuth 2.0 configuradas que podem ser identificadas Amazon Kendra usando um nome de usuário, senha, um ID de cliente gerado e um segredo do cliente. Consulte a ServiceNow documentação sobre a autenticação OAuth 2.0 para obter mais informações.
- Verifique se cada documento é exclusivo em ServiceNow e entre outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de ServiceNow autenticação em um AWS Secrets Manager segredo e, se estiver usando a API, anotou o ARN do segredo.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de ServiceNow dados Amazon Kendra a. Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de ServiceNow dados, você deve fornecer os detalhes necessários da sua fonte de ServiceNow dados para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou ServiceNow para Amazon Kendra verPré-requisitos.

Console

Para se conectar Amazon Kendra a ServiceNow

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

   > **Note**
   >
   > Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha ServiceNow conector e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o ServiceNow conector com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. ServiceNow host — Insira a URL do ServiceNow host. O formato do URL do host que você insere é *your-domain.service-now.com*.

b. ServiceNow versão — Selecione a versão da sua ServiceNow instância. Você pode selecionar entre Roma, Sandiego, Tóquio ou outros.

c. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

d. Autenticação — Escolha entre a autenticação básica e a autenticação Oauth 2.0.

e. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de ServiceNow autenticação. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta. Insira as seguintes informações na janela:

i. Senha: um nome para sua senha. O prefixo 'AmazonKendra- ServiceNow -' é adicionado automaticamente ao seu nome secreto.

ii. Se estiver usando a autenticação básica, insira o nome secreto, o nome de usuário e a senha da sua conta. ServiceNow

Se estiver usando a autenticação OAuth2.0, insira o nome secreto, nome de usuário, senha, ID do cliente e segredo do cliente que você criou na sua conta. ServiceNow

iii. Salve e adicione seu segredo.

f. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

g. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

h. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

i. Escolha Próximo.

7. Na página Configurações de sincronização, insira as seguintes informações:

a. Para artigos de conhecimento, escolha entre as seguintes opções:

- Artigos de conhecimento: escolha indexar artigos de conhecimento.
- Anexos de artigos de conhecimento: escolha indexar anexos de artigos de conhecimento.
- Tipo de artigos de conhecimento — Escolha entre Somente artigos públicos e artigos de conhecimento com base na consulta de ServiceNow filtro com base em seu caso de uso. Se você selecionar Incluir artigos com base na consulta de ServiceNow filtro, deverá inserir uma consulta de filtro copiada da sua ServiceNow conta. *Exemplos de consultas de filtro incluem: workflow_state=draft^EQ,*

kb_knowledge_base=dfc19531bf2021003f07e2c1ac0739ab^text ISNOTEMPTY^EQ, article_type=text^active=true^EQ.

> **Important**
>
> Se você optar por rastrear Somente artigos públicos, Amazon Kendra rastreia somente artigos de conhecimento aos quais tenha sido atribuída uma função de acesso público em. ServiceNow

- Incluir artigos com base em um filtro de descrição curta: especifique padrões de expressão regular para incluir ou excluir artigos específicos.

b. Para Itens do catálogo de serviços:

- Itens do catálogo de serviço: escolha indexar os itens do catálogo de serviço.
- Anexos de itens de catálogo de serviço: escolha indexar anexos de itens de catálogo.
- Itens do catálogo de serviço ativos: escolha indexar os itens do catálogo de serviço ativos.
- Itens do catálogo de serviço inativos: escolha indexar os itens do catálogo de serviço inativos.
- Consulta de filtro — Escolha incluir itens do catálogo de serviços com base em um filtro definido na sua ServiceNow instância. *Exemplos de consultas de filtro incluem: short_descriptionLIKEAccess^category=2809952237b1300054b6a3549dbe5dd4 nameSTARTSWITHService^active=true^EQ*.
- Incluir itens do catálogo de serviços com base no filtro de descrição resumida: especifique um padrão regex para incluir itens específicos do catálogo.

c. Para incidentes:

- Incidentes: escolha indexar incidentes de serviço.
- Anexos de incidentes: opte por indexar anexos de incidentes.
- Incidentes ativos: escolha indexar incidentes ativos.
- Incidentes inativos: escolha indexar incidentes inativos.
- Tipo de incidente ativo: escolha entre Todos os incidentes, Incidentes abertos, Aberto: incidentes não atribuídos e Incidentes resolvidos, dependendo do caso de uso.

- Consulta de filtro — Escolha incluir incidentes com base em um filtro definido na sua ServiceNow instância. *Exemplos de consultas de filtro incluem: short_descriptionLIKETest^urgency=3^state=1^EQ, priority=2^category=software^EQ* .
- Incluir incidentes com base em um filtro de descrição resumida: especifique um padrão de regex para incluir incidentes específicos.

d. Para Configuração adicional:

- Informações de ACL: as listas de controle de acesso para entidades que você selecionou são incluídas por padrão. Desmarcar uma lista de controle de acesso tornará públicos todos os arquivos dessa categoria. As opções de ACL são automaticamente desativadas para entidades não selecionadas. Para artigos públicos, a ACL não é aplicada.
- Para Tamanho máximo do arquivo — Especifique o limite de tamanho do arquivo em MBs que o Amazon Kendra rastreará. O Amazon Kendra rastreará somente os arquivos dentro do limite de tamanho que você definir. O tamanho padrão do arquivo é 50 MB. O tamanho máximo do arquivo deve ser maior que 0MB e menor ou igual a 50MB.
- Padrões de regex de anexos: adicione padrões de expressão regular para incluir ou excluir determinados arquivos anexados de catálogos, artigos de conhecimento e incidentes. Você pode adicionar até 100 padrões.

e. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

- Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
- Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

f. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

g. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

a. Mapeamentos de campo padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra a ServiceNow

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `SERVICENOWV2` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- URL do host — especifique a versão da instância do ServiceNow host. Por exemplo, *`your-domain.service-now.com`*.
- Tipo de autenticação — especifique o tipo de autenticação que você usa, seja `basicAuth` ou `OAuth2` para sua ServiceNow instância.
- ServiceNow versão da instância — Especifique a ServiceNow instância que você usa, se`Tokyo`, `SandiegoRome`, ou`Others`.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial

falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:

- `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
- `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação que você criou em sua conta. ServiceNow

  Se você usar uma autenticação básica, a senha deverá conter uma estrutura JSON com as seguintes chaves:

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

- Se você usar as credenciais do cliente OAuth2, o segredo será armazenado em uma estrutura JSON com as seguintes chaves:

  ```
  {
      "username": "user name",
      "password": "password",
      "clientId": "client id",
      "clientSecret": "client secret"
  }
  ```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o ServiceNow conector e. Amazon Kendra Para obter mais informações, consulte IAM funções para fontes ServiceNow de dados.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Filtros de inclusão e exclusão: especifique se deseja incluir ou excluir determinados arquivos anexados usando os nomes dos arquivos e os tipos de arquivo de artigos de conhecimento, catálogos de serviços e incidentes.

  > **Note**
  >
  > A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Documentos específicos para indexar — Você pode usar uma ServiceNow consulta para especificar os documentos que deseja de uma ou mais bases de conhecimento, incluindo bases de conhecimento privadas. O acesso às bases de conhecimento é determinado pelo usuário que você usa para se conectar à ServiceNow instância. Para obter mais informações, consulte Como especificar documentos para indexar com uma consulta.
- Parâmetros de indexação: você também pode optar por especificar se deseja:
  - Indexar artigos de conhecimento, catálogos de serviços ou os dois Se você optar por indexar artigos de conhecimento, itens do catálogo de serviços e incidentes, deverá fornecer o nome do ServiceNow campo mapeado para o campo de conteúdo do documento de índice no Amazon Kendra índice.
  - Indexar anexos a artigos de conhecimento e itens de catálogo de serviço e incidentes.
  - Inclua artigos de conhecimento, itens do catálogo de serviços e incidentes com base no padrão de filtro `short description`.
  - Escolha filtrar itens e incidentes do catálogo de serviços ativos e inativos.
  - Escolha filtrar incidentes com base no tipo de incidente.
  - Escolha quais entidades devem ter a ACL rastreada.
  - Você pode usar uma ServiceNow consulta para especificar os documentos que deseja de uma ou mais bases de conhecimento, incluindo bases de conhecimento privadas. O acesso às bases de conhecimento é determinado pelo usuário que você usa para se conectar à

ServiceNow instância. Para obter mais informações, consulte Como especificar documentos para indexar com uma consulta.

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Mapeamentos de campo — Escolha mapear os campos da fonte de ServiceNow dados para os Amazon Kendra campos de índice. Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o esquema ServiceNow do modelo.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte ServiceNow de dados, consulte:

- Introdução ao Amazon Kendra anúncio do ServiceNow conector atualizado (V2) para Amazon Kendra

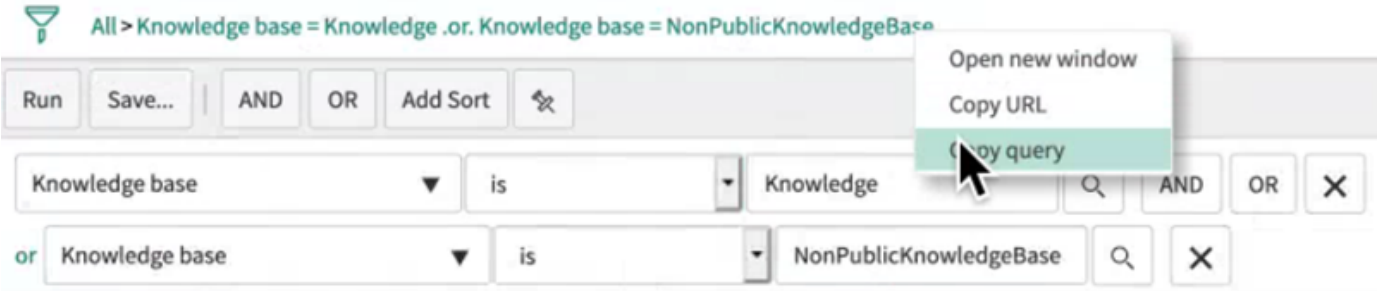
Especificando documentos para indexar com uma consulta

Você pode usar uma ServiceNow consulta para especificar os documentos que deseja incluir em um Amazon Kendra índice. Ao usar uma consulta, você pode especificar várias bases de conhecimento, incluindo bases de conhecimento privadas. O acesso às bases de conhecimento é determinado pelo usuário que você usa para se conectar à ServiceNow instância.

Para criar uma consulta, você usa o construtor de ServiceNow consultas. Você pode usar o construtor para criar a consulta e testar se a consulta retorna a lista correta de documentos.

Para criar uma consulta usando o ServiceNow console

1. Faça login no ServiceNow console.
2. No menu à esquerda, escolha Conhecimento, em seguida, Artigos e Tudo.
3. Na parte superior da página, selecione o ícone de filtro.
4. Use o criador de consultas para criar a consulta.
5. Quando a consulta estiver concluída, clique com o botão direito do mouse na consulta e escolha Copiar consulta para copiar a consulta do criador de consultas. Salve essa consulta para usar em Amazon Kendra.



Certifique-se de não alterar nenhum parâmetro de consulta ao copiar a consulta. Se algum dos parâmetros da consulta não for reconhecido, ServiceNow tratará o parâmetro como vazio e não o usará para filtrar os resultados.

Slack

O Slack é um aplicativo de comunicação empresarial que permite aos usuários enviar mensagens e anexos por meio de vários canais públicos e privados. Você pode usar Amazon Kendra para indexar seus canais públicos e privados do Slack, enviar bots e arquivar mensagens, arquivos e anexos, mensagens diretas e em grupo. Também é possível selecionar conteúdo específico para filtrar.

> **Note**
>
> Amazon Kendra agora suporta um conector Slack atualizado.
> O console foi atualizado automaticamente para você. Todos os novos conectores que você criar no console usarão a arquitetura atualizada. Se você usa a API, agora deve usar o TemplateConfigurationobjeto em vez do `SlackConfiguration` objeto para configurar seu conector.
> Os conectores configurados usando o console antigo e a arquitetura de API continuarão funcionando conforme configurados. No entanto, você não poderá editá-los ou atualizá-los. Se você quiser editar ou atualizar a configuração do conector, deverá criar um novo conector. Recomendamos migrar o fluxo de trabalho do conector para a versão atualizada. O suporte para conectores configurados usando a arquitetura mais antiga está programado para terminar em junho de 2024.

Você pode se conectar Amazon Kendra à sua fonte de dados do Slack usando o Amazon Kendra console ou a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte Amazon Kendra de dados do Slack, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra O conector de fonte de dados do Slack é compatível com os seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Sincronizações de conteúdo completas e incrementais
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Slack, faça essas alterações no Slack e AWS nas contas.

No Slack, verifique se você:

- Configurei um token OAuth do usuário do Slack Bot ou um token OAuth do usuário do Slack. Você pode escolher qualquer um dos tokens para se conectar Amazon Kendra à sua fonte de dados do Slack. É necessário um token para ser usado como suas credenciais de autenticação. Consulte a documentação do Slack sobre tokens de acesso para obter mais informações.

  > **Note**
  >
  > Se você usar o token do bot como parte das credenciais do Slack, não poderá indexar mensagens diretas, mensagens de grupo e deverá adicionar o token do bot ao canal que deseja indexar.

  > **Note**
  >
  > Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- Anotou o ID da equipe do Slack Workspace a partir do URL da página principal do Slack Workspace. Por exemplo, *`https://app.slack.com/client/T0123456789/...`* e que *`T0123456789`* é o ID da equipe.
- Foram adicionados os seguintes escopos/permissões do Oauth:

Escopo do token do usuário	Escopo do token de bot
• channels:history • channels:read • emoji:read • files:read	• channels:history • canais:gerenciar • channels:read • conversations.connect:manage

Escopo do token do usuário	Escopo do token de bot
• groups:history • groups:read • im:history • im:read • mpim:history • mpim:read • team:read • users.profile:read • users:read • usuários:read.email	• conversations.connect:read • files:read • groups:history • groups:read • im:history • im:read • mpim:history • mpim:read • reações:ler • team:read • usergroups:read • users.profile:read • users:read • usuários:read.email

- Verifique se cada documento é exclusivo no Slack e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação do Slack em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um novo Secrets Manager segredo ao conectar sua fonte de dados do Slack a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Slack, você deve fornecer os detalhes necessários da sua fonte de dados do Slack para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Slack para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao Slack

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

   > **Note**
   >
   > Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector Slack e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Slack com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

a. Para o ID da equipe do espaço de trabalho do Slack — o ID da equipe do seu espaço de trabalho do Slack. Você pode encontrar o ID da sua equipe no URL da página principal do seu espaço de trabalho do Slack. Por exemplo, *`https://app.slack.com/client/T0123456789/...`* e que *`T0123456789`* é o ID da equipe.

b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

c. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do Slack. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Slack-' é adicionado automaticamente ao seu nome secreto.

B. Para o token do Slack — insira os valores da credencial de autenticação que você configurou no Slack.

ii. Salve e adicione seu segredo.

d. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

e. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

f. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

g. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Selecione o tipo de conteúdo — Selecione as entidades ou os tipos de conteúdo do Slack que você deseja rastrear. Você pode escolher entre todos os canais, canais públicos, canais privados, mensagens em grupo e mensagens privadas.

b. Selecione a data de início do rastreamento — insira a data em que você deseja começar a rastrear seu conteúdo.

c. Para configuração adicional — escolha incluir mensagens de bots e arquivadas e use padrões de expressão regular para incluir ou excluir determinado conteúdo.

> **Note**
>
> Se você optar por incluir tanto os IDs quanto os nomes dos canais, o conector do Amazon Kendra Slack priorizará os IDs dos canais em vez dos nomes dos canais.
> Se você optar por incluir determinadas mensagens privadas e em grupo, o conector do Amazon Kendra Slack ignorará todas as mensagens privadas e em grupo e rastreará somente as mensagens privadas e de grupo que você especificar.

d. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

   - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
   - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

e. Em Cronograma de execução de sincronização, em Frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

f. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

   b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

   c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Slack

Você deve especificar um JSON do esquema da fonte de dados usando a API TemplateConfiguration. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `SLACK` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- ID da equipe do espaço de trabalho do Slack: o ID da equipe do Slack que você copiou do URL da página principal do Slack.
- Desde a data — a data para começar a rastrear os dados da sua equipe do Slack Workspace. A data deve seguir este formato: yyyy-mm-dd.
- Modo de sincronização — especifique como Amazon Kendra você deve atualizar seu índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização. Escolha uma das seguintes opções:
  - `FORCED_FULL_CRAWL`para indexar todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados é sincronizada com seu índice.
  - `FULL_CRAWL`para indexar somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
  - `CHANGE_LOG`para indexar somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

- Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.
- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Slack. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
    "slackToken": "token"
}
```

- IAM role — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e chamar as APIs públicas necessárias para o conector Slack e. Amazon Kendra Para obter mais informações, consulte Funções para o IAM das fontes de dados do Slack.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Canais específicos — filtre por canais públicos ou privados e especifique determinados canais pelo ID.
- Tipos de canais e mensagens — Se você Amazon Kendra deve indexar seus canais públicos e privados, suas mensagens diretas e de grupo e suas mensagens de bot e arquivadas. Se você usar o token do bot como parte das credenciais do Slack, adicione o token do bot ao canal que deseja indexar. Você não pode indexar mensagens diretas e mensagens de grupo usando um token de bot.

- Retrospectiva: você pode escolher configurar um `lookBack` parâmetro para que o conector do Slack rastreie o conteúdo atualizado ou excluído até um número específico de horas antes da última sincronização do conector.
- Filtros de inclusão e exclusão — especifique se deseja incluir ou excluir determinado conteúdo do Slack. Se você usar o token do bot como parte das credenciais do Slack, adicione o token do bot ao canal que deseja indexar. Você não pode indexar mensagens diretas e mensagens de grupo usando um token de bot.

> **Note**
>
> A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Slack para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice`_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte o Esquema do modelo do Slack.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Slack, consulte:

- Descubra o conhecimento nos espaços de trabalho do Slack com a pesquisa inteligente usando o conector do Slack Amazon Kendra

Zendesk

O Zendesk é um sistema de gerenciamento de relacionamento com o cliente que ajuda as empresas a automatizar e aprimorar as interações de suporte ao cliente. Você pode usar Amazon Kendra para indexar seus tickets de suporte do Zendesk, comentários do ticket, anexos do ticket, artigos da central de ajuda, comentários de artigos, anexos de comentários de artigos, tópicos da comunidade do Guide, publicações da comunidade e comentários de publicações da comunidade.

Você pode filtrar pelo nome da organização se quiser indexar tickets que estejam somente dentro de uma organização específica. Você também pode escolher definir uma data de rastreamento para quando quiser começar a rastrear dados do Zendesk.

Você pode se conectar Amazon Kendra à sua fonte de dados do Zendesk usando o Amazon Kendra console e a TemplateConfigurationAPI.

Para solucionar problemas do conector da fonte de dados do Amazon Kendra Zendesk, consulteSolucionar problemas de origens de dados.

Tópicos

- Atributos compatíveis
- Pré-requisitos
- Instruções de conexão
- Saiba mais

Atributos compatíveis

Amazon Kendra O conector de fonte de dados do Zendesk é compatível com os seguintes recursos:

- Mapeamentos de campos
- Controle de acesso do usuário
- Filtros de inclusão/exclusão
- Registro de alterações, sincronizações completas e incrementais de conteúdo
- Nuvem privada virtual (VPC)

Pré-requisitos

Antes de poder usar Amazon Kendra para indexar sua fonte de dados do Zendesk, faça essas alterações em seu Zendesk e AWS em suas contas.

No , verifique se você:

- Criou uma conta administrativa da Zendesk Suite (Profissional/Empresarial).
- Anotou o URL de host do Zendesk. Por exemplo, *https://{sub-domain}.zendesk.com/*.

> **Note**
>
> (Local/servidor) Amazon Kendra verifica se as informações do endpoint incluídas são iguais às informações do endpoint especificadas nos AWS Secrets Manager detalhes de configuração da fonte de dados. Isso ajuda a proteger contra o problema de assistência confusa, que é um problema de segurança em que um usuário não tem permissão para realizar uma ação, mas usa o Amazon Kendra como proxy para acessar a senha configurada e realizar a ação. Se você alterar posteriormente as informações do endpoint, crie uma nova senha para sincronizar essas informações.

- Configurou um token OAuth 2.0 contendo um ID do cliente, segredo do cliente, nome de usuário e senha. O token OAuth 2.0 é necessário para ser usado como suas credenciais de autenticação. Consulte a documentação do Zendesk sobre a configuração de tokens OAuth 2.0 para obter mais informações.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

- O seguinte escopo OAuth 2.0 foi adicionado:
  - leitura
- Opcional: instalou um certificado SSL para permitir a conexão ao Amazon Kendra .
- Verificou se cada documento é exclusivo no Zendesk e outras fontes de dados que você planeja usar para o mesmo índice. Cada fonte de dados que você deseja usar para um índice não deve

conter o mesmo documento em todas as fontes de dados. Os IDs de documentos são globais para um índice e devem ser exclusivos por índice.

No seu Conta da AWS, verifique se você tem:

- Criou um Amazon Kendra índice e, se estiver usando a API, anotei o ID do índice.
- Criou uma IAM função para sua fonte de dados e, se estiver usando a API, anotei o ARN da IAM função.

> **Note**
>
> Se você alterar o tipo de autenticação e as credenciais, deverá atualizar sua IAM função para acessar a ID AWS Secrets Manager secreta correta.

- Armazenou suas credenciais de autenticação de em uma senha do AWS Secrets Manager e, se estiver usando a API, anotou o ARN da senha.

> **Note**
>
> Recomendamos que você atualize ou alterne regularmente as credenciais e as senhas. Forneça somente o nível de acesso necessário para sua própria segurança. Não recomendamos que você reutilize credenciais e senhas nas fontes de dados e nas versões 1.0 e 2.0 do conector (quando for aplicável).

Se você não tiver uma IAM função ou segredo existente, poderá usar o console para criar uma nova IAM função e um Secrets Manager segredo ao conectar sua fonte de dados do Zendesk a. Amazon Kendra Se você estiver usando a API, deverá fornecer o ARN de uma IAM função e Secrets Manager segredo existentes e um ID de índice.

Instruções de conexão

Para se conectar Amazon Kendra à sua fonte de dados do Zendesk, você deve fornecer os detalhes necessários da sua fonte de dados do Zendesk para que Amazon Kendra possa acessar seus dados. Se você ainda não configurou o Zendesk para Amazon Kendra, consultePré-requisitos.

Console

Para se conectar Amazon Kendra ao Zendesk

1. Faça login no AWS Management Console e abra o Amazon Kendra console.
2. No painel de navegação esquerdo, escolha Índices e, em seguida, escolha o índice que deseja usar na lista de índices.

> **Note**
>
> Você pode escolher definir ou editar as configurações de Controle de acesso do usuário em Configurações do índice.

3. Na página Introdução, escolha Adicionar fonte de dados.
4. Na página Adicionar fonte de dados, escolha Conector Zendesk e, em seguida, escolha Adicionar conector. Se estiver usando a versão 2 (se aplicável), escolha o conector Zendesk com a tag “V2.0".
5. Na página Especar detalhes da fonte de dados, insira as seguintes informações:

   a. Em Nome e descrição, em Nome da fonte de dados: insira um nome para a fonte de dados. Você pode incluir hífens, mas não espaços.

   b. (Opcional) Descrição: insira uma descrição opcional para a fonte de dados.

   c. Em Idioma padrão — Escolha um idioma para filtrar seus documentos para o índice. A menos que você especifique o contrário, o idioma padrão é o inglês. O idioma especificado nos metadados do documento substitui o idioma selecionado.

   d. Em Tags, em Adicionar nova tag — Inclua tags opcionais para pesquisar e filtrar seus recursos ou monitorar seus AWS custos.

   e. Escolha Próximo.

6. Na página Definir seção e segurança, insira as informações a seguir:

   a. URL do Zendesk: insira o URL do Zendesk. Por exemplo, *https://{sub-domain}.zendesk.com/*.

   b. Autorização — Ative ou desative as informações da lista de controle de acesso (ACL) para seus documentos, se você tiver uma ACL e quiser usá-la para controle de acesso. A ACL especifica quais documentos os usuários e grupos podem acessar. As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

   c. AWS Secrets Manager segredo — Escolha um segredo existente ou crie um novo Secrets Manager segredo para armazenar suas credenciais de autenticação do

Zendesk. Se optar por criar uma nova senha, uma janela secreta do AWS Secrets Manager será aberta.

i. Insira as seguintes informações em Criar uma janela de senha do AWS Secrets Manager :

A. Senha: um nome para sua senha. O prefixo 'AmazonKendra-Zendesk' é adicionado automaticamente ao seu nome secreto.

B. Em ID do cliente, segredo do cliente, nome de usuário, senha — insira os valores da credencial de autenticação configurados no Zendesk.

ii. Salve e adicione seu segredo.

d. Nuvem privada virtual (VPC): você pode escolher usar uma VPC. Nesse caso, você deve adicionar sub-redes e grupos de segurança da VPC.

e. Rastreador de identidade — especifique se deseja ativar o rastreador Amazon Kendra de identidade. O rastreador de identidade usa as informações da lista de controle de acesso (ACL) dos seus documentos para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Se você tiver uma ACL para seus documentos e optar por usá-la, também poderá optar por ativar o rastreador de identidade para configurar a filtragem de contexto Amazon Kendra do usuário dos resultados da pesquisa. Caso contrário, se o rastreador de identidade estiver desativado, todos os documentos poderão ser pesquisados publicamente. Se você quiser usar o controle de acesso para seus documentos e o rastreador de identidade estiver desativado, você também pode usar a PutPrincipalMappingAPI para carregar informações de acesso de usuários e grupos para filtragem de contexto do usuário.

f. IAM função — Escolha uma IAM função existente ou crie uma nova IAM função para acessar as credenciais do repositório e indexar o conteúdo.

> **Note**
>
> IAM as funções usadas para índices não podem ser usadas para fontes de dados. Se você não tiver certeza se uma função existente é usada para um índice ou perguntas frequentes, escolha Criar uma nova função para evitar erros.

g. Escolha Próximo.

7. Na página Configurar configurações de sincronização, insira as seguintes informações:

a. Selecionar conteúdo — Selecione os tipos de conteúdo que você deseja rastrear desde tickets até artigos da central de ajuda, tópicos da comunidade e muito mais.

b. Nome da organização — Insira os nomes da organização da Zendesk para filtrar o conteúdo.

c. Data de início da sincronização — insira a data a partir da qual você deseja começar a rastrear seu conteúdo.

d. Padrões Regex: adicionar padrões de expressão regular para incluir ou excluir determinados arquivos. Você pode adicionar até 100 padrões.

e. Modo de sincronização: escolha como você deseja atualizar o índice quando o conteúdo da fonte de dados for alterado. Quando você sincroniza sua fonte de dados Amazon Kendra pela primeira vez, todo o conteúdo é rastreado e indexado por padrão. Você deve executar uma sincronização completa dos seus dados se a sincronização inicial falhar, mesmo que você não escolha a sincronização completa como opção de modo de sincronização.

   - Sincronização completa: indexe todo o conteúdo de forma atualizada, substituindo o conteúdo existente sempre que sua fonte de dados for sincronizada com seu índice.
   - Sincronização nova e modificada: indexe somente conteúdo novo e modificado sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.
   - Sincronização nova, modificada e excluída: indexe somente conteúdo novo, modificado e excluído sempre que sua fonte de dados for sincronizada com seu índice. Amazon Kendra pode usar o mecanismo da sua fonte de dados para rastrear alterações no conteúdo e indexar o conteúdo que foi alterado desde a última sincronização.

f. Em Sincronização, cronograma de execução para frequência — Escolha com que frequência sincronizar o conteúdo da fonte de dados e atualizar seu índice.

g. Escolha Próximo.

8. Na página Definir mapeamentos de campo, insira as seguintes informações:

   a. Campos de fonte de dados padrão — Selecione entre os campos de fonte de dados padrão Amazon Kendra gerados que você deseja mapear para o seu índice.

b. Adicionar campo: para adicionar campos de fonte de dados personalizados, crie um nome de campo de índice para mapear e o tipo de dados do campo.

c. Escolha Próximo.

9. Na página Revisar e criar, verifique se as informações inseridas estão corretas e selecione Adicionar fonte de dados. Você também pode optar por editar as informações a partir desta página. Sua fonte de dados aparecerá na página Fontes de dados depois que a fonte de dados for adicionada com sucesso.

API

Para se conectar Amazon Kendra ao Zendesk

Você deve especificar um JSON do esquema da fonte de dados usando a TemplateConfigurationAPI. Você deve fornecer as seguintes informações:

- Fonte de dados — especifique o tipo de fonte de dados como `ZENDESK` quando você usa o esquema TemplateConfigurationJSON. Também especifique a fonte de dados como `TEMPLATE` quando você chama a CreateDataSourceAPI.
- URL do host: forneça o URL do host como parte da configuração da conexão ou dos detalhes do endpoint do repositório. Por exemplo, *`https://yoursubdomain.zendesk.com`*.
- Registro de alterações — Se Amazon Kendra deve usar o mecanismo de registro de alterações da fonte de dados do Zendesk para determinar se um documento deve ser atualizado no índice.

> **Note**
>
> Use o log de alterações se o Amazon Kendra não quiser digitalizar todos os documentos. Se seu registro de alterações for grande, talvez leve Amazon Kendra menos tempo para digitalizar os documentos na fonte de dados do Zendesk do que para processar o registro de alterações. Se estiver sincronizando a fonte de dados do Zendesk com o índice pela primeira vez, todos os documentos serão digitalizados.

- Nome de recurso secreto da Amazon (ARN) — Forneça o nome de recurso da Amazon (ARN) de um Secrets Manager segredo que contenha as credenciais de autenticação da sua conta do Zendesk. A senha deve conter uma estrutura JSON com as seguintes chaves:

```
{
```

```
    "hostUrl": "https://yoursubdomain.zendesk.com",
    "clientId": "client ID",
    "clientSecret": "Zendesk client secret",
    "userName": "Zendesk user name",
    "password": "Zendesk password"
}
```

- IAM função — Especifique `RoleArn` quando você liga `CreateDataSource` para fornecer uma IAM função com permissões para acessar seu Secrets Manager segredo e para chamar as APIs públicas necessárias para o conector do Zendesk e. Amazon Kendra Para obter mais informações, consulte Funções do IAM para as fontes de dados do Zendesk.

Você também pode adicionar os seguintes recursos opcionais:

- Nuvem privada virtual (VPC): especifique a `VpcConfiguration` quando ao chamar `CreateDataSource`. Para ter mais informações, consulte Configurando Amazon Kendra para usar um Amazon VPC.
- Tipos de documento/conteúdo — especifique se deseja rastrear:
  - Tíquetes de suporte, comentários de tíquetes e/ou anexos de comentários de tickets
  - Artigos da central de ajuda, anexos de artigos e comentários de artigos
  - Tópicos, publicações ou comentários da comunidade do guia
- Filtros de inclusão e exclusão — especifique se deseja incluir ou excluir determinado conteúdo do Slack. Se você usar o token do bot como parte das credenciais do Slack, adicione o token do bot ao canal que deseja indexar. Você não pode indexar mensagens diretas e mensagens de grupo usando um token de bot.

  > **Note**
  >
  > A maioria das fontes de dados usa padrões de expressão regular, que são padrões de inclusão ou exclusão chamados de filtros. Se você especificar um filtro de inclusão, somente o conteúdo que corresponda ao filtro de inclusão será indexado. Qualquer documento que não corresponda ao filtro de inclusão não é indexado. Se especificar um filtro de inclusão e exclusão, os documentos que corresponderem ao filtro de exclusão não serão indexados, mesmo que correspondam ao filtro de inclusão.

- Filtragem de contexto do usuário e controle de acesso —Amazon Kendra rastreia a lista de controle de acesso (ACL) de seus documentos, se você tiver uma ACL para seus documentos.

As informações da ACL são usadas para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Para obter mais informações, consulte Filtrar o contexto do usuário.

- Mapeamentos de campo: escolha mapear os campos de fonte de dados do Zendesk para os campos de índice do Amazon Kendra . Para obter mais informações, consulte Mapear campos de fonte de dados.

> **Note**
>
> O campo do corpo do documento ou o corpo do documento equivalente para seus documentos é obrigatório Amazon Kendra para pesquisar seus documentos. Você deve mapear o nome do campo do corpo do documento na fonte de dados para o nome do campo de índice `_document_body`. Todos os demais campos são opcionais.

Para ver uma lista de outras chaves JSON importantes a serem configuradas, consulte Esquema de modelo do Zendesk.

Saiba mais

Para saber mais sobre a integração Amazon Kendra com sua fonte de dados do Zendesk, consulte:

- Descubra os insights do Zendesk com a pesquisa Amazon Kendra inteligente

Mapeando campos de fontes de dados

Amazon Kendra conectores de fonte de dados podem mapear campos de documentos ou de conteúdo da sua fonte de dados para campos no seu Amazon Kendra índice. Por padrão, todo conector é projetado para rastrear campos específicos da fonte de dados. Os campos padrão da fonte de dados e suas propriedades não podem ser alterados nem personalizados. No Amazon Kendra console, os campos padrão e as propriedades do campo padrão que não podem ser editados ficam esmaecidos.

Amazon Kendra os conectores também permitem que você mapeie campos personalizados de documentos ou conteúdos de sua fonte de dados para campos personalizados em seu índice. Por exemplo, se você tiver um campo na fonte de dados chamado “departamento” que contém

informações do departamento de um documento, ele pode ser mapeado para um campo de índice chamado “Departamento”. Dessa forma, você pode usar o campo ao consultar documentos.

Você também pode mapear campos Amazon Kendra reservados ou comuns, como`_created_at`. Se sua fonte de dados tiver um campo chamado “creation_date”, você poderá mapeá-lo para o campo Amazon Kendra reservado equivalente chamado. `_created_at` Para obter mais informações sobre campos Amazon Kendra reservados, consulte Atributos ou campos do documento.

Você pode mapear campos para a maioria das fontes de dados. Você pode criar mapeamentos de campo para as seguintes fontes de dados:

- Adobe Experience Manager
- Alfresco
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (Windows)
- Amazon FSx (EM UM NetApp TOQUE)
- Amazon RDS/Aurora
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon Kendra Rastreador da Web
- Amazon WorkDocs
- Box
- Confluence
- Dropbox
- Drupal
- GitHub
- Google Workspace Drives
- 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

Ao armazenar os documentos em um bucket do S3 ou fonte de dados do S3, especifique os campos usando um arquivo de metadados JSON. Para obter mais informações, consulte conectores de fonte de dados do S3.

O mapeamento dos campos da fonte de dados para um campo de índice é um processo de três etapas:

1. Crie um índice. Para obter mais informações, consulte Criar um índice.
2. Atualize o índice para adicionar campos.
3. Crie uma fonte de dados e inclua mapeamentos de campo para mapear campos reservados e quaisquer campos personalizados para Amazon Kendra indexar campos.

Para atualizar o índice para adicionar campos personalizados, use o console para editar os mapeamentos dos campos da fonte de dados e adicionar um campo personalizado ou usar a UpdateIndexAPI. Você pode adicionar um total de 500 campos personalizados ao seu índice.

Para fontes de dados do banco de dados, se o nome da coluna do banco de dados corresponder ao nome de um campo reservado, o campo e a coluna serão mapeados automaticamente.

Com a UpdateIndexAPI, você adiciona campos reservados e personalizados usando`DocumentMetadataConfigurationUpdates`.

O exemplo de JSON a seguir usa `DocumentMetadataConfigurationUpdates` para adicionar um campo chamado “Departamento” ao índice.

```
"DocumentmetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE"
   }
]
```

Ao criar o campo, você tem a opção de definir como o campo é usado para pesquisa. Você pode escolher entre as seguintes opções:

- Exibível: determina se o campo é retornado na resposta da consulta. O padrão é `true`.
- Facetável: indica que o campo pode ser usado para criar facetas. O padrão é `false`.
- Pesquisável: determina se o campo é usado na pesquisa. O padrão é `true` para campos de string e `false` para campos de número e data.
- Classificável: indica que o campo pode ser usado para classificar os resultados da pesquisa. Ele só pode ser definido para campos de data, número e sequência de caracteres. Ele não pode ser definido para campos de lista de strings.

O exemplo de JSON a seguir usa `DocumentMetadataConfigurationUpdates` para adicionar um campo chamado “Departamento” ao índice.

```
"DocumentMetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE",
       "Search": {
           "Facetable": true
       }
   }
]
```

Usando campos de documentos Amazon Kendra reservados ou comuns

Com a UpdateIndex API, você pode criar campos reservados ou comuns usando `DocumentMetadataConfigurationUpdates` e especificando o nome do campo de índice Amazon Kendra reservado para mapear para seu atributo de documento/nome de campo equivalente. Você também pode criar campos personalizadas. Se você usa um conector de fonte de dados, a maioria inclui mapeamentos de campo que mapeiam os campos do documento da fonte de dados para campos de Amazon Kendra índice. Se usar o console, atualize os campos selecionando a fonte de dados, a ação de edição e, em seguida, prosseguindo para a seção de mapeamentos de campo para configurar a fonte de dados.

Você pode configurar o objeto `Search` para definir um campo como exibível, facetável, pesquisável e classificável. Configure o objeto `Relevance` para definir a ordem de classificação, a duração do aumento ou o período de tempo de um campo a ser aplicado ao aumento, à atualização, ao valor de importância e aos valores de importância mapeados para valores de campo específicos. Se usar o console, defina as configurações de pesquisa de um campo selecionando a opção de faceta no menu de navegação. Para definir o ajuste de relevância, selecione a opção de pesquisar o índice no menu de navegação, insira uma consulta e use as opções do painel lateral para ajustar a relevância da pesquisa. Você não pode alterar o tipo de campo depois de criar o campo.

Amazon Kendra tem os seguintes campos de documento reservados ou comuns que você pode usar:

- `_authors`: uma lista de um ou mais autores responsáveis pelo conteúdo do documento.
- `_category`: uma categoria que coloca um documento em um grupo específico.
- `_created_at`: a data e a hora no formato ISO 8601 em que o documento foi criado. Por exemplo, 2012-03-25T12:30:10+01:00 é o formato de data e hora ISO 8601 para 25 de março de 2012 às 12h30 (mais 10 segundos) no horário da Europa Central.
- `_data_source_id`: o identificador da fonte de dados que contém o documento.
- `_document_body`: o conteúdo do documento de trabalho.
- `_document_id`: o identificador exclusivo de cada documento.
- `_document_title`: o título do documento.
- `_excerpt_page_number`: o número da página em um arquivo PDF em que o trecho do documento aparece. Se o índice foi criado antes de 8 de setembro de 2020, você deve reindexar os documentos antes de poder usar esse atributo.

- `_faq_id`: se for um documento do tipo pergunta e resposta (Perguntas frequentes), um identificador exclusivo para as Perguntas frequentes.
- `_file_type`: o tipo de arquivo do documento, como pdf ou doc.
- `_last_updated_at`: a data e a hora no formato ISO 8601 em que o documento foi atualizado pela última vez. Por exemplo, 2012-03-25T12:30:10+01:00 é o formato de data e hora ISO 8601 para 25 de março de 2012 às 12h30 (mais 10 segundos) no horário da Europa Central.
- `_source_uri`: o URI em que o documento está disponível. Por exemplo, o URI do documento no site da empresa.
- `_version`: um identificador para a versão específica de um documento.
- `_view_count`: o número de vezes que o documento foi visualizado.
- `_language_code`(String): o código de um idioma que se aplica ao documento. O padrão é inglês se você não especificar um idioma. Para obter mais informações sobre os idiomas suportados, incluindo os códigos, consulte Adicionar documentos em outros idiomas além do inglês.

Para campos personalizados, você cria esses campos usando `DocumentMetadataConfigurationUpdates` com a API `UpdateIndex`, assim como faz ao criar um campo reservado ou comum. Você deve definir o tipo de dados apropriado para o campo personalizado. Se usar o console, atualize os campos selecionando a fonte de dados, a ação de edição e, em seguida, prosseguindo para a seção de mapeamentos de campo para configurar a fonte de dados. Algumas fontes de dados não oferecem suporte à adição de novos campos ou campos personalizados. Você não pode alterar o tipo de campo depois de criar o campo.

Estes são os tipos que podem ser definidos em campos personalizados:

- Data
- Número
- String
- Lista de strings

Se você adicionou documentos ao índice usando a BatchPutDocumentAPI, `Attributes` lista os campos/atributos dos seus documentos e cria campos usando o objeto. `DocumentAttribute`

Para documentos indexados de uma fonte de Amazon S3 dados, você cria campos usando um arquivo de metadados JSON que inclui as informações dos campos.

Ao usar um banco de dados compatível como fonte de dados, poderá configura os campos usando a opção de mapeamentos de campo.

Adicionar documentos em outros idiomas além do inglês

Você pode indexar documentos em vários idiomas. Se não especificar um idioma, o Amazon Kendra indexa documentos em uma fonte de dados em inglês por padrão. Inclua o código do idioma de um documento nos metadados do documento como um campo. Consulte Mapeamentos de campo e Atributos personalizados para obter mais informações sobre o campo `_language_code` de um documento.

Você pode especificar o código do idioma para todos os seus documentos em sua fonte de dados ao ligar CreateDataSource. Se um documento não tiver um código de idioma especificado em um campo de metadados, o documento será indexado usando o código de idioma especificado para todos os documentos no nível da fonte de dados. No console, indexe os documentos em um idioma compatível somente no nível da fonte de dados. Acesse as Fontes de dados, em seguida, a página Especificar detalhes da fonte de dados e escolha um idioma no menu suspenso Idioma.

Você também pode pesquisar ou consultar documentos em um idioma compatível. Para obter mais informações, consulte Pesquisar em idiomas.

Os seguintes idiomas e os códigos são suportados (inglês ou `en` por padrão, se você não especificar um idioma). Essa tabela inclui idiomas que oferecem Amazon Kendra suporte à pesquisa semântica completa, bem como idiomas que oferecem suporte apenas à correspondência simples de palavras-chave. Os idiomas que oferecem suporte à pesquisa semântica completa são marcados com um asterisco e estão em negrito na tabela a seguir. O inglês (idioma padrão) também é suportado com a pesquisa semântica completa.

Nome do idioma	Código do idioma
Árabe	`ar`
Armênio	`hy`
Basco	`eu`
Bengali	`bn`
Búlgaro	`bg`

Nome do idioma	Código do idioma
Catalão	`ca`
Chinês: simplificado e tradicional*	`zh`
Tcheco	`cs`
Dinamarquês	`da`
Holandês	`nl`
Finlandês	`fi`
Francês: inclui francês (Canadá) *	`fr`
Galego	`gl`
Alemão*	`de`
Grego	`el`
Hindi	`hi`
Húngaro	`hu`
Indonésio	`id`
Irlandês	`ga`
Italiano	`it`
Japonês*	`ja`
Coreano*	`ko`
Letão	`lv`
Lituano	`lt`
Norueguês	`no`

Nome do idioma	Código do idioma
Persa	`fa`
Português	`pt`
Português (Brasil)	`pt-BR`
Romeno	`ro`
Russo	`ru`
Sorani	`ckb`
Espanhol: inclui espanhol (México)*	`es`
Sueco	`sv`
Turco	`tr`

*A pesquisa semântica é suportada para o idioma.

Para idiomas que oferecem suporte à pesquisa semântica, os seguintes recursos são suportados.

- Relevância do documento além da simples correspondência de palavras-chave.
- Perguntas frequentes além da simples correspondência de palavras-chave.
- Extraindo respostas de documentos com base na compreensão Amazon Kendra de leitura.
- Buckets de confiança (muito alto, alto, médio e baixo) dos resultados da pesquisa.

Para idiomas que não oferecem suporte à pesquisa semântica, a correspondência simples de palavras-chave é suportada para relevância do documento e perguntas frequentes.

Sinônimos (incluindo sinônimos personalizados), aprendizado e feedback incrementais e sugestões de consulta são compatíveis somente no inglês (idioma padrão).

Configurando Amazon Kendra para usar um Amazon VPC

Amazon Kendra pode se conectar a uma nuvem privada virtual (VPC) que você criou Amazon Virtual Private Cloud para indexar o conteúdo armazenado em fontes de dados executadas em sua nuvem

privada. Ao criar um conector de fonte de dados, você pode fornecer identificadores de grupo de segurança e sub-rede para a sub-rede que contém sua fonte de dados. Com essas informações, Amazon Kendra cria uma interface de rede elástica que é usada para se comunicar com segurança com sua fonte de dados na sua VPC.

Para configurar um conector de fonte de Amazon Kendra dados com Amazon VPC, você pode usar a operação AWS Management Console ou a CreateDataSourceAPI. Se você usa o console, conecta uma VPC durante o processo de configuração do conector.

> **Note**
>
> O Amazon VPC recurso é opcional ao configurar um conector Amazon Kendra de fonte de dados. Se sua fonte de dados estiver acessível pela Internet pública, você não precisará ativar o Amazon VPC recurso. Nem todos os conectores Amazon Kendra de fonte de dados são compatíveis Amazon VPC.

Se sua fonte de dados não estiver em execução Amazon VPC e não estiver acessível pela Internet pública, primeiro conecte sua fonte de dados à sua VPC usando uma rede privada virtual (VPN). Em seguida, você pode conectar sua fonte de dados Amazon Kendra usando uma combinação de Amazon VPC AWS Virtual Private Network e. Para obter informações sobre como configurar uma VPN, consulte a AWS VPN documentação.

Tópicos

- Configurando o Amazon VPC suporte para conectores Amazon Kendra
- Configurar uma fonte Amazon Kendra de dados à qual se conectar Amazon VPC
- Conectar um banco de dados em uma VPC
- Solução de problemas de conexão VPC

Configurando o Amazon VPC suporte para conectores Amazon Kendra

Para configurar Amazon VPC para uso com seus Amazon Kendra conectores, siga as etapas a seguir.

Etapas

- Etapa 1. Crie Amazon VPC sub-redes para Amazon Kendra
- Etapa 2. Crie grupos Amazon VPC de segurança para Amazon Kendra

- Etapa 3. Configure sua fonte de dados externa e Amazon VPC

Etapa 1. Crie Amazon VPC sub-redes para Amazon Kendra

Crie ou escolha uma Amazon VPC sub-rede existente que Amazon Kendra possa ser usada para acessar sua fonte de dados. As sub-redes preparadas devem estar em uma das seguintes zonas Regiões da AWS de disponibilidade:

- Oeste dos EUA (Oregon) /us-west-2 —usw2-az1, usw2-az2, usw2-az3
- Leste dos EUA (Norte da Virgínia) /us-east-1—use1-az1, use1-az2, use1-az4
- Leste dos EUA (Ohio) /us-east-2—use2-az1, use2-az2, use2-az3
- Ásia-Pacífico (Tóquio) /ap-northeast-1—apne1-az1, apne1-az2, apne1-az4
- Ásia-Pacífico (Mumbai) /ap-south—aps1-az1, aps1-az2, aps1-az3
- Ásia-Pacífico (Singapura) /ap-southeast-1—apse1-az1, apse1-az2, apse1-az3
- Ásia-Pacífico (Sydney) /ap-southeast-2—apse2-az1, apse2-az2, apse2-az3
- Canadá (Central) /ca-central-1—cac1-az1, cac1-az2, cac1-az4
- Europa (Irlanda) /eu-west-1-1-az1euw1-az1, uew1-az2, euw1-az3
- Europa (Londres) /eu-west-2—usw2-az1, usw2-az2, usw2-az3

Sua fonte de dados deve estar acessível a partir das sub-redes que você forneceu ao Amazon Kendra conector.

Para obter mais informações sobre como configurar Amazon VPC sub-redes, consulte Sub-redes para você no Guia do usuário Amazon VPC da Amazon VPC.

Se for Amazon Kendra necessário rotear a conexão entre duas ou mais sub-redes, você poderá preparar várias sub-redes. Por exemplo, a sub-rede que contém sua fonte de dados está sem endereços IP. Nesse caso, você pode fornecer uma sub-rede adicional Amazon Kendra com endereços IP suficientes e conectada à primeira sub-rede. Se você listar várias sub-redes, as sub-redes devem conseguir se comunicar entre si.

Etapa 2. Crie grupos Amazon VPC de segurança para Amazon Kendra

Para conectar seu conector de fonte de Amazon Kendra dados Amazon VPC, você deve preparar um ou mais grupos de segurança da sua VPC para atribuir. Amazon Kendra Os grupos de segurança

serão associados à interface de rede elástica criada por Amazon Kendra. Essa interface de rede controla o tráfego de entrada e saída de e para o Amazon VPC acesso às Amazon Kendra sub-redes.

Certifique-se de que as regras de saída do seu grupo de segurança permitam que o tráfego dos conectores da fonte de Amazon Kendra dados acesse as sub-redes e a fonte de dados com a qual você vai sincronizar. Por exemplo, você pode usar um MySQL conector para sincronizar a partir de um MySQL banco de dados. Se você estiver usando a porta padrão, os grupos de segurança devem permitir o acesso Amazon Kendra à porta 3306 no host que executa o banco de dados.

Recomendamos que você configure um grupo de segurança padrão com os seguintes valores Amazon Kendra para uso:

- Regras de entrada — Se você optar por deixar isso vazio, todo o tráfego de entrada será bloqueado.
- Regras de saída — adicione uma regra para permitir que todo o tráfego de saída Amazon Kendra possa iniciar as solicitações de sincronização da sua fonte de dados.
  - Versão IP — IPv4
  - Tipo — Todo o tráfego
  - Protocolo — Todo o tráfego
  - Alcance de portas — Todos
  - Destino — 0.0.0.0/0

Para obter mais informações sobre como configurar grupos Amazon VPC de segurança, consulte Regras de grupos de segurança no Guia do usuário da Amazon VPC.

Etapa 3. Configure sua fonte de dados externa e Amazon VPC

Certifique-se de que sua fonte de dados externa tenha as configurações de permissões e de rede corretas Amazon Kendra para acessá-la. Você pode encontrar instruções detalhadas sobre como configurar suas fontes de dados na seção de pré-requisitos de cada página do conector.

Além disso, verifique suas Amazon VPC configurações e certifique-se de que sua fonte de dados externa possa ser acessada pela sub-rede à qual você atribuirá. Amazon Kendra Para fazer isso, recomendamos que você crie uma Amazon EC2 instância na mesma sub-rede com os mesmos grupos de segurança e teste o acesso à sua fonte de dados a partir dessa Amazon EC2 instância. Para obter mais informações, consulte Solução de problemas de Amazon VPC conexão.

Configurar uma fonte Amazon Kendra de dados à qual se conectar Amazon VPC

Ao adicionar uma nova fonte de dados Amazon Kendra, você pode usar o Amazon VPC recurso se o conector da fonte de dados selecionado suportar esse recurso.

Você pode configurar uma nova fonte de Amazon Kendra dados Amazon VPC ativada usando a AWS Management Console ou a Amazon Kendra API. Especificamente, use a operação de CreateDataSourceAPI e, em seguida, use o `VpcConfiguration` parâmetro para fornecer as seguintes informações:

- `SubnetIds`— Uma lista de identificadores de sub-redes Amazon VPC
- `SecurityGroupIds`— Uma lista de identificadores de grupos de Amazon VPC segurança

Se você usa o console, fornece as Amazon VPC informações necessárias durante a configuração do conector. Para usar o console para habilitar o recurso Amazon VPC para um conector, primeiro você escolhe um Amazon VPC. Em seguida, você fornece identificadores de todas as sub-redes da Amazon VPC e identificadores de qualquer grupo de segurança da Amazon VPC. Você pode escolher as sub-redes da Amazon VPC e os grupos de segurança da Amazon VPC que você criou em Configurar a Amazon VPC ou usar qualquer um existente.

Tópicos

- Visualizando Amazon VPC identificadores
- Verificando sua IAM função de fonte de dados

Visualizando Amazon VPC identificadores

Os identificadores para sub-redes e grupos de segurança são configurados no console. Amazon VPC Para visualizar os identificadores, use os procedimentos a seguir.

Para visualizar identificadores de sub-rede

1. Faça login AWS Management Console e abra o console da Amazon VPC em https://console.aws.amazon.com/vpc/.
2. No painel de navegação, escolha Sub-redes.
3. Na lista Sub-redes, escolha a sub-rede que contém seu servidor de banco de dados.

4. Na guia Detalhes, anote o identificador no campo ID da sub-rede.

Para visualizar identificadores de grupos de segurança

1. Faça login AWS Management Console e abra o console da Amazon VPC em https://console.aws.amazon.com/vpc/.
2. No painel de navegação, escolha Grupos de segurança.
3. Na lista de grupo de segurança, escolha o grupo para o qual você deseja obter o identificador.
4. Na guia Detalhes, anote o identificador no campo ID do grupo de segurança.

Verificando sua IAM função de fonte de dados

Certifique-se de que sua função de fonte de dados AWS Identity and Access Management IAM (conector) contenha permissões para acessar seu Amazon VPC.

Se você usar o console para criar uma nova função para sua IAM função, Amazon Kendra adicionará automaticamente as permissões corretas à sua IAM função em seu nome. Se você usa a API ou usa uma IAM função existente, verifique se sua função contém permissões de acesso Amazon VPC. Para verificar se você tem as permissões corretas, consulte IAM papéis para VPC.

Você pode modificar uma fonte de dados existente para usar uma Amazon VPC sub-rede diferente. No entanto, verifique a IAM função da fonte de dados e, se necessário, modifique-a para refletir a alteração para que o conector da fonte de Amazon Kendra dados funcione corretamente.

Conectar um banco de dados em uma VPC

O exemplo a seguir mostra como conectar um MySQL banco de dados executado em uma nuvem privada virtual (VPC). O exemplo pressupõe que você esteja começando com sua VPC padrão e que precise criar um MySQL banco de dados. Se você já tem uma VPC, verifique se ela está configurada conforme mostrado. Se você tiver um MySQL banco de dados, poderá usá-lo em vez de criar um novo.

Etapas

- Etapa 1: Configurar uma VPC
- Etapa 2: criar e configurar grupos de segurança
- Etapa 3: criar um banco de dados

- Etapa 4: criar um conector de fonte de dados

Etapa 1: Configurar uma VPC

Configure sua VPC para que você tenha uma sub-rede privada e um grupo de segurança Amazon Kendra para acessar um MySQL banco de dados em execução na sub-rede. As sub-redes fornecidas na configuração da VPC devem estar na região Oeste dos EUA (Oregon), na região Leste dos EUA (Norte da Virgínia) ou na região Europa (Irlanda).

Para configurar uma VPC usando Amazon VPC

1. Faça login AWS Management Console e abra o console da Amazon VPC em https://console.aws.amazon.com/vpc/.
2. No painel de navegação, escolha Tabelas de rotas e selecione Criar tabela de rotas.
3. Para o campo Nome, insira**`Private subnet route table`**. No menu suspenso VPC, selecione sua VPC e escolha Criar tabela de rotas. Para retornar à lista de tabelas, escolha Fechar.
4. No painel de navegação, escolha Gateways NAT e, em seguida, escolha Criar gateway NAT.
5. No menu suspenso Sub-rede, escolha a sub-rede que é a sub-rede pública. Anote o ID da sub-rede.
6. Se você não tiver um endereço IP elástico, escolha Criar novo EIP, Criar um gateway NAT e, em seguida, escolha Fechar.
7. No painel de navegação, escolha Tabelas de rotas.
8. Na lista da tabela de rotas, escolha a tabela de rotas de sub-rede privada que você criou na etapa 3. Em Ações, escolha Editar rotas.
9. Escolha Adicionar rota. Para o destino, insira **`0.0.0.0/0`** para permitir todo o tráfego de saída para a Internet. Em Destino, escolha Gateway NAT e, em seguida, escolha o gateway criado na etapa 4. Escolha Salvar alterações e, em seguida, escolha Fechar.
10. No menu Ações, escolha Editar associações de sub-redes.
11. Escolha as sub-redes que você deseja que sejam privadas. Não escolha a sub-rede com o gateway NAT que você anotou anteriormente. Escolha Salvar associações quando terminar.

Etapa 2: criar e configurar grupos de segurança

Em seguida, configure os grupos de segurança para o banco de dados

Para criar e configurar grupos de segurança

1. Faça login AWS Management Console e abra o console da Amazon VPC em https://console.aws.amazon.com/vpc/.
2. Na descrição da VPC, observe o CIDR IPv4.
3. No painel de navegação, escolha Grupos de segurança e, em seguida, escolha Criar grupo de segurança.
4. Em Nome do grupo de segurança, insira **`DataSourceInboundSecurityGroup`**. Forneça uma descrição e escolha a VPC na lista. Escolha Criar grupo de segurança e, em seguida, escolha Fechar.
5. Escolha a guia Regras de entrada.
6. Escolha Editar regras de entrada e, em seguida, escolha Adicionar regra
7. Para um banco de dados, insira o número da porta para o intervalo de portas. Por exemplo, para MySQL é e**`3306`**, para HTTPS, é**`443`**. Em Fonte, insira o Encaminhamento Entre Domínios Sem Classificação (CIDR) da VPC. Escolha Salvar regras e escolha Fechar.

O grupo de segurança permite que qualquer pessoa dentro da VPC se conecte ao banco de dados e permite conexões de saída com a internet.

Etapa 3: criar um banco de dados

Crie um banco de dados para armazenar os documentos ou use o banco de dados existente.

Para obter instruções sobre como criar um MySQL banco de dados, consulte MySQL.

Etapa 4: criar um conector de fonte de dados

Depois de configurar sua VPC e criar seu banco de dados, você pode criar um conector de fonte de dados para o banco de dados. Para obter informações sobre conectores de banco de dados Amazon Kendra compatíveis, consulte Conectores compatíveis.

Para seu banco de dados, certifique-se de configurar sua VPC, as sub-redes privadas que você criou em sua VPC e o grupo de segurança que você criou em sua VPC.

Solução de problemas de conexão VPC

Se você encontrar algum problema com sua conexão de nuvem privada virtual (VPC), verifique se IAM as permissões, as configurações do grupo de segurança e as tabelas de rotas da sub-rede estão configuradas corretamente.

Uma causa potencial de falha na sincronização do conector da fonte de dados é que a fonte de dados pode estar inacessível a partir da sub-rede à qual você atribuiu. Amazon Kendra Para solucionar esse problema, recomendamos que você crie uma Amazon EC2 instância com as mesmas Amazon VPC configurações. Em seguida, tente acessar a fonte de dados dessa Amazon EC2 instância usando chamadas da API REST ou outros métodos (com base no tipo específico da sua fonte de dados).

Se você acessar com êxito a fonte de dados da Amazon EC2 instância que você criou, isso significa que sua fonte de dados pode ser acessada por essa sub-rede. Portanto, seu problema de sincronização não está relacionado ao fato de sua fonte de dados estar inacessível pelo Amazon VPC.

Se você não conseguir acessar sua Amazon EC2 instância a partir da configuração da VPC e validá-la com a Amazon EC2 instância que você criou, precisará solucionar mais problemas. Por exemplo, se você tiver um Amazon S3 conector cuja sincronização falhou com erros sobre problemas de conexão, você pode configurar uma Amazon EC2 instância com a mesma Amazon VPC configuração que você atribuiu ao seu Amazon S3 conector. Em seguida, use essa instância do Amazon EC2 para testar se a sua Amazon VPC foi configurada corretamente.

Veja a seguir um exemplo de configuração de uma Amazon EC2 instância para solucionar problemas de Amazon VPC conexão com uma fonte de Amazon S3 dados.

Tópicos

- Etapa 1: executar uma Amazon EC2 instância
- Etapa 2: conectar-se à Amazon EC2 instância
- Etapa 3: testar o Amazon S3 acesso

Etapa 1: executar uma Amazon EC2 instância

1. Faça login no AWS Management Console e abra o console do Amazon EC2 em https://console.aws.amazon.com/ec2/.
2. Selecione Executar uma instância.

3. Escolha Configurações de rede, escolha Editar e faça o seguinte:

   a. Escolha a mesma VPC e sub-rede às quais você atribuiu. Amazon Kendra

   b. Para Firewall (grupos de segurança), escolha Selecionar grupo de segurança existente. Em seguida, selecione o grupo de segurança ao qual você atribuiu Amazon Kendra.

      > **Note**
      >
      > O grupo de segurança deve permitir o tráfego de saída para o. Amazon S3

   c. Defina Atribuição automática de IP público como Desabilitar.

   d. Em Detalhes avançados, faça o seguinte:

      - Em Perfil de instância do IAM, selecione Criar novo perfil do IAM para criar e anexar um perfil de IAM instância à sua instância. Certifique-se de que o perfil tenha permissões de acesso Amazon S3. Para obter mais informações, consulte Como posso conceder acesso à minha Amazon EC2 instância a um Amazon S3 bucket? em AWS re:Post.
      - Deixe todas as outras configurações como padrão.

   e. Analise e execute a Amazon EC2 instância.

Etapa 2: conectar-se à Amazon EC2 instância

Depois que sua Amazon EC2 instância estiver em execução, acesse a página de detalhes da instância e conecte-se à sua instância. Para fazer isso, use as etapas em Conecte-se às suas instâncias sem exigir um endereço IPv4 público usando o EC2 Instance Connect Endpoint no Guia do Amazon EC2 usuário para instâncias Linux.

Etapa 3: testar o Amazon S3 acesso

Depois de se conectar ao terminal da Amazon EC2 instância, execute um AWS CLI comando para testar a conexão dessa sub-rede privada com seu Amazon S3 bucket.

Para testar o Amazon S3 acesso, digite o seguinte AWS CLI comando no AWS CLI: `aws s3 ls`

Depois que o AWS CLI comando for executado, analise o seguinte:

- Se você configurou IAM as permissões necessárias corretamente e sua Amazon S3 configuração está correta, você deve ver uma lista dos seus Amazon S3 buckets.

- Se você ver erros de permissãoAccess Denied, como, é provável que sua configuração de VPC esteja correta, mas algo esteja errado com suas IAM permissões ou com sua política Amazon S3 de bucket.

Se o comando estiver atingindo o tempo limite, é provável que sua conexão esteja expirando porque a configuração da VPC está incorreta e a instância do Amazon EC2 não pode acessar o Amazon S3 a partir da sua sub-rede. Reconfigure sua VPC e tente novamente.

Excluindo um índice, uma fonte de dados ou documentos carregados em lote

Esta seção mostra como excluir um índice, um repositório de fontes de dados de documentos em seu índice ou documentos no índice que você carregou em lote.

Tópicos

- Excluir um índice
- Excluir uma fonte de dados
- Excluindo documentos carregados em lote

Excluir um índice

Você pode excluir um índice do Amazon Kendra quando não estiver mais usando o índice. Por exemplo, exclua um índice quando:

- Você não está mais usando o índice e deseja reduzir as cobranças em sua conta AWS. Um índice do Amazon Kendra acumula cobranças enquanto está em execução, independentemente de você fazer consultas no índice ou não.
- Você deseja reconfigurar o índice para uma edição diferente do Amazon Kendra. Exclua o índice existente e crie um novo com a edição diferente.
- Você atingiu o número máximo de índices em sua conta e não deseja exceder sua cota. Exclua um índice existente e adicione um novo. Para obter informações sobre o número máximo de índices que podem ser criados, consulte Cotas.

Para excluir um índice, use o console, o AWS Command Line Interface, o script AWS CloudFormation ou a API `DeleteIndex`. A exclusão de um índice remove o índice e todas as fontes de dados e dados do documento associados. A exclusão de um índice não remove os documentos originais do seu armazenamento.

A exclusão de um índice é uma operação assíncrona. Quando você começa a excluir um índice, o status do índice muda par a`DELETING`. Ele permanece no estado `DELETING` até que todas as informações relacionadas ao índice sejam removidas. Depois que o índice é excluído, ele não aparece mais nos resultados de uma chamada para a API ListIndices. Se você chamar a API DescribeIndex com o identificador do índice excluído, receberá uma exceção`ResourceNotFound`.

Para excluir um índice (console)

1. Faça login no AWS Management Console e abra o console Amazon Kendra em https://console.aws.amazon.com/kendra/.
2. No painel de navegação, escolha Índices e depois escolha o índice a ser excluído.
3. Selecione Excluir para excluir o índice selecionada.

Para excluir um índice (CLI)

- Na AWS CLI, use o comando a seguir. O comando a seguir é formatado para Linux e macOS. Para Windows, substitua o caractere de continuação de linha do Unix (\) pelo circunflexo (^).

```
aws kendra delete-index \
   --id index-id
```

Excluir uma fonte de dados

Você exclui uma fonte de dados quando deseja remover as informações contidas na fonte de dados do índice do Amazon Kendra. Por exemplo, exclua uma fonte de dados quando:

- Uma fonte de dados está configurada incorretamente. Exclua a fonte de dados, aguarde a conclusão da exclusão da fonte de dados e, em seguida, recrie-a.
- Você migrou documentos de uma fonte de dados para outra. Exclua a fonte de dados original e recrie-a no novo local.
- Você atingiu o limite de fontes de dados para um índice. Exclua uma das fontes de dados existentes e adicione uma nova. Para obter mais informações sobre o número de fontes de dados que você pode criar, consulte Cotas.

Para excluir uma fonte de dados, use o console, o AWS Command Line Interface (AWS CLI), a API `DeleteDataSource` ou um script AWS CloudFormation. A exclusão de uma fonte de dados remove todas as informações sobre a fonte de dados do índice. Se você quiser apenas interromper a sincronização da fonte de dados, altere a programação de sincronização da fonte de dados para “executar sob demanda”.

A exclusão de uma fonte de dados é uma operação assíncrona. Quando você começa a excluir uma fonte de dados, o status da fonte de dados muda para `DELETING`. Ele permanece no estado

`DELETING` até que todas as informações relacionadas à fonte de dados sejam removidas. Depois que a fonte de dados é excluída, ela não aparece mais nos resultados de uma chamada para a API ListIndices. Se você chamar a API DescribeIndex com a fonte de dados excluída, receberá uma exceção `ResourceNotFound`.

> **Note**
>
> Excluir uma fonte de dados inteira ou ressincronizar o índice após excluir documentos específicos de uma fonte de dados pode levar até uma hora ou mais, dependendo do número de documentos que você deseja excluir.

Como excluir uma fonte de dados (console)

1. Faça login no AWS Management Console e abra o console Amazon Kendra em https://console.aws.amazon.com/kendra/.
2. No painel de navegação, escolha Índices e, em seguida, escolha o índice que contém a fonte de dados a ser excluída.
3. No painel de navegação, escolha Fontes dos dados.
4. Escolha a fonte de dados para remover.
5. Escolha Excluir para excluir a fonte de dados.

Como excluir uma fonte de dados (CLI)

- Na AWS Command Line Interface, use o comando a seguir. O comando a seguir é formatado para Linux e macOS. Para Windows, substitua o caractere de continuação de linha do Unix (\) pelo circunflexo (^).

```
aws kendra delete-data-source \
   --id data-source-id \
   --index-id index-id
```

Ao excluir uma fonte de dados, o Amazon Kendra remove todas as informações armazenadas sobre a fonte de dados. O Amazon Kendra remove todos os dados do documento armazenados no índice e todos os históricos e métricas de execução associados à fonte de dados. A exclusão de uma fonte de dados não remove os documentos originais de armazenamento.

Os documentos na fonte de dados podem ser incluídos na contagem de documentos retornada pela API `DescribeIndex` enquanto o Amazon Kendra exclui uma fonte de dados. Documentos da fonte de dados podem aparecer nos resultados da pesquisa enquanto o Amazon Kendra exclui a fonte de dados.

O Amazon Kendra libera os recursos de uma fonte de dados quando você chama a API `DeleteDataSource`ou opta por excluir a fonte de dados no console. Se estiver excluindo a fonte de dados para reduzir o número de fontes de dados para abaixo do limite, poderá criar uma nova fonte de dados imediatamente.

Se estiver excluindo uma fonte de dados e depois criando outra fonte de dados para os dados do documento, aguarde até que a primeira fonte de dados seja excluída antes de sincronizar a nova fonte de dados.

Você pode excluir uma fonte de dados que está em processo de sincronização com o Amazon Kendra. A sincronização é interrompida e a fonte de dados é removida. Se você tentar iniciar uma sincronização quando a fonte de dados estiver sendo excluída, você receberá uma exceção `ConflictException`.

Você não pode excluir uma fonte de dados se o índice associado estiver no estado `DELETING`. A exclusão de um índice exclui todas as fontes de dados do índice. Você pode começar a excluir um índice enquanto a fonte de dados desse índice estiver no estado `DELETING`.

Se você tiver duas fontes de dados apontando para os mesmos documentos, como duas fontes de dados apontando para o mesmo bucket do Amazon S3, os documentos no índice podem ficar inconsistentes quando uma das fontes de dados for excluída. Quando duas fontes de dados fazem referência aos mesmos documentos, somente uma cópia dos dados do documento é armazenada no índice. A remoção de uma fonte de dados remove os dados de índice dos documentos. A outra fonte de dados não sabe que os documentos foram removidos, então Amazon Kendra não os reindexará corretamente na próxima vez em que for sincronizado. Quando você tem duas fontes de dados apontando para o mesmo local do documento, exclua as duas fontes de dados e recrie uma.

Excluindo documentos carregados em lote

Você pode excluir documentos diretamente de um índice usando a API BatchDeleteDocument. Você não pode excluir documentos diretamente usando o console. Se usar o console, é possível excluir documentos específicos do repositório de fontes de dados e sincronizar novamente com o índice ou excluir todo o conector da fonte de dados.

A exclusão de um índice usando `BatchDeleteDocument` é uma operação assíncrona. Depois de chamar a API `BatchDeleteDocument`, use API BatchGetDocumentStatus para monitorar o progresso da exclusão dos documentos. Quando um documento é excluído do índice, Amazon Kendra retorna `NOT_FOUND` como status.

> **Note**
>
> A exclusão de documentos de um índice usando `BatchDeleteDocument` pode levar até uma hora ou mais, dependendo do número de documentos excluídos.

Para excluir documentos carregados em lote de um índice (CLI)

- Na AWS Command Line Interface, use o comando a seguir. O comando a seguir é formatado para Linux e macOS. Para Windows, substitua o caractere de continuação de linha do Unix (\) pelo circunflexo (^).

```
aws kendra batch-delete-document \
   --index-id index-id \
   --document-id-list 'doc-id-1' 'doc-id-2'
```

Enriquecendo seus documentos durante a absorção

Você pode alterar os campos ou atributos de metadados do conteúdo e do documento durante o processo de absorção de documentos. Com o atributo do Amazon Kendra de Enriquecimento personalizado de documentos, você pode criar, modificar ou excluir atributos e conteúdo do documento ao absorver os documentos. Amazon Kendra Isso significa que você pode manipular e absorver os dados conforme necessário.

Esse atributo oferece controle sobre como os documentos são tratados e absorvidos. Amazon Kendra Por exemplo, você pode limpar as informações de identificação pessoal nos metadados do documento ao absorver os documentos. Amazon Kendra

Outra forma de usar esse atributo é invocar uma função do Lambda no AWS Lambda para executar o reconhecimento óptico de caracteres (OCR) em imagens, tradução em texto e outras tarefas para preparar os dados para pesquisa ou análise. Por exemplo, você pode invocar uma função para executar o OCR em imagens. A função pode interpretar texto de imagens e tratar cada imagem como um documento textual. Uma empresa que recebe pesquisas de clientes enviadas por e-mail e as armazena como imagens pode gerar essas imagens como documentos de texto no Amazon Kendra. A empresa pode então pesquisar informações valiosas sobre pesquisas com clientes no Amazon Kendra.

Você pode usar operações básicas para aplicar como a primeira análise dos dados e, em seguida, usar uma função do Lambda para aplicar operações mais complexas nos dados. Por exemplo, use uma operação básica para simplesmente remover todos os valores no campo de metadados do documento 'Customer_ID' e depois aplicar uma função do Lambda para extrair texto das imagens do texto nos documentos.

Como funciona o enriquecimento personalizado de documentos

O processo geral de enriquecimento personalizado de documentos é o seguinte:

1. Configure o enriquecimento personalizado de documentos ao criar ou atualizar a fonte de dados ou ao indexar os documentos diretamente no Amazon Kendra.
2. O Amazon Kendra aplica configurações em linha ou lógica básica para alterar os dados. Para obter mais informações, consulte the section called “Operações básicas para alterar metadados”.
3. Se optar por configurar a manipulação avançada de dados, Amazon Kendra poderá aplicá-la nos documentos originais brutos ou nos documentos estruturados e analisados. Para obter mais

informações, consulte the section called “Funções do Lambda: extrair e alterar metadados ou conteúdo”.

4. Os documentos alterados são inseridos em Amazon Kendra

Em qualquer momento desse processo, se a configuração não for válida, Amazon Kendra gerará um erro.

Ao escolher as APIs CreateDataSource, UpdateDataSource ou BatchputDocument, você fornece a configuração personalizada de enriquecimento de documentos. Se escolher `BatchPutDocument`, deverá configurar o enriquecimento personalizado de documentos com cada solicitação. Se você usa o console, seleciona o índice e, em seguida, selecione Enriquecimentos de documentos para configurar o enriquecimento personalizado de documentos.

Se você usar o enriquecimentos de documentos no console, poderá optar por configurar somente as operações básicas, somente as funções do Lambda ou as duas, como ao usar a API. Selecione Avançar nas etapas do console para optar por não configurar operações básicas e somente as funções do Lambda, incluindo se deseja aplicar aos dados originais (pré-extração) ou estruturados (pós-extração). Você só pode salvar as configurações ao concluir todas as etapas no console. As configurações do documento não serão salvas se você não concluir todas as etapas.

Operações básicas para alterar metadados

Você pode manipular os campos e o conteúdo do documento usando a lógica básica. Isso inclui remover valores em um campo, modificar valores em um campo usando uma condição ou criar um campo. Para manipulações avançadas que vão além do que você pode manipular usando a lógica básica, invoque uma função do Lambda. Para obter mais informações, consulte the section called “Funções do Lambda: extrair e alterar metadados ou conteúdo”.

Para aplicar a lógica básica, especifique o campo de destino que deseja manipular usando o objeto DocumentAttributeTarget. Forneça a chave de atributo. Por exemplo, a chave “Departamento” é um campo ou atributo que contém todos os nomes de departamentos associados aos documentos. Você também pode especificar um valor a ser usado no campo de destino se uma determinada condição for atendida. Defina a condição usando o objeto DocumentAttributeCondition. por exemplo, se o campo “Source_URI” contiver “financeiro” como valor de URI, o campo de destino “Departmento” será preenchido previamente com o valor de destino “Financeiro” para o documento. Você também pode excluir os valores do atributo do documento de destino.

Para aplicar a lógica básica usando o console, selecione o índice e, em seguida, selecione Enriquecimentos de documentos no menu de navegação. Acesse Configurar operações básicas para aplicar manipulações básicas aos campos e ao conteúdo do documento.

O exemplo a seguir é do uso da lógica básica para remover todos os números de identificação do cliente no campo “Customer_ID”.

Exemplo 1: remoção dos números de identificação do cliente associados aos documentos

Dados antes da aplicação da manipulação básica.

Document_ID	Body_Text	Customer_ID
1	Lorem Ipsum.	CID1234
2	Lorem Ipsum.	CID1235
3	Lorem Ipsum.	CID1236

Dados antes da aplicação da manipulação básica.

Document_ID	Body_Text	Customer_ID
1	Lorem Ipsum.	
2	Lorem Ipsum.	
3	Lorem Ipsum.	

Veja a seguir um exemplo do uso da lógica básica para criar um campo chamado “Departamento” e preencher previamente esse campo com os nomes dos departamentos com base nas informações do campo do “Source_URI”. Isso define a condição estabelecendo que, se o campo “Source_URI” contiver “financial” (financeiro) como valor de URI, o campo de destino “Department” (Departamento) será preenchido previamente com o valor de destino “Financeiro” para o documento.

Exemplo 2: criar o campo “Departamento” e preenchê-lo previamente com os nomes dos departamentos associados aos documentos usando uma condição.

Dados antes da aplicação da manipulação básica.

Document_ID	Body_Text	Source_URI
1	Lorem Ipsum.	financeiro/1
2	Lorem Ipsum.	financeiro/2
3	Lorem Ipsum.	financeiro/3

Dados antes da aplicação da manipulação básica.

Document_ID	Body_Text	Source_URI	Departamento
1	Lorem Ipsum.	financeiro/1	Financeiro
2	Lorem Ipsum.	financeiro/2	Financeiro
3	Lorem Ipsum.	financeiro/3	Financeiro

> **Note**
>
> O Amazon Kendra não poderá criar um campo de documento de destino se ele ainda não tiver sido criado como um campo de índice. Depois de criar o campo de índice, você poderá criar um campo de documento usando`DocumentAttributeTarget`. Em seguida, o Amazon Kendra mapeia o campo de metadados do documento recém-criado para o campo de índice.

O código a seguir é um exemplo de como configurar a manipulação básica de dados para remover os números de identificação do cliente associados aos documentos.

Console

Para configurar a manipulação básica de dados para remover números de identificação do cliente

1. No painel de navegação esquerdo, em Índices, selecione Enriquecimentos de documentos e, em seguida, selecione Adicionar enriquecimento de documentos.

2. Na página Configurar operações básicas, escolha na lista suspensa a fonte de dados que você deseja alterar os campos e o conteúdo do documento. Em seguida, escolha no menu suspenso o nome do campo do documento “Customer_ID”, selecione no menu suspenso o nome do campo do índice “Customer_ID” e selecione no menu suspenso a ação de destino Excluir. Em seguida, selecione Adicionar operação básica.

CLI

Para configurar a manipulação básica de dados para remover números de identificação do cliente

```
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 a manipulação básica de dados para remover números de identificação do cliente

```
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 a manipulação básica de dados para remover números de identificação do cliente

```
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");
    }
}
```

Funções do Lambda: extrair e alterar metadados ou conteúdo

Você pode manipular os campos e o conteúdo do documento usando as funções do Lambda. Isso é útil se você quiser ir além da lógica básica e aplicar manipulações avançadas de dados. Por

exemplo, ao usar o reconhecimento óptico de caracteres (OCR), que interpreta texto de imagens e trata cada imagem como um documento de texto. Ou recuperando a data e hora atual em um determinado fuso horário e inserindo a data e hora em que há um valor vazio para um campo de data.

Você pode aplicar a lógica básica primeiro e depois usar uma função do Lambda para manipular ainda mais os dados, ou vice-versa. Você também pode optar por aplicar somente uma função do Lambda.

O Amazon Kendra pode invocar uma função do Lambda para aplicar manipulações avançadas de dados durante o processo de absorção como parte do CustomDocumentEnrichmentConfiguration. Você pode especificar uma função que inclui permissão para executar a função do Lambda e acessar o bucket do Amazon S3 para armazenar a saída de suas manipulações de dados: consulte funções de acesso do IAM.

O Amazon Kendra pode aplicar uma função do Lambda em seus documentos originais ou brutos ou nos documentos estruturados e analisados. Você pode configurar uma função do Lambda que pegue os dados originais ou brutos e aplique as manipulações de dados usando o PreExtractionHookConfiguration. Você também pode configurar uma função do Lambda que usa os documentos estruturados e aplica suas manipulações de dados usando o PostExtractionHookConfiguration. O Amazon Kendra extrai os metadados e o texto do documento para estruturar os documentos. As funções do Lambda devem seguir as estruturas obrigatórias de solicitação e resposta. Para obter mais informações, consulte the section called “Contratos de dados para funções do Lambda”.

Para aplicar a função do Lambda usando o console, selecione o índice e, em seguida, selecione Enriquecimentos de documentos no menu de navegação. Acesse Configurar funções do Lambda para configurar uma função do Lambda.

Você pode configurar somente uma função do Lambda para `PreExtractionHookConfiguration` e somente mais uma função do Lambda para `PostExtractionHookConfiguration`. No entanto, essa função do Lambda pode invocar outras funções necessárias. Você pode configurar `PreExtractionHookConfiguration` e `PostExtractionHookConfiguration` e/ou um só deles. A função do Lambda para `PreExtractionHookConfiguration` não deve exceder um tempo de execução de 5 minutos e a função do Lambda para `PostExtractionHookConfiguration` não deve exceder o tempo de execução de 1 minuto. A configuração do enriquecimento personalizado de documentos naturalmente leva mais tempo para absorver os documentos no Amazon Kendra do que se a opção não estiver configurada.

Você pode configurar o Amazon Kendra para invocar uma função do Lambda somente se uma condição for atendida. Por exemplo, você pode especificar uma condição determinando que, se houver valores de data e hora vazios, o Amazon Kendra deverá invocar uma função para inserir a data e hora atuais.

Veja a seguir um exemplo do uso de uma função do Lambda para executar o OCR para interpretar texto de imagens e armazenar esse texto em um campo chamado “Document_Image_Text”.

Exemplo 1: extraindo texto de imagens para criar documentos textuais

Dados antes da aplicação da manipulação avançada.

Document_ID	Document_Image
1	image_1.png
2	image_2.png
3	image_3.png

Dados depois da aplicação da manipulação avançada.

Document_ID	Document_Image	Document_Image_Text
1	image_1.png	Resposta de pesquisa enviada por e-mail
2	image_2.png	Resposta de pesquisa enviada por e-mail
3	image_3.png	Resposta de pesquisa enviada por e-mail

Veja a seguir um exemplo do uso de uma função do Lambda para inserir a data e hora atual para valores de data vazios. Isso usa a condição de que, se o valor do campo de data for “nulo”, ele deve ser substituído pela data e hora atual.

Exemplo 2: substituindo valores vazios no campo Last_Updated pela data e hora atual.

Dados antes da aplicação da manipulação avançada.

Document_ID	Body_Text	Last_Updated
1	Lorem Ipsum.	1º de janeiro de 2020
2	Lorem Ipsum.	
3	Lorem Ipsum.	1.º de julho de 2020

Dados depois da aplicação da manipulação avançada.

Document_ID	Body_Text	Last_Updated
1	Lorem Ipsum.	1º de janeiro de 2020
2	Lorem Ipsum.	1.º de dezembro de 2021
3	Lorem Ipsum.	1.º de julho de 2020

O código a seguir é um exemplo de configuração de uma função do Lambda para manipulação avançada de dados nos dados originais brutos.

Console

Para configurar uma função do Lambda para manipulação avançada de dados nos dados originais brutos

1. No painel de navegação esquerdo, em Índices, selecione Enriquecimentos de documentos e, em seguida, selecione Adicionar enriquecimento de documentos.
2. Na página Configurar funções do Lambda, na seção Lambda para pré-extração, selecione nos menus suspensos o ARN da função do Lambda e o bucket do Amazon S3. Adicione a função de acesso do IAM ao selecionar a opção de criar uma nova função no menu suspenso. Isso cria as permissões necessárias do Amazon Kendra para criar o enriquecimento do documento.

CLI

Para configurar uma função do Lambda para manipulação avançada de dados nos dados originais brutos

```
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 uma função do Lambda para manipulação avançada de dados nos dados originais brutos

```
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 uma função do Lambda para manipulação avançada de dados nos dados originais brutos

```
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 dados para funções do Lambda

As funções do Lambda para manipulação avançada de dados interagem com os contratos de dados do Amazon Kendra. Os contratos são as estruturas obrigatórias de solicitação e resposta das funções do Lambda. Se as funções do Lambda não seguirem essas estruturas, o Amazon Kendra gerará um erro.

A função do Lambda para `PreExtractionHookConfiguration` deve esperar a seguinte estrutura de solicitação:

```
{
    "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>
}
```

A estrutura `metadata`, que inclui a estrutura `CustomDocumentAttribute`, é a seguinte:

```
{
    "attributes": [<CustomDocumentAttribute<]
}

CustomDocumentAttribute
{
    "name": <str>,
    "value": <CustomDocumentAttributeValue>
}

CustomDocumentAttributeValue
{
    "stringValue": <str>,
    "integerValue": <int>,
    "longValue": <long>,
    "stringListValue": list<str>,
    "dateValue": <str>
}
```

A função do Lambda para `PreExtractionHookConfiguration` deve seguir a seguinte estrutura de resposta:

```
{
    "version": <str>,
    "dataBlobStringEncodedInBase64": <str>, //In the case of a data blob
    "s3ObjectKey": <str>, //In the case of an S3 bucket
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

A função do Lambda para `PostExtractionHookConfiguration` deve esperar a seguinte estrutura de solicitação:

```
{
    "version": <str>,
    "s3Bucket": <str>,
    "s3ObjectKey": <str>,
    "metadata": <Metadata>
}
```

A função do Lambda para `PostExtractionHookConfiguration` deve seguir a seguinte estrutura de resposta:

```
PostExtractionHookConfiguration Lambda Response
{
    "version": <str>,
    "s3ObjectKey": <str>,
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

O documento alterado é enviado para o bucket do Amazon S3. O documento alterado deve seguir o formato mostrado em the section called “Formato de documento estruturado”.

Formato de documento estruturado

O Amazon Kendra carrega o documento estruturado em um determinado bucket do Amazon S3. O documento estruturado segue esse formato:

```
Kendra document

{
   "textContent": <TextContent>
}

TextContent
{
  "documentBodyText": <str>
}
```

Exemplo de uma função do Lambda que adere aos contratos de dados

O código Python a seguir é um exemplo de uma função do Lambda que aplica manipulação avançada dos campos de metadados `_authors`, `_document_title` e do conteúdo do corpo nos documentos brutos ou originais.

No caso do conteúdo corporal estiver em um bucket do 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"]}}
        ]
    }
```

No caso do conteúdo corporal estiver em um bucket do blob de dados

```
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"]}}
        ]
    }
```

O código Python a seguir é um exemplo de uma função do Lambda que aplica manipulação avançada dos campos de metadados `_authors`, `_document_title` e do conteúdo do corpo nos documentos estruturados ou analisados.

```
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"]}}
        ]
    }
```

Pesquisando um índice

Para pesquisar um Amazon Kendra índice, você usa a API de consulta. A API de `Query` retorna informações sobre os documentos indexados que você usa no aplicativo. Esta seção mostra como fazer uma consulta, realizar filtros e interpretar a resposta que recebe da API de `Query`.

Para pesquisar documentos com os quais você indexou Amazon Lex, use Amazon Kendra a AMAZON. KendraSearchIntent. Para ver um exemplo de configuração Amazon Kendra com Amazon Lex, consulte Criação de um bot de perguntas frequentes para um Amazon Kendra índice.

Tópicos

- Consultar um índice
- Navegar em um índice
- Apresentar resultados da pesquisa
- Pesquisa tabular de HTML
- Sugestões de consulta
- Corretor ortográfico de consulta
- Filtragem e pesquisa de facetas
- Filtragem no contexto do usuário
- Respostas de consulta e tipos de resposta
- Ajuste e classificação de respostas
- Reduzir/expandir os resultados da consulta

Consultar um índice

Ao pesquisar seu índice, Amazon Kendra usa todas as informações fornecidas sobre seus documentos para determinar os documentos mais relevantes para os termos de pesquisa inseridos. Alguns dos itens Amazon Kendra considerados são:

- O texto ou o corpo do documento.
- O título do documento.
- Campos de texto personalizados que você marcou como pesquisáveis.

- O campo de data que você indicou deve ser usado para determinar a “atualização” de um documento.
- Qualquer outro campo que possa fornecer informações relevantes.

Amazon Kendra também pode filtrar a resposta com base em qualquer filtro de campo/atributo que você possa ter definido para a pesquisa. Por exemplo, se tiver um campo personalizado chamado “departamento”, poderá filtrar a resposta para retornar somente documentos de um departamento chamado “jurídico”. Para obter mais informações, consulte Campos ou atributos personalizados.

Os resultados da pesquisa retornados são classificados pela relevância que Amazon Kendra determina cada documento. Os resultados são paginados para que você possa mostrar uma página por vez para o usuário.

Para pesquisar documentos com os quais você indexou Amazon Lex, use Amazon Kendra a AMAZON. KendraSearchIntent. Para ver um exemplo de configuração Amazon Kendra com Amazon Lex, consulte Criação de um bot de perguntas frequentes para um Amazon Kendra índice.

O exemplo a seguir mostra como pesquisar um índice. Amazon Kendra determina o tipo de resultado da pesquisa (resposta, documento, pergunta-resposta) que é mais adequado para a consulta. Você não pode configurar Amazon Kendra para retornar um tipo específico de resposta de pesquisa (resposta, documento, pergunta-resposta) a uma consulta.

Para obter informações sobre as respostas de consulta, consulte Respostas de consulta e tipos de resposta.

Pré-requisitos

Antes de usar a API de consulta para consultar um índice:

- Configure as permissões necessárias para um índice e conecte-se à sua fonte de dados ou carregue seus documentos em lote. Para obter mais informações, consulte Funções do IAM. Use o nome de recurso da Amazon da função ao chamar a API para criar um conector de índice e fonte de dados ou fazer upload em lote de documentos.
- Configure o AWS Command Line Interface, um SDK ou acesse o Amazon Kendra console. Para obter mais informações, consulte Configurar Amazon Kendra.
- Crie um índice e conecte-se a uma fonte de dados de documentos ou faça upload de documentos em lote. Para obter mais informações, consulte Criar um índice e Criar de um conector de fonte de dados.

Pesquisar um índice (console)

Você pode usar o Amazon Kendra console para pesquisar e testar seu índice. Faça consultas e visualize os resultados.

Para pesquisar um índice com o console

1. Faça login no AWS Management Console e abra o Amazon Kendra console em http://console.aws.amazon.com/kendra/.
2. No painel de navegação, escolha Índices.
3. Escolha seu índice.
4. No menu de navegação, escolha a opção de pesquisar no índice.
5. Na caixa de texto Filtro, insira uma consulta e, depois, pressione Enter.
6. Amazon Kendra retorna os resultados da pesquisa.

Você também poderá obter o ID da consulta para a pesquisa selecionando o ícone de lâmpada no painel lateral.

Pesquisar um índice (SDK)

Para pesquisar um índice com Python ou Java

- O exemplo a seguir pesquisa um índice. Altere o valor de `query` para sua consulta de pesquisa `index_id` e/ou `indexId` para o identificador do índice que você deseja pesquisar.

  Você também poderá obter o ID da consulta para a pesquisa como parte dos elementos de resposta ao chamar a 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");
        }
    }
}
```

Pesquisar um índice (Postman)

Você pode usar o Postman para consultar e testar seu Amazon Kendra índice.

Para pesquisar um índice usando o Postman

1. Crie uma nova coleção no Postman e defina o tipo de solicitação como POST.
2. Insira o URL do endpoint. Por exemplo, https://kendra.<region>.amazonaws.com.
3. Selecione a guia Autorização e insira as seguintes informações:

   - Tipo – Selecione a Assinatura da AWS .
   - AccessKey—Insira a chave de acesso gerada ao criar um IAM usuário.
   - SecretKey—Insira a chave secreta gerada ao criar um IAM usuário.
   - AWS Região — Insira a região do seu índice. Por exemplo, Oeste-EUA-2.
   - Nome do serviço – Insira kendra. Isso diferencia maiúsculas de minúsculas, então deve ser minúsculo.

   > **Warning**
   >
   > Se digitar o nome do serviço incorreto ou não usar letras minúsculas, um erro será gerado ao selecionar Enviar para enviar a solicitação: “A credencial deve ter como escopo o serviço correto 'kendra'.”
   > Você também deverá verificar se inseriu a chave de acesso e a chave secreta corretas.

4. Selecione a guia Cabeçalhos e insira as seguintes informações de chave e valor:

   - Chave: X-Amz-Target

     Valor: com.amazonaws.kendra. AWSKendraFrontendService.Consulta
   - Chave: Codificação de conteúdo

     Valor: amz-1.0

5. Selecione a guia Corpo e faça o seguinte:

   - Escolha o tipo JSON bruto para o corpo da solicitação.
   - Insira um JSON que inclua seu ID de índice e o texto da consulta.

```
{
    "IndexId": "index-id",
    "QueryText": "enter a query here"
```

```
}
```

> **Warning**
>
> Se seu JSON não usar a indentação correta, um erro será gerado: "". SerializationException Verifique o recuo em seu JSON.

6. Selecione Enviar (próximo ao canto superior direito).

Pesquisar com sintaxe de consulta avançada

Crie consultas mais específicas do que consultas simples de palavra-chave ou linguagem natural usando sintaxe ou operadores de consulta avançados. Isso inclui intervalos, booleanos, curingas e muito mais. Ao usar operadores, poderá dar mais contexto à sua consulta e refinar ainda mais os resultados da pesquisa.

Amazon Kendra suporta os seguintes operadores.

- Booleano: lógica para limitar ou ampliar a pesquisa. Por exemplo, limita a pesquisa do `amazon AND sports`, a pesquisar somente documentos que contenham os dois termos.
- Parênteses: lê termos de consulta aninhados em ordem de precedência. Por exemplo, `(amazon AND sports) NOT rainforest` lê `(amazon AND sports)` antes `NOT rainforest`.
- Intervalos: valores de data ou intervalo numérico. Os intervalos podem ser inclusivos, exclusivos ou ilimitados. Por exemplo, pesquise documentos que foram atualizados pela última vez entre 1º de janeiro de 2020 e 31 de dezembro de 2020, incluindo essas datas.
- Campos: usa um campo específico para limitar a pesquisa. Por exemplo, pesquisar documentos que tenham “Estados Unidos” no campo “localização”.
- Curingas: correspondem parcialmente a uma sequência de caracteres de texto. Por exemplo, `Cloud*` poderia corresponder CloudFormation. Amazon Kendra atualmente só suporta curingas finais.
- Citações exatas: corresponde exatamente a uma sequência de caracteres de texto. Por exemplo, documentos que contêm o `"Amazon Kendra" "pricing"`.

Use uma combinação de qualquer um dos operadores acima.

Observe que o uso excessivo de operadores ou consultas altamente complexas pode afetar a latência da consulta. Os curingas são algumas das operadoras mais caras em termos de latência. Uma regra geral é que quanto mais termos e operadores você usa, maior o impacto potencial na latência. Outros fatores que afetam a latência incluem o tamanho médio dos documentos indexados, o tamanho do seu índice, qualquer filtragem nos resultados da pesquisa e a carga geral do seu índice. Amazon Kendra

Booleano

Combine ou exclua palavras usando os operadores booleanos `AND`, `OR` e `NOT`.

Veja a seguir exemplos do uso de operadores booleanos.

`amazon AND sports`

Retorna resultados de pesquisa que contêm os termos “amazon” e “sports”v no texto, como Amazon Prime Video Sports ou outro conteúdo similar.

`sports OR recreation`

Retorna resultados da pesquisa que contêm os termos “esportes” ou “entretenimento”, ou ambos, no texto.

`amazon NOT rainforest`

Retorna resultados da pesquisa que contêm o termo “amazônia”, mas não o termo “floresta tropical” no texto. Isso é para pesquisar documentos sobre a empresa Amazon, não sobre a Floresta Amazônica.

Parênteses

Consulte palavras aninhadas em ordem de precedência usando parênteses. Os parênteses indicam Amazon Kendra como uma consulta deve ser lida.

Veja a seguir exemplos do uso de operadores de parênteses.

`(amazon AND sports) NOT rainforest`

Retorna documentos que contêm os termos “amazônia” e “esportes” no texto, mas não o termo “floresta tropical”. Isso serve para pesquisar vídeos esportivos do Amazon Prime ou outro conteúdo similar, não esportes de aventura na Floresta Amazônica. Os parênteses ajudam a indicar que o `amazon AND sports` deve ser lido antes de `NOT rainforest`. A consulta não deve ser lida como `amazon AND (sports NOT rainforest)`.

`(amazon AND (sports OR recreation)) NOT rainforest`

Retorna documentos que contêm os termos “esportes” ou “entretenimento”, ou ambos, e o termo “amazon”. Mas isso não inclui o termo “floresta tropical”. Isso serve para pesquisar vídeos esportivos ou entretenimentos no Amazon Prime, não esportes de aventura na Floresta Amazônica. Os parênteses ajudam a indicar que os `sports OR recreation` deve ser lido antes de combinar com 'amazon', que é lido antes de `NOT rainforest`. A consulta não deve ser lida como `amazon AND (sports OR (recreation NOT rainforest))`.

Intervalos

Use uma faixa de valores para filtrar os resultados da pesquisa. Você especifica um atributo e os valores do intervalo. Isso pode ser data ou tipo numérico.

Os intervalos de datas devem estar nos seguintes formatos:

- Epoch
- YYYY
- mm-AAAA
- dd-mm-AAAA
- dd-mm-AAAA'T'HH

Você também poderá especificar se deseja incluir ou excluir os valores mais baixos e mais altos do intervalo.

Veja a seguir exemplos do uso de operadores de alcance.

`_processed_date:>2019-12-31 AND _processed_date:<2021-01-01`

Retorna documentos que foram processados em 2020 – maiores que 31 de dezembro de 2019 e menos de 1º de janeiro de 2021.

`_processed_date:>=2020-01-01 AND _processed_date:<=2020-12-31`

Retorna documentos que foram processados em 2020 – maiores ou iguais a 1º de janeiro de 2020 e menores ou iguais a 31 de dezembro de 2020.

`_document_likes:<1`

Retorna documentos com zero curtidas ou sem feedback do usuário – menos de 1 curtida.

Especifique se um intervalo deve ser tratado como inclusivo ou exclusivo dos valores de intervalo fornecidos.

Inclusive

`_last_updated_at:[2020-01-01 TO 2020-12-31]`

Documentos de devolução atualizados pela última vez em 2020, inclusive os dias 1º de dezembro de 2020 e 31 de dezembro de 2020.

Exclusive

`_last_updated_at:{2019-12-31 TO 2021-01-01}`

Os documentos de devolução foram atualizados pela última vez em 2020, inclusive os dias 31 de dezembro de 2019 e 1º de janeiro de 2021.

Para intervalos ilimitados que não são inclusivos nem exclusivos, basta usar os operadores.**<** and **>**. Por exemplo, `_last_updated_at:>2019-12-31 AND _last_updated_at:<2021-01-01`.

Campos

Limite sua pesquisa para retornar somente documentos que atendam a um valor em um campo específico. O campo pode ser de qualquer tipo.

Veja a seguir exemplos do uso de operadores de contexto em nível de campo.

`status:"Incomplete" AND financial_year:2021`

Retorna documentos do exercício financeiro de 2021 com o status de incompleto.

`(sports OR recreation) AND country:"United States" AND level:"professional"`

Retorna documentos que discutem esportes profissionais ou entretenimento nos Estados Unidos.

Curingas

Amplie sua pesquisa para considerar variantes de palavras e frases usando o operador curinga. Isso é útil ao pesquisar variantes de nome. Amazon Kendra atualmente só suporta curingas finais. O número de caracteres de prefixo para um curinga à direita deve ser maior que dois.

Veja a seguir exemplos do uso de operadores curinga.

`Cloud*`

Retorna documentos que contêm variantes como CloudFormation CloudWatch e.

`kendra*aws`

Retorna documentos que contêm variantes, como kendra.amazonaws.

`kendra*aws*`

Retorna documentos que contêm variantes, como kendra.amazonaws.com

Cotações exatas

Use aspas para pesquisar uma correspondência exata de uma parte do texto.

Veja a seguir exemplos do uso de aspas.

`"Amazon Kendra" "pricing"`

Retorna documentos que contêm a frase do “Amazon Kendra” e o termo “preços”. Os documentos devem conter o “Amazon Kendra” e os “preços” para que os resultados sejam retornados.

`"Amazon Kendra" "pricing" cost`

Retorna documentos que contêm a frase do “Amazon Kendra” e o termo “preço” e, opcionalmente, o termo “custo”. Os documentos devem conter o “Amazon Kendra” e “preços” para retornar aos resultados, mas podem não necessariamente incluir “custo”.

Sintaxe de consulta inválida

Amazon Kendra emite um aviso se houver problemas com a sintaxe da consulta ou se ela não for suportada atualmente pelo Amazon Kendra. Para obter mais informações, consulte a Documentação da API para avisos de consulta.

As consultas a seguir são exemplos de sintaxe de consulta inválida.

`_last_updated_at:<2021-12-32`

Data inválida. O dia 32 não existe no calendário gregoriano, que é usado pelo Amazon Kendra.

`_view_count:ten`

Valor numérico inválido. Os dígitos devem ser usados para representar valores numéricos.

`nonExistentField:123`

Pesquisa de campo inválida. O campo deve existir para usar a pesquisa de campo.

```
Product:[A TO D]
```

Intervalo inválido. Valores numéricos ou datas devem ser usados para intervalos.

```
OR Hello
```

Booleano inválido. Os operadores devem ser usados com termos e colocados entre termos.

Pesquisando em idiomas

Pesquise documentos em um idioma com suporte. Você passa o código do idioma no AttributeFilterpara retornar documentos filtrados no idioma escolhido. Digite a consulta em um idioma com suporte.

Se você não especificar um idioma, Amazon Kendra consulta documentos em inglês por padrão. Para obter mais informações sobre os idiomas com suporte, inclusive seus códigos, consulte Adicionar documentos em outros idiomas além do inglês.

Para pesquisar documentos em um idioma com suporte pelo console, selecione seu índice e, em seguida, selecione a opção de pesquisar seu índice no menu de navegação. Escolha o idioma para o qual você deseja devolver os documentos selecionando as configurações de pesquisa e, em seguida, selecionando um idioma no menu suspenso Idioma.

Os exemplos a seguir mostram como pesquisar documentos em espanhol.

Para pesquisar um índice em espanhol no console

1. Faça login no AWS Management Console e abra o Amazon Kendra console em http://console.aws.amazon.com/kendra/.
2. No menu de navegação, escolha Índices e escolha seu índice.
3. No menu de navegação, escolha a opção de pesquisar no índice.
4. Nas configurações de pesquisa, selecione o menu suspenso Idiomas e escolha Espanhol.
5. Insira uma consulta na caixa de texto e pressione enter.
6. Amazon Kendra retorna os resultados da pesquisa em espanhol.

Para pesquisar um índice em espanhol usando a CLI, Python ou Java

- O exemplo a seguir pesquisa um índice em espanhol. Altere o valor `searchString` da sua consulta de pesquisa e o valor `indexID` do identificador do índice a ser pesquisado. O código do idioma para o espanhol é `es`. É possível substituí-lo pelo seu próprio código de 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");
        }
    }
}
```

Recuperar passagens

Use a API de Retrieve como um recuperador para sistemas de RAG (geração aumentada de recuperação).

Os sistemas de RAG usam inteligência artificial generativa para criar aplicativos de resposta a perguntas. Os sistemas de RAG consistem em um recuperador e LLM (grandes modelos de linguagem). Dada uma consulta, o recuperador identifica os trechos de texto mais relevantes de um acervo de documentos e os envia ao LLM para fornecer a resposta mais útil. Em seguida, o LLM analisa os trechos ou passagens de texto relevantes e gera uma resposta abrangente para a consulta.

A API de `Retrieve` analisa trechos de texto ou trechos chamados de passagens e retorna as principais passagens que são mais relevantes para a consulta.

Assim como a API de Query, a API de `Retrieve` também busca informações relevantes usando a pesquisa semântica. A pesquisa semântica leva em conta o contexto da consulta de pesquisa, além de todas as informações disponíveis dos documentos indexados. No entanto, por padrão, a API de `Query` retorna apenas trechos de até 100 palavras simbólicas. Com a API de `Retrieve`, recupere passagens mais longas de até 200 palavras simbólicas e até 100 passagens semanticamente relevantes. Isso não inclui respostas do tipo pergunta-resposta ou perguntas frequentes do índice. As passagens são trechos de texto que podem ser extraídos semanticamente de vários documentos e de várias partes do mesmo documento. Se, em casos extremos, seus documentos não produzirem nenhuma passagem usando a API de `Retrieve`, você também poderá usar a API de `Query` e seus tipos de respostas.

Você também poderá fazer o seguinte com a API de `Retrieve`:

- Substitua o aumento no nível do índice
- Filtrar com base nos campos ou atributos do documento
- Filtrar com base no acesso do usuário ou do grupo aos documentos
- Veja o bucket de pontuação de confiança para obter um resultado de passagem recuperado. O bucket de confiança fornece uma classificação relativa que indica o grau de confiança do Amazon Kendra de que a resposta é relevante para a consulta.

> **Note**
>
> No momento, os compartimentos de pontuação de confiança estão disponíveis somente em inglês.

Você também poderá incluir certos campos na resposta que podem fornecer informações adicionais úteis.

Atualmente, a API de `Retrieve` não oferece suporte a todos os recursos com suporte pela API de `Query`. Os seguintes recursos não têm suporte: consulta usando sintaxe de consulta avançada, correções ortográficas sugeridas para consultas, facetagem, sugestões de consultas para preencher automaticamente consultas de pesquisa e aprendizado incremental. Observe que nem todos os recursos se aplicam à `Retrieve` API. Todos os lançamentos futuros da `Retrieve` API serão documentados neste guia.

A API de `Retrieve` compartilha o número de unidades de capacidade de consulta que definida para seu índice. Para obter mais informações sobre o que está incluído em uma única unidade de capacidade e a capacidade base padrão de um índice, consulte Como ajustar a capacidade.

> **Note**
>
> Você não pode adicionar capacidade se estiver usando a Amazon Kendra Developer Edition; você só pode adicionar capacidade ao usar a Amazon Kendra Enterprise Edition. Para obter mais informações sobre o que está incluído nas edições do Desenvolvedor e Enterprise, consulte Edições do Amazon Kendra.

Veja a seguir um exemplo do uso da API de `Retrieve` para recuperar as 100 passagens mais relevantes dos documentos em um índice para a consulta de "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 em um índice

Você pode procurar documentos por seus atributos ou facetas sem precisar digitar uma consulta de pesquisa. Amazon Kendra O Navegar do índice pode ajudar seus usuários a descobrir documentos navegando livremente em um índice sem uma consulta específica em mente. Isso também ajuda seus usuários a navegar amplamente em um índice como ponto de partida em suas pesquisas.

O Navegar do índice só pode ser usado para pesquisar por atributo ou faceta do documento com um tipo de classificação. Não é possível pesquisar um índice inteiro usando o Navegar do índice. Se o texto da consulta estiver ausente, Amazon Kendra solicitará um filtro de atributo do documento ou uma faceta e um tipo de classificação.

Para permitir a navegação no índice usando a API de consulta, você deve incluir AttributeFilterou Facet e. SortingConfiguration Para permitir a navegação pelo índice no console, selecione seu índice em Índices no menu de navegação e, em seguida, selecione a opção de pesquisar seu índice. Na caixa de pesquisa, pressione a tecla Enter duas vezes. Selecione o menu suspenso Filtrar resultados da pesquisa para escolher um filtro e selecione o menu suspenso Classificar para escolher um tipo de classificação.

Veja a seguir um exemplo de navegação em um índice de documentos no idioma espanhol em ordem decrescente da data de criação do 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");
        }
    }
}
```

Apresentar resultados da pesquisa

Destaque determinados documentos nos resultados da pesquisa quando seus usuários fazem determinadas consultas. Isso ajuda a tornar os resultados mais visíveis e proeminentes para seus usuários. Os resultados em destaque são separados da lista normal de resultados e exibidos na parte superior da página de pesquisa. Experimente apresentar documentos diferentes para consultas diferentes ou garantir que determinados documentos tenham a visibilidade que merecem.

Você mapeia consultas específicas para documentos específicos para serem exibidas nos resultados. Se uma consulta contiver uma correspondência exata, um ou mais documentos específicos aparecerão nos resultados da pesquisa.

Por exemplo, especifique que, se seus usuários emitirem a consulta “novos produtos 2023", selecione os documentos intitulados “Novidades” e “Em breve” para aparecer na parte superior da página de resultados da pesquisa. Isso ajuda a garantir que esses documentos sobre novos produtos tenham a visibilidade que merecem.

Amazon Kendra não duplica os resultados da pesquisa se um resultado já estiver selecionado para ser exibido na parte superior da página de resultados da pesquisa. Um resultado em destaque não é novamente classificado como o primeiro resultado se já estiver acima de todos os outros resultados.

Para destacar determinados resultados, especifique uma correspondência exata de uma consulta de texto completo, não uma correspondência parcial de uma consulta usando uma palavra-chave ou frase contida em uma consulta. Por exemplo, se especificar apenas a consulta “Kendra” em um conjunto de resultados em destaque, consultas como: “Como a Kendra classifica semanticamente os resultados?” não renderizará os resultados em destaque. Os resultados em destaque são projetados para consultas específicas, em vez de consultas com escopo muito amplo. Amazon Kendra naturalmente lida com consultas de tipo de palavra-chave para classificar os documentos mais úteis nos resultados da pesquisa, evitando a apresentação excessiva de resultados com base em palavras-chave simples.

Se houver determinadas consultas que seus usuários usam com frequência, especifique essas consultas para obter resultados em destaque. Por exemplo, se analisar suas principais consultas usando a Análise do Amazon Kendra e encontrar essas consultas específicas, como: “Como o Kendra classifica semanticamente os resultados?” e 'pesquisa semântica kendra', são usadas com frequência, então essas consultas podem ser úteis para especificar o documento intitulado 'pesquisa 101'.Amazon Kendra

Amazon Kendra trata as consultas de resultados em destaque como não diferenciando maiúsculas de minúsculas. Amazon Kendra converte uma consulta em minúsculas e substitui os caracteres de espaço em branco à direita por um único espaço. Amazon Kendra corresponde a todos os outros caracteres como estão quando você especifica suas consultas para resultados em destaque.

Você cria um conjunto de resultados em destaque que você mapeia para determinadas consultas usando a CreateFeaturedResultsSetAPI. Se usa o console, seleciona seu índice e, em seguida, seleciona Resultados em destaque no menu de navegação para criar um conjunto de resultados em destaque. Crie até 50 conjuntos de resultados em destaque por índice, até quatro documentos a serem apresentados por conjunto e até 49 textos de consulta por conjunto de resultados em destaque. Solicite o aumento desses limites entrando em contato com o Suporte.

Selecionar o mesmo documento em vários conjuntos de resultados em destaque. No entanto, você não deve usar o mesmo texto de consulta de correspondência exata em vários conjuntos. As consultas que você especifica para resultados em destaque devem ser exclusivas por conjunto de resultados em destaque para cada índice.

Organize a ordem dos documentos ao selecionar até quatro documentos em destaque. Se usa a API, a ordem em que você lista os documentos em destaque é a mesma exibida nos resultados em destaque. Se usa o console, pode simplesmente arrastar e soltar a ordem dos documentos ao selecionar documentos para serem exibidos nos resultados.

O controle de acesso, em que determinados usuários e grupos têm acesso a determinados documentos e outros não, ainda é respeitado ao configurar os resultados em destaque. Isso também vale para a filtragem de contexto do usuário. Por exemplo, o usuário A pertence ao grupo de empresas “Estagiários”, que não deve acessar documentos sobre segredos da empresa. Se o usuário A inserir uma consulta que contenha um documento secreto da empresa, o usuário A não verá esse documento em destaque nos resultados. Isso também vale para qualquer outro resultado na página de resultados da pesquisa. Você também poderá usar tags para controlar o acesso a um conjunto de resultados em destaque, que é um recurso do Amazon Kendra para o qual você controla o acesso.

Veja a seguir um exemplo de criação de um conjunto de resultados em destaque com as consultas “novos produtos 2023", “novos produtos disponíveis” mapeadas para os documentos intitulados “Nocidades” (doc-id-1) e “Em breve” (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.")
```

Pesquisa tabular de HTML

O recurso de pesquisa tabular do Amazon Kendra pode pesquisar e extrair respostas de tabelas incorporadas em documentos HTML. Ao pesquisar seu índice, Amazon Kendra inclua um trecho de uma tabela se for relevante para a consulta e forneça informações úteis.

Amazon Kendra examina todas as informações no corpo do texto de um documento, incluindo informações úteis em tabelas. Por exemplo, um índice contém relatórios comerciais com tabelas sobre custos operacionais, receitas e outras informações financeiras. Para a consulta, “qual é o custo operacional anual de 2020-2022?” , Amazon Kendra pode retornar um trecho de uma tabela que contém as colunas relevantes “Operações (milhões de dólares)” e “Exercício financeiro” e linhas da tabela contendo valores de renda para 2020, 2021 e 2022. O trecho da tabela é incluído no resultado, junto com o título do documento, um link para o documento completo e quaisquer outros campos do documento que você escolher incluir.

Trechos da tabela podem ser exibidos nos resultados da pesquisa, independentemente de as informações serem encontradas em uma célula de uma tabela ou em várias células. Por exemplo, Amazon Kendra pode exibir um trecho de tabela personalizado para cada um desses tipos de consultas:

- “cartão de crédito com maior taxa de juros em 2020"
- “cartão de crédito com maior taxa de juros de 2020 a 2022"
- “os 3 cartões de crédito com maior taxa de juros em 2020 a 2022"
- “cartões de crédito com taxas de juros inferiores a 10%”
- “todos os cartões de crédito com juros baixos disponíveis”

Amazon Kendra destaca a célula ou células da tabela que são mais relevantes para a consulta. As células mais relevantes com suas linhas, colunas e nomes de colunas correspondentes são exibidas no resultado da pesquisa. O trecho da tabela exibe até cinco colunas e três linhas, dependendo de quantas células da tabela são relevantes para a consulta e quantas colunas estão disponíveis na tabela original. A célula mais relevante é exibida no trecho da tabela, junto com as próximas células mais relevantes.

A resposta inclui o bucket de confiança (`MEDIUM`, `HIGH` e `VERY_HIGH`) para mostrar a relevância da resposta da tabela para a consulta. Se o valor de uma célula da tabela for `VERY_HIGH` em confidência, ele se tornará a “resposta principal” e será destacado. Para valores de células de tabela que são `HIGH` confiáveis, eles são destacados. Para valores de células de tabela que são

MEDIUM confiáveis, eles não são destacados. A confiança geral da resposta da tabela é retornada na resposta. Por exemplo, se uma tabela contém principalmente células da tabela com confiança HIGH, a confiança geral retornada na resposta para a resposta da tabela é confiança HIGH.

Por padrão, as tabelas não recebem um nível mais alto de importância ou mais peso do que outros componentes de um documento. Em um documento, se uma tabela for apenas ligeiramente relevante para uma consulta, mas houver um parágrafo altamente relevante, Amazon Kendra retornará um trecho do parágrafo. Os resultados da pesquisa exibem o conteúdo que fornece a melhor resposta possível e as informações mais úteis, no mesmo documento ou em outros documentos. Se a confiança de uma tabela ficar abaixo da confiança MEDIUM, o trecho da tabela não será retornado na resposta.

Para usar a pesquisa tabular em um índice existente, reindexe seu conteúdo.

Amazon Kendra a pesquisa tabular suporta sinônimos (incluindo sinônimos personalizados). Amazon Kendra só oferece suporte a documentos em inglês com tabelas HTML que estejam dentro da tag da tabela.

O exemplo a seguir mostra um trecho da tabela incluído no resultado da consulta. Para visualizar um exemplo de JSON com respostas de consulta, incluindo trechos de tabelas, consulte Respostas e tipos de consulta.

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");
        }
    }
}
```

Sugestões de consulta

As sugestões de consulta do Amazon Kendra podem ajudar seus usuários a digitar suas consultas de pesquisa com mais rapidez e orientar suas pesquisas.

Amazon Kendra sugere consultas relevantes para seus usuários com base em uma das seguintes opções:

- Consultas populares no histórico de consultas ou no log de consultas
- O conteúdo dos campos/atributos do documento

Defina sua preferência para usar o histórico de consultas ou os campos do documento definindo os `SuggestionTypes` como `QUERY` ou `DOCUMENT_ATTRIBUTES` e chamando GetQuerySuggestions. Por padrão, Amazon Kendra usa o histórico de consultas para basear sugestões. Se o histórico da consulta e os campos do documento estiverem ativados quando você ligar UpdateQuerySuggestionsConfige você não tiver definido sua `SuggestionTypes` preferência para usar os campos do documento, Amazon Kendra use o histórico da consulta.

Se usa o console, pode basear as sugestões de consulta no histórico da consulta ou nos campos do documento. Primeiro, selecione seu índice e, em seguida, selecione Sugestões de consulta em Enriquecimentos no menu de navegação. Em seguida, selecione Configurar sugestões de consulta. Depois de configurar as sugestões de consulta, você é direcionado para um console de pesquisa onde pode selecionar os campos Histórico da consulta ou Documento no painel direito e inserir uma consulta de pesquisa na barra de pesquisa.

Por padrão, as sugestões de consulta usando o histórico de consultas e os campos do documento são ativadas sem custo adicional. Desative esses tipos de sugestões de consulta a qualquer momento usando a API de `UpdateQuerySuggestionsConfig`. Para desativar as sugestões de consulta com base no histórico de consultas, defina `Mode` como `DISABLED` ao chamar `UpdateQuerySuggestionsConfig`. Para desativar as sugestões de consulta com base nos campos do documento, defina `AttributeSuggestionsMode` como `INACTIVE` na configuração

dos campos do documento e, em seguida, chame `UpdateQuerySuggestionsConfig>`. Se usa o console, pode desativar as sugestões de consulta nas Configurações de sugestões de consulta.

As sugestões de consulta não diferenciam maiúsculas de minúsculas. Amazon Kendra converte o prefixo da consulta e a consulta sugerida em minúsculas, ignora todas as aspas simples e duplas e substitui vários caracteres de espaço em branco por um único espaço. Amazon Kendra corresponde a todos os outros caracteres especiais do jeito que estão. Amazon Kendra não mostra nenhuma sugestão se um usuário digitar menos de dois caracteres ou mais de 60 caracteres.

Tópicos

- Sugestões de consulta usando o histórico de consultas
- Sugestões de consulta usando campos do documento
- Bloqueie determinadas consultas ou conteúdos de campos do documento de sugestões

Sugestões de consulta usando o histórico de consultas

Tópicos

- Configurações para selecionar consultas para sugestões
- Sugestões claras, mantendo o histórico de consultas
- Não há sugestões disponíveis

Você pode optar por sugerir consultas relevantes para seus usuários com base em consultas populares no histórico de consultas ou no registro de consultas. Amazon Kendra usa todas as consultas que seus usuários pesquisam e aprendem com essas consultas para fazer sugestões aos usuários. Amazon Kendra sugere consultas populares aos usuários quando eles começam a digitar a consulta. Amazon Kendra sugere uma consulta se o prefixo ou os primeiros caracteres da consulta corresponderem ao que o usuário começa a digitar como consulta.

Por exemplo, um usuário começa a digitar a consulta “próximos eventos”. O Amazon Kendra aprendeu com o histórico de consultas que muitos usuários pesquisaram “próximos eventos 2050" várias vezes. O usuário observa os “próximos eventos 2050" aparecerem diretamente abaixo da barra de pesquisa, preenchendo automaticamente a consulta de pesquisa. O usuário seleciona essa sugestão de consulta e o documento “Novos eventos: o que está acontecendo em 2050" é retornado nos resultados da pesquisa.

Você pode especificar como Amazon Kendra seleciona consultas qualificadas para sugerir aos seus usuários. Por exemplo, você pode especificar que uma sugestão de consulta deve ter sido pesquisada por pelo menos 10 usuários exclusivos (o padrão é três), ter sido pesquisada nos últimos 30 dias e não conter nenhuma palavra ou frase da sua lista de bloqueio. Amazon Kendra exige que uma consulta tenha pelo menos um resultado de pesquisa e contenha pelo menos uma palavra com mais de quatro caracteres.

Configurações para selecionar consultas para sugestões

Defina as seguintes configurações para selecionar consultas para sugestões usando a API de UpdateQuerySuggestionsConfig:

- Modo – As sugestões de consulta usando o histórico de consultas são `ENABLED` ou `LEARN_ONLY`. O Amazon Kendra ativa as sugestões de consulta por padrão. O modo `LEARN_ONLY` desativa as sugestões de consulta. Se desativado, Amazon Kendra continua aprendendo sugestões, mas não faz sugestões de consulta aos usuários.
- Janela de tempo do log de consultas – Quão recentes são suas consultas na janela de tempo do log de consultas. A janela de tempo é um valor inteiro para o número de dias do dia atual até os dias anteriores.
- Consultas sem informações do usuário – Defina `TRUE` para incluir todas as consultas ou `FALSE` para incluir somente consultas com informações do usuário. Use essa configuração se seu aplicativo de pesquisa incluir informações do usuário, como a ID do usuário, quando um usuário fizer uma consulta. Por padrão, essa configuração não filtra as consultas se não houver informações específicas do usuário associadas às consultas. No entanto, você poderá usar essa configuração para fazer sugestões somente com base em consultas que incluam informações do usuário.
- Usuários exclusivos – O número mínimo de usuários exclusivos que precisam pesquisar uma consulta para que a consulta se qualifique para sugerir aos seus usuários. Esse número é um valor inteiro.
- Contagem de consultas – O número mínimo de vezes que uma consulta deve ser pesquisada para que ela esteja qualificada para ser sugerida aos seus usuários. Esse número é um valor inteiro.

Essas configurações afetam a forma como as consultas são selecionadas como consultas populares para sugerir aos seus usuários. A forma como você ajusta suas configurações dependerá de suas necessidades específicas, por exemplo:

- Se seus usuários costumam pesquisar uma vez por mês, em média, você poderá definir o número de dias na janela de tempo do log de consultas para 30 dias. Ao usar essa configuração, você captura a maioria das consultas recentes de seus usuários antes que elas se tornem desatualizadas na janela de tempo.
- Se apenas um pequeno número de suas consultas incluir informações do usuário e você não quiser sugerir consultas com base em uma amostra pequena, defina as consultas para incluir todos os usuários.
- Se definir consultas populares como sendo pesquisadas por pelo menos 10 usuários exclusivos e pesquisadas pelo menos 100 vezes, defina os usuários exclusivos como 10 e a contagem de consultas como 100.

> **Warning**
>
> Suas alterações nas configurações podem não entrar em vigor imediatamente. Acompanhe as alterações nas configurações usando a API de DescribeQuerySuggestionsConfig. O tempo para que suas configurações atualizadas entrem em vigor depende das atualizações que você fizer e do número de consultas de pesquisa em seu índice. O Amazon Kendra atualiza automaticamente as sugestões a cada 24 horas, depois de alterar uma configuração ou depois de aplicar uma lista de bloqueio.

CLI

Para recuperar sugestões 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 atualizar as sugestões de consulta

Por exemplo, para alterar a janela de tempo do log de consultas e o número mínimo de vezes que uma consulta deve ser pesquisada:

```
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 sugestões 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 atualizar as sugestões de consulta

Por exemplo, para alterar a janela de tempo do log de consultas e o número mínimo de vezes que uma consulta deve ser pesquisada:

```
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.")
```

Sugestões claras, mantendo o histórico de consultas

Esclareça as sugestões de consulta usando a API de ClearQuerySuggestions. A limpeza de sugestões exclui somente as sugestões de consulta existentes, não as consultas no histórico de consultas. Quando você limpa as sugestões, Amazon Kendra aprende novas sugestões com base nas novas consultas adicionadas ao registro de consultas a partir do momento em que você apagou as sugestões.

CLI

Para limpar as sugestões de consulta

```
aws  kendra clear-query-suggestions \
 --index-id index-id
```

Python

Para limpar as sugestões 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.")
```

Não há sugestões disponíveis

Se não vê sugestões de consulta, talvez por um dos seguintes motivos:

- Não há consultas suficientes em seu índice com as quais você Amazon Kendra possa aprender.
- Suas configurações de sugestões de consulta são muito rígidas, fazendo com que a maioria das consultas seja filtrada das sugestões.
- Você apagou sugestões recentemente e Amazon Kendra ainda precisa de tempo para que novas consultas se acumulem a fim de aprender novas sugestões.

Verifique suas configurações atuais usando a API de DescribeQuerySuggestionsConfig.

Sugestões de consulta usando campos do documento

Tópicos

- Configurações para selecionar campos para sugestões
- Controle do usuário nos campos do documento

Opte por sugerir consultas relevantes para seus usuários com base no conteúdo dos campos do documento. Em vez de usar o histórico de consultas para sugerir outras consultas relevantes populares, você pode usar as informações contidas em um campo do documento que são úteis para o preenchimento automático da consulta. Amazon Kendra procura conteúdo relevante em

campos definidos como `Suggestable` e que estejam estreitamente alinhados com a consulta do seu usuário. Em seguida, Amazon Kendra sugere esse conteúdo para seu usuário quando ele começar a digitar sua consulta.

Por exemplo, se você especificar o campo de título no qual basear as sugestões e um usuário começar a digitar a consulta “Como a Amazon Ken... ', o título mais relevante 'Como Amazon Kendra funciona' pode ser sugerido para completar automaticamente a pesquisa. O usuário vê “Como Amazon Kendra funciona” aparecer diretamente abaixo da barra de pesquisa, preenchendo automaticamente a consulta de pesquisa. O usuário seleciona essa sugestão de consulta e o documento “Como Amazon Kendra funciona” é retornado nos resultados da pesquisa.

Use o conteúdo de qualquer campo de `String` e tipo de documento `StringList`, para sugerir uma consulta definindo o campo `Suggestable` como parte da configuração de seus campos para sugestões de consulta. Você também poderá usar uma lista de bloqueios para que os campos de documentos sugeridos que contenham determinadas palavras ou frases não sejam exibidos aos seus usuários. Use uma lista de bloqueio. A lista de bloqueios se aplica se você definir sugestões de consulta para usar o histórico de consultas ou os campos do documento.

Configurações para selecionar campos para sugestões

Defina as seguintes configurações para selecionar campos do documento para sugestões de uso do AttributeSuggestionsConfig e chamada da API deUpdateQuerySuggestionsConfig para atualizar as configurações no nível do índice:

- Modo de sugestões de campos/atributos — As sugestões de consulta usando campos de documentos são ou. `ACTIVE` `INACTIVE` Amazon Kendra ativa as sugestões de consulta por padrão.
- Campos/atributos sugestionáveis – Os nomes de campo ou chaves de campo nos quais basear sugestões. Esses campos devem ser definidos como `TRUE` para `Suggestable`, como parte da configuração dos campos. Substitua a configuração dos campos no nível da consulta e, ao mesmo tempo, manter a configuração no nível do índice. Use a GetQuerySuggestionsAPI para alterar `AttributeSuggestionConfig` no nível da consulta. Essa configuração no nível da consulta pode ser útil para experimentar rapidamente o uso de diferentes campos do documento sem precisar atualizar a configuração no nível do índice.
- Campos/atributos adicionais – Os campos adicionais que você deseja incluir na resposta para uma sugestão de consulta. Esses campos são usados para fornecer informações adicionais na resposta; no entanto, eles não são usados para basear sugestões.

> **Warning**
>
> Suas alterações nas configurações podem não entrar em vigor imediatamente. Acompanhe as alterações nas configurações usando a API de DescribeQuerySuggestionsConfig. O tempo para que suas configurações atualizadas entrem em vigor depende das atualizações que você fizer. Amazon Kendra atualiza automaticamente as sugestões a cada 24 horas, depois de alterar uma configuração ou depois de aplicar uma lista de bloqueio.

CLI

Para recuperar sugestões de consulta e substituir a configuração dos campos do documento no nível da consulta, em vez de precisar alterar a configuração no nível do 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 atualizar as sugestões de consulta

Por exemplo, para alterar a configuração dos campos do documento no nível do 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 sugestões de consulta e substituir a configuração dos campos do documento no nível da consulta, em vez de precisar alterar a configuração no nível do 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 atualizar as sugestões de consulta

Por exemplo, para alterar a configuração dos campos do documento no nível do 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.")
```

Controle do usuário nos campos do documento

Aplique a filtragem de contexto do usuário aos campos do documento nos quais deseja basear as sugestões de consulta. Isso filtra as informações do campo do documento com base no acesso do usuário ou do grupo aos documentos. Por exemplo, um estagiário pesquisa no portal da empresa e não tem acesso a um documento ultrassecreto da empresa. Portanto, as consultas sugeridas com base no título do documento ultrassecreto, ou em qualquer outro campo sugestionável, não são mostradas ao estagiário.

Indexe seus documentos com uma lista de controle de acesso (ACL), definindo quais usuários e grupos têm acesso a quais documentos. Em seguida, aplique a filtragem de contexto do usuário aos campos do seu documento para sugestões de consulta. A filtragem de contexto de usuário atualmente definida para seu índice é a mesma filtragem de contexto de usuário aplicada à configuração dos campos do documento para sugestões de consulta. A filtragem de contexto do usuário faz parte da configuração dos campos do documento. Você usa o AttributeSuggestionsGetConfig e chama o GetQuerySuggestions.

Bloqueie determinadas consultas ou conteúdos de campos do documento de sugestões

Uma lista de bloqueio Amazon Kendra impede de sugerir determinadas consultas aos seus usuários. Uma lista de bloqueio é uma lista de palavras ou frases que você deseja excluir das sugestões de consulta. Amazon Kendra exclui consultas que contêm uma correspondência exata das palavras ou frases na lista de bloqueio.

Use uma lista de bloqueio para se proteger contra palavras ou frases ofensivas que geralmente aparecem no histórico de consultas ou nos campos do documento e que o Amazon Kendra pode selecionar como sugestões. Uma lista de bloqueio também pode Amazon Kendra impedir a sugestão de consultas que contenham informações que não estão prontas para serem divulgadas ou anunciadas publicamente. Por exemplo, seus usuários frequentemente consultam sobre uma próxima versão de um possível novo produto. No entanto, você não quer sugerir o produto porque não está pronto para lançá-lo. Você poderá bloquear as consultas que contêm o nome do produto e as informações do produto das sugestões.

Crie uma lista de bloqueio para consultas usando a API de CreateQuerySuggestionsBlockList. Você coloca cada palavra ou frase de bloco em uma linha separada em um arquivo de texto. Em seguida, você carrega o arquivo de texto no seu bucket do Amazon S3 e fornece o caminho ou a localização do arquivo. Amazon S3 Amazon Kendra atualmente suporta a criação de apenas uma lista de bloqueio.

Você pode substituir o arquivo de texto das palavras e frases bloqueadas no seu Amazon S3 bucket. Para atualizar a lista de bloqueios Amazon Kendra, use a UpdateQuerySuggestionsBlockListAPI.

Use a API de DescribeQuerySuggestionsBlockList para obter o status da sua lista de bloqueios. O `DescribeQuerySuggestionsBlockList` também poderá fornecer outras informações úteis, como:

- Quando sua lista de bloqueios foi atualizada pela última vez
- Quantas palavras ou frases estão na sua lista de bloqueio atual
- Mensagens de erro úteis ao criar uma lista de bloqueio

Você também poderá usar a API de ListQuerySuggestionsBlockLists para obter uma lista dos resumos da lista de bloqueios para um índice.

Para excluir sua lista de bloqueios, use a DeleteQuerySuggestionsBlockListAPI.

Suas atualizações na lista de bloqueios podem não entrar em vigor imediatamente. Acompanhar as atualizações usando a API `DescribeQuerySuggestionsBlockList`.

CLI

Para criar uma lista de bloqueios

```
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 atualizar uma lista de bloqueios

```
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 excluir uma lista de bloqueio

```
aws kendra delete-query-suggestions-block-list \
 --index-id index-id \
 --id block-list-id
```

Python

Para criar uma lista de bloqueios

```
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 atualizar uma lista de bloqueios

```
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 excluir uma lista de bloqueio

```
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.")
```

Corretor ortográfico de consulta

O Corretor ortográfico do Amazon Kendra sugere correções ortográficas para uma consulta. Isso pode ajudar você a reduzir ao mínimo as ocorrências de zero resultados de pesquisa e a retornar resultados relevantes. Seus usuários podem receber zero resultados de pesquisa de consultas com ortografia incorreta, sem resultados correspondentes ou sem documentos retornados. Ou seus usuários podem receber resultados de pesquisa irrelevantes de consultas com erros ortográficos.

O corretor ortográfico foi projetado para sugerir correções para palavras com erros ortográficos com base nas palavras que aparecem em seus documentos indexados e na proximidade com que uma palavra corrigida corresponde a uma palavra com ortografia incorreta. Por exemplo, se a palavra

“statements” (“demonstrações”) aparecer em seus documentos indexados, isso pode ser semelhante à palavra incorreta “statments” (“colocações”) na consulta “statments (colocações”) financeiras de fim de ano”.

O verificador ortográfico retorna as palavras pretendidas ou corrigidas que substituem as palavras com erros ortográficos no texto da consulta original. Por exemplo, “implantar a pesquisa do kendre” pode retornar “implantar a pesquisa Kendra”. Você também pode usar os locais de deslocamento fornecidos na API para destacar ou colocar em itálico as palavras corrigidas retornadas em uma consulta em seu aplicativo de front-end. No console, as palavras corrigidas são destacadas ou em itálico por padrão. Por exemplo, “implantar a pesquisa Kendra”.

Para termos específicos de negócios ou especializados que aparecem em seus documentos indexados, o verificador ortográfico não interpreta esses termos como erros de ortografia na consulta. Por exemplo, 'amazon macie' não é corrigido para 'amazon mace'.

Para palavras com hífen, como “fim-de-ano”, o corretor ortográfico as trata como palavras individuais para sugerir correções para essas palavras. Por exemplo, a correção sugerida para “final de ano” pode ser “fim de ano”.

Para os `DOCUMENT` e tipos de resposta de consulta de `QUESTION_ANSWER`, o corretor ortográfico sugere correções para palavras com ortografia incorreta com base nas palavras no corpo do documento. O corpo do documento é mais confiável do que o título para sugerir correções que se aproximam das palavras escritas incorretamente. Para tipos de resposta de consulta de `ANSWER`, o corretor ortográfico sugere correções com base nas palavras do documento padrão de perguntas e respostas em seu índice.

Você pode ativar o Corretor Ortográfico usando o SpellCorrectionConfigurationobjeto. Você define `IncludeQuerySpellCheckSuggestions` como `TRUE`. O Corretor ortográfico é ativado por padrão no console. Ele é incorporado ao console por padrão.

O Corretor ortográfico também pode sugerir correções ortográficas para consultas em vários idiomas, não apenas em inglês. Para obter uma lista dos idiomas com suporte pelo Corretor Ortográfico, consulte os idiomas Suportados pelo Amazon Kendra.

Usar o corretor ortográfico de consulta com limites padrão

O corretor ortográfico foi projetado com certos padrões ou limites. A seguir está uma lista dos limites atuais que se aplicam quando você ativa sugestões de correção ortográfica.

- As correções ortográficas sugeridas não podem ser retornadas para palavras com menos de três caracteres ou mais de 30 caracteres. Para permitir mais de 30 caracteres ou menos de três caracteres, entre em contato com o Suporte.
- As correções ortográficas sugeridas não podem restringir as sugestões com base no controle de acesso do usuário ou na sua lista de controle de acesso para filtragem de contexto do usuário. As correções ortográficas são baseadas em todas as palavras em seus documentos indexados, sejam elas restritas a determinados usuários ou não. Se quiser evitar que certas palavras apareçam nas correções ortográficas sugeridas para consultas, não ative o `SpellCorrectionConfiguration`.
- As correções ortográficas sugeridas não podem ser retornadas para palavras que incluam números. Por exemplo, "como 2 não br8k ubun2”.
- As correções ortográficas sugeridas não podem usar palavras que não aparecem em seus documentos indexados.
- As correções ortográficas sugeridas não podem usar palavras com menos de 0,01% de frequência em seus documentos indexados. Para alterar o limite de 0,01%, entre em contato com o Suporte.

Filtragem e pesquisa de facetas

Melhore os resultados da pesquisa ou a resposta da API de consulta usando filtros. Os filtros restringem os documentos na resposta aos que se aplicam diretamente à consulta. Para criar sugestões de pesquisa facetada, use a lógica booleana para filtrar atributos específicos do documento da resposta ou documentos que não correspondem a critérios específicos. Especifique as facetas usando o parâmetro `Facets` na API de `Query`.

Para pesquisar documentos com os quais você indexou Amazon Lex, use Amazon Kendra a AMAZON. KendraSearchIntent. Para ver um exemplo de configuração Amazon Kendra com Amazon Lex, consulte Criação de um bot de perguntas frequentes para um Amazon Kendra índice. Você também pode fornecer um filtro para a resposta usando AttributeFilter. Esse é o filtro de consulta em JSON durante a configuração do `AMAZON.KendraSearchIntent`. Para fornecer um filtro de atributo ao configurar uma intenção de pesquisa no console, acesse o editor de intenção e escolha a consulta no Amazon Kendra para fornecer um filtro de consulta em JSON. Para obter mais informações sobre o `AMAZON.KendraSearchIntent`, consulte o guia de documentação do Amazon Lex.

Facetas

As facetas são visualizações com escopo de um conjunto de resultados de pesquisa. Por exemplo, forneça resultados de pesquisa para cidades em todo o mundo, onde os documentos são filtrados por uma cidade específica à qual estão associados. Ou crie facetas para exibir os resultados de um autor específico.

Use um atributo de documento ou campo de metadados associado a um documento como uma faceta para que seus usuários possam pesquisar por categorias ou valores dentro dessa faceta. Você também poderá exibir facetas aninhadas nos resultados da pesquisa para que seus usuários possam pesquisar não apenas por uma categoria ou campo, mas também por uma subcategoria ou subcampo.

O exemplo a seguir mostra como obter informações de faceta do atributo personalizado “Cidade”.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        Facets = [
            {
                "DocumentAttributeKey" : "City"
            }
        ]
        )
```

Use facetas aninhadas para restringir ainda mais a pesquisa. Por exemplo, o atributo do documento ou a faceta “Cidade” inclui um valor chamado “Seattle”. Além disso, o atributo ou faceta do documento "CityRegion" inclui os valores “Norte” e “Sul” para documentos atribuídos a “Seattle”. Exiba facetas aninhadas com suas contagens nos resultados da pesquisa para que os documentos possam ser pesquisados não apenas por cidade, mas também por uma região dentro de uma cidade.

Observe que facetas aninhadas podem afetar a latência da consulta. Uma regra geral é que quanto mais facetas aninhadas você usa, maior o impacto potencial na latência. Outros fatores que afetam a latência incluem o tamanho médio dos documentos indexados, o tamanho do seu índice, consultas altamente complexas e a carga geral do seu índice do Amazon Kendra .

O exemplo a seguir mostra como obter informações de facetas para o atributo personalizado CityRegion "", como uma faceta aninhada em “Cidade”.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        Facets = [
            {
                "DocumentAttributeKey" : "City",
                "Facets": [
                    {
                        "DocumentAttributeKey" : "CityRegion"
                    }
                ]
            }
        ]
        )
```

As informações de facetas, como a contagem de documentos, são retornadas na matriz de respostas dos `FacetResults`. Você usa o conteúdo para exibir sugestões de pesquisa facetada em seu aplicativo. Por exemplo, se o atributo do documento “Cidade” contiver a cidade à qual uma pesquisa pode ser aplicada, use essas informações para exibir uma lista de pesquisas de cidades. Os usuários podem escolher uma cidade para filtrar os resultados da pesquisa. Para fazer a pesquisa facetada, chame a API de consulta e use o atributo de documento escolhido para filtrar os resultados.

Exiba até 10 valores de faceta por faceta para uma consulta e somente uma faceta aninhada dentro de uma faceta. Se você deseja aumentar esses limites, entre em contato com o Suporte. Se você quiser limitar o número de valores de faceta por faceta a menos de 10, especifique isso no objeto `Facet`.

O exemplo de resposta do JSON a seguir mostra facetas com escopo definido no atributo do documento “Cidade”. A resposta inclui a contagem de documentos para o valor da faceta.

```
{
    'FacetResults': [
        {
            'DocumentAttributeKey': 'City',
            'DocumentAttributeValueCountPairs': [
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Dubai'
                    }
```

```
                },
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Seattle'
                    }
                },
                {
                    'Count': 1,
                    'DocumentAttributeValue': {
                        'StringValue': 'Paris'
                    }
                }
            ]
        }
    ]
```

Você também poderá exibir informações de facetas de uma faceta aninhada, como uma região dentro de uma cidade, para filtrar ainda mais os resultados da pesquisa.

O exemplo de resposta JSON a seguir mostra facetas com escopo no atributo de documento "CityRegion", como uma faceta aninhada em “Cidade”. A resposta inclui a contagem de documentos para os valores das facetas aninhadas.

```
{
    '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'
                                    }
                                }
                            ]
                        }
                    ]
                }
            }
        ]
}
```

Ao usar um campo de lista de sequência de caracteres para criar facetas, os resultados de facetas retornados são baseados no conteúdo da lista de sequência de caracteres. Por exemplo, se você tiver um campo de lista de sequência de caracteres que contém dois itens, um com a lista “bassê”, “cachorro salsicha” e outro com o valor “husky”, você receberá três facetas do `FacetResults`.

Para ter mais informações, consulte Respostas de consulta e tipos de resposta.

Usando atributos do documento para filtrar os resultados da pesquisa

Por padrão, a `Query` retorna todos os resultados da pesquisa. Para filtrar as respostas, realize operações lógicas nos atributos do documento. Por exemplo, se quiser documentos apenas para uma cidade específica, poderá filtrar os atributos personalizados do documento “Cidade” e “Estado”. Você usa AttributeFilterpara criar uma operação booleana nos filtros que você fornece.

A maioria dos atributos pode ser usada para filtrar respostas para todos os tipos de resposta. No entanto, o atributo `_excerpt_page_number` só é aplicável aos tipos de resposta `ANSWER` ao filtrar respostas.

O exemplo a seguir mostra como realizar uma operação E lógica filtrando em uma cidade específica, Seattle, e estado, Washington.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {'AndAllFilters':
            [
                {"EqualsTo": {"Key": "City","Value": {"StringValue": "Seattle"}}},
```

```
                {"EqualsTo": {"Key": "State","Value": {"StringValue": "Washington"}}}
            ]
            }
        )
```

O exemplo a seguir mostra como realizar uma operação OU lógica para quando qualquer uma das chaves `Fileformat`, `Author` ou `SourceURI` corresponder aos 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"}}}
            ]
            }
        )
```

Para campos de `StringList`, use os filtros de atributos `ContainsAny` ou `ContainsAll` para retornar documentos com a sequência de caracteres especificada. O exemplo a seguir mostra como retornar todos os documentos que têm os valores “Seattle” ou “Portland” em seu atributo de `Locations` personalizado.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {
                "ContainsAny": { "Key": "Locations", "Value": { "StringListValue":
 [ "Seattle", "Portland"] }}
            }
        )
```

Filtrando os atributos de cada documento nos resultados da pesquisa

Amazon Kendra retorna os atributos do documento para cada documento nos resultados da pesquisa. Filtre determinados atributos do documento que deseja incluir na resposta como parte dos

resultados da pesquisa. Por padrão, todos os atributos do documento atribuídos a um documento são retornados na resposta.

No exemplo a seguir, somente os atributos `_source_uri` e `_author` do documento são incluídos na resposta de um documento.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        RequestedDocumentAttributes = ["_source_uri", "_author"]
        )
```

Filtragem no contexto do usuário

Filtre os resultados da pesquisa de um usuário com base no acesso do usuário ou do grupo aos documentos. Use um token de usuário, ID de usuário ou atributo de usuário para filtrar documentos. O Amazon Kendra também pode mapear usuários para seus grupos. Opte por usar o AWS IAM Identity Center como sua armazenamento/fonte de identidade.

A filtragem de contexto do usuário é um tipo de pesquisa personalizada com o benefício de controlar o acesso aos documentos. Por exemplo, nem todas as equipes que pesquisam informações no portal da empresa devem acessar documentos ultrassecretos da empresa, nem esses documentos são relevantes para todos os usuários. Somente usuários específicos ou grupos de equipes com acesso a documentos ultrassecretos devem ver esses documentos nos resultados de pesquisa.

Quando um documento é indexado Amazon Kendra, uma lista de controle de acesso (ACL) correspondente é ingerida para a maioria dos documentos. A ACL especifica quais nomes de usuário e grupos têm acesso permitido ou negado ao documento. Documentos sem uma ACL são documentos públicos.

Amazon Kendra pode extrair as informações do usuário ou do grupo associadas a cada documento para a maioria das fontes de dados. Por exemplo, um documento no Quip pode incluir uma lista de “compartilhamento” de usuários selecionados que têm acesso ao documento. Use um bucket do S3 como fonte de dados, forneça um arquivo JSON para sua ACL e inclua o caminho do S3 para esse arquivo como parte da configuração da fonte de dados. Se você adicionar documentos diretamente a um índice, você especifica a ACL no objeto Principal como parte do objeto de documento na BatchPutDocumentAPI.

Você pode usar a CreateAccessControlConfigurationAPI para reconfigurar seu controle de acesso existente em nível de documento sem indexar todos os seus documentos novamente. Por exemplo, seu índice contém documentos ultrassecretos da empresa que somente determinados funcionários ou usuários devem acessar. Um desses usuários deixa a empresa ou muda para uma equipe que deveria ser impedida de acessar documentos ultrassecretos. O usuário ainda tem acesso a documentos ultrassecretos porque o usuário teve acesso quando seus documentos foram indexados anteriormente. Crie uma configuração específica de controle de acesso para o usuário com acesso negado. Posteriormente, atualize a configuração do controle de acesso para permitir o acesso caso o usuário retorne à empresa e se junte novamente à equipe "ultrassecreta". Configure novamente o controle de acesso para seus documentos conforme as circunstâncias mudarem.

Para aplicar sua configuração de controle de acesso a determinados documentos, você chama a BatchPutDocumentAPI com o `AccessControlConfigurationId` incluído no objeto Document. Se você usar um bucket do S3 como fonte de dados, atualize o `.metadata.json` com o `AccessControlConfigurationId` e sincronize sua fonte de dados. Amazon Kendra atualmente só oferece suporte à configuração de controle de acesso para fontes de dados e documentos do S3 indexados usando a `BatchPutDocument` API.

Filtragem por token de usuário

Ao consultar um índice, você poderá usar um token de usuário para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Quando você emite uma consulta, Amazon Kendra extrai e valida o token, extrai e verifica as informações do usuário e do grupo e executa a consulta. Todos os documentos aos quais o usuário tem acesso, incluindo documentos públicos, são devolvidos. Para obter mais informações, consulte Controle de acesso de usuário baseado em token.

Você fornece o token do usuário no UserContextobjeto e o passa na API de consulta.

Veja a seguir como incluir um token de usuário.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserToken = {
        Token = "token"
    })
```

Você poderá mapear usuários para grupos. Ao usar a filtragem de contexto de usuário, não é necessário incluir todos os grupos aos quais um usuário pertence ao emitir a consulta. Com a

PutPrincipalMappingAPI, você pode mapear usuários para seus grupos. Se não quiser usar a API de `PutPrincipalMapping`, deverá fornecer o nome do usuário e todos os grupos aos quais o usuário pertence ao emitir uma consulta. Você também pode obter níveis de acesso de grupos e usuários na sua fonte de identidade do IAM Identity Center usando o UserGroupResolutionConfigurationobjeto.

Filtragem por ID de usuário e grupo

Ao consultar um índice, você pode usar o ID do usuário e o grupo para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo aos documentos. Quando você emite uma consulta, Amazon Kendra verifica as informações do usuário e do grupo e executa a consulta. Todos os documentos relevantes para a consulta à qual o usuário tem acesso, inclusive documentos públicos, são retornados.

Você também poderá filtrar os resultados da pesquisa por fontes de dados às quais usuários e grupos têm acesso. Especificar uma fonte de dados é útil se um grupo estiver vinculado a várias fontes de dados, mas você desejar que o grupo acesse apenas documentos de uma determinada fonte de dados. Por exemplo, os grupos “Pesquisa”, “Engenharia” e “Vendas e Marketing” estão todos vinculados aos documentos da empresa armazenados nas fontes de dados do Confluence e Salesforce. No entanto, a equipe de “Vendas e Marketing” só precisa acessar documentos relacionados ao cliente armazenados no Salesforce. Assim, quando usuários de vendas e marketing pesquisam documentos relacionados ao cliente, eles podem visualizar documentos da Salesforce em seus resultados. Usuários que não trabalham em vendas e marketing não veem documentos do Salesforce em seus resultados de pesquisa.

Você fornece as informações do usuário, dos grupos e das fontes de dados no UserContextobjeto e as transmite na API de consulta. O ID do usuário e a lista de grupos e fontes de dados devem corresponder ao nome especificado no objeto Entidade principal para identificar o usuário, os grupos e as fontes de dados. Com o objeto `Principal`, adicione um usuário, grupo ou fonte de dados a uma lista de permissões ou a uma lista de negação para acessar um documento.

Se desejar, especifique um dos seguintes itens:

- Informações de usuários e grupos e informações de fontes de dados (opcionais).
- Somente as informações do usuário se você mapear seus usuários para grupos e fontes de dados usando a PutPrincipalMappingAPI. Você também pode obter níveis de acesso de grupos e usuários na sua fonte de identidade do IAM Identity Center usando o UserGroupResolutionConfigurationobjeto.

Se essas informações não estiverem incluídas na consulta, Amazon Kendra retornará todos os documentos. Se você fornecer essas informações, somente documentos com IDs de usuário, grupos e fontes de dados correspondentes serão retornados.

Veja a seguir como incluir ID de usuário, grupos e fontes de dados.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserId = {
        UserId = "user1"
    },
    Groups = {
        Groups = ["Sales and Marketing"]
    },
    DataSourceGroups = {
        DataSourceGroups = [{"DataSourceId" : "SalesforceCustomerDocsGroup", "GroupId":
 "Sales and Marketing"}]
    })
```

Filtrando por atributo do usuário

Ao consultar um índice, use os atributos integrados `_user_id` e `_group_id` para filtrar os resultados da pesquisa com base no acesso do usuário e do grupo aos documentos. Configure até 100 identificadores de grupo. Quando você emite uma consulta, Amazon Kendra verifica as informações do usuário e do grupo e executa a consulta. Todos os documentos relevantes para a consulta à qual o usuário tem acesso, incluindo documentos públicos, são retornados.

Você fornece os atributos de usuário e grupo no AttributeFilterobjeto e os transmite na API de consulta.

O exemplo a seguir mostra uma solicitação que filtra a resposta da consulta com base na ID do usuário e nos grupos “RH” e “TI” aos quais o usuário pertence. A consulta retornará qualquer documento que tenha o usuário ou os grupos “RH” ou “TI” na lista de permissões. Se o usuário ou qualquer um dos grupos estiver na lista de negação de um documento, o documento não será retornado.

```
response = kendra.query(
        QueryText = query,
        IndexId = index,
```

```
        AttributeFilter = {
            "OrAllFilters": [
                {
                    "EqualsTo": {
                        "Key": "_user_id",
                        "Value": {
                            "StringValue": "user1"
                        }
                     }
                },
                {
                    "EqualsTo": {
                        "Key": "_group_ids",
                        "Value": {
                            "StringListValue": ["HR", "IT"]
                        }
                    }
                }
            ]
        }
        )
```

Você também poderá especificar qual fonte de dados um grupo pode acessar no objeto `Principal`.

> **Note**
>
> A filtragem de contexto do usuário não é um controle de autenticação ou autorização para seu conteúdo. Ele não faz autenticação de usuário no usuário e nos grupos enviados para a API de `Query`. Cabe ao seu aplicativo garantir que as informações do usuário e do grupo enviadas à API de `Query` sejam autenticadas e autorizadas.

Há uma implementação da filtragem de contexto do usuário para cada fonte de dados. A seção a seguir descreve cada implementação.

Tópicos

- Filtragem de contexto do usuário para documentos adicionados diretamente a um índice
- Filtragem de contexto do usuário para perguntas frequentes
- Filtragem de contexto do usuário para fontes de dados

Filtragem de contexto do usuário para documentos adicionados diretamente a um índice

Quando você adiciona documentos diretamente a um índice usando a BatchPutDocumentAPI, Amazon Kendra obtém informações do usuário e do grupo do `AccessControlList` campo do documento. Você fornece uma lista de controle de acesso (ACL) para seus documentos e a ACL é ingerida com seus documentos.

Você especifica a ACL no objeto Principal como parte do objeto Documento na API de `BatchPutDocument` . Você fornece as seguintes informações:

- O acesso que o usuário ou grupo deve ter. Você poderá dizer `ALLOW` ou `DENY`.
- O tipo de entidade. Você poderá dizer `USER` ou `GROUP`.
- O nome do usuário ou grupo.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para perguntas frequentes

Quando você adiciona uma FAQ a um índice, Amazon Kendra obtém informações do usuário e do grupo do `AccessControlList` objeto/campo do arquivo JSON da FAQ. Você também poderá usar um arquivo CSV de perguntas frequentes com campos ou atributos personalizados para controle de acesso.

Você fornece as seguintes informações:

- O acesso que o usuário ou grupo deve ter. Você poderá dizer `ALLOW` ou `DENY`.
- O tipo de entidade. Você poderá dizer `USER` ou `GROUP`.
- O nome do usuário ou grupo.

Para obter mais informações, consulte Arquivos de perguntas frequentes.

Filtragem de contexto do usuário para fontes de dados

Amazon Kendra também rastreia as informações da lista de controle de acesso (ACL) de usuários e grupos de conectores de fonte de dados compatíveis. Isso é útil para a filtragem de contexto do

usuário, em que os resultados da pesquisa são filtrados com base no acesso do usuário ou do grupo aos documentos.

Tópicos

- Filtragem de contexto do usuário para as fontes de dados do Gerenciador de experiência da Adobe
- Filtragem de contexto do usuário para as fontes de dados do Alfresco
- Filtragem de contexto do usuário para fontes de dados Aurora (MySQL)
- Filtragem de contexto do usuário para fontes de dados (PostgreSQL) do Aurora
- Filtragem de contexto do usuário para fontes Amazon FSx de dados
- Filtragem de contexto do usuário para fontes de dados do banco de dados
- Filtragem de contexto do usuário para fontes de dados (Microsoft SQL Server) do Amazon RDS
- Filtragem de contexto do usuário para fontes de dados Amazon RDS (MySQL)
- Filtragem de contexto do usuário para Amazon RDS fontes de dados (Oracle)
- Filtragem de contexto do usuário para fontes de dados (PostgreSQL) do Amazon RDS
- Filtragem de contexto do usuário para fontes de dados do Amazon S3
- Filtragem de contexto do usuário para fontes Amazon WorkDocs de dados
- Filtragem de contexto do usuário para fontes de dados do Box
- Filtragem de contexto do usuário para fontes de dados do Confluence
- Filtragem de contexto do usuário para fontes de dados do Dropbox
- Filtragem de contexto do usuário para fontes de dados do Drupal
- Filtragem de contexto do usuário para fontes GitHub de dados
- Filtragem de contexto do usuário para fontes de dados do Gmail
- Filtragem de contexto do usuário para fontes de dados do Google Drive
- Filtragem de contexto do usuário para fontes de dados do IBM DB2
- Filtragem de contexto do usuário para fontes de dados do Jira
- Filtragem de contexto de usuário para fontes de dados do Microsoft Exchange
- Filtragem de contexto de usuário para fontes de OneDrive dados da Microsoft
- Filtragem de contexto de usuário para fontes de dados Microsoft OneDrive v2.0
- Filtragem de contexto de usuário para fontes de SharePoint dados da Microsoft

- Filtragem de contexto de usuário para fontes de dados do Microsoft SQL Server
- Filtragem de contexto do usuário para fontes de dados do Microsoft Teams
- Filtragem de contexto do usuário para fontes de dados do Microsoft Yammer
- Filtragem de contexto do usuário para fontes de dados MySQL
- Filtragem de contexto do usuário para fontes de dados do Oracle Database
- Filtragem de contexto do usuário para fontes de dados do PostgreSQL
- Filtragem de contexto do usuário para fontes de dados do Quip
- Filtragem de contexto do usuário para fontes de dados do Salesforce
- Filtragem de contexto do usuário para fontes de dados do ServiceNow
- Filtragem de contexto do usuário para fontes de dados do Slack
- Filtragem de contexto do usuário para fontes de dados do Zendesk

Filtragem de contexto do usuário para as fontes de dados do Gerenciador de experiência da Adobe

Quando você usa uma fonte de dados do Adobe Experience Manager, Amazon Kendra obtém as informações do usuário e do grupo da instância do Adobe Experience Manager.

Os IDs do grupo e do usuário são mapeados da seguinte forma:

- `_group_ids` – Os IDs de grupo existem no conteúdo do Gerenciador de experiência da Adobe onde há permissões de acesso definidas. Eles são mapeados a partir dos nomes dos grupos no Gerenciador de experiência da Adobe.
- `_user_id` – Os IDs de usuário existem no conteúdo do Gerenciador de experiência da Adobe onde há permissões de acesso definidas. Eles são mapeados dos e-mails do usuário como IDs no Gerenciador de experiência da Adobe.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para as fontes de dados do Alfresco

Quando você usa uma fonte de dados do Alfresco, Amazon Kendra obtém as informações do usuário e do grupo da instância do Alfresco.

Os IDs do grupo e do usuário são mapeados da seguinte forma:

- `_group_ids` – Existem IDs de grupo no Alfresco em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos nomes do sistema dos grupos (não dos nomes de exibição) no Alfresco.
- `_user_id` – Existem IDs de usuário no Alfresco em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como IDs no Alfresco.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados Aurora (MySQL)

Quando você usa uma fonte de dados Aurora (MySQL), Amazon Kendra obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma Aurora fonte de dados de banco de dados (MySQL) tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados (PostgreSQL) do Aurora

Quando você usa uma fonte de dados Aurora (PostgreSQL) Amazon Kendra , obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma fonte de dados de banco de dados Aurora (PostgreSQL) tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes Amazon FSx de dados

Quando você usa uma fonte de Amazon FSx dados, Amazon Kendra obtém informações de usuários e grupos do serviço de diretório da Amazon FSx instância.

As IDs do Amazon FSx grupo e do usuário são mapeadas da seguinte forma:

- `_group_ids` – Há IDs de grupo no Amazon FSx em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos nomes dos grupos do sistema no serviço de diretório do Amazon FSx.
- `_user_id`—As IDs de usuário existem Amazon FSx em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos nomes de usuário do sistema no serviço de diretório do Amazon FSx.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do banco de dados

Quando você usa uma fonte de dados de banco de dados Amazon Aurora PostgreSQL, como, Amazon Kendra obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no AclConfigurationobjeto como parte do DatabaseConfigurationobjeto na CreateDataSourceAPI.

Uma fonte de dados de banco de dados tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados (Microsoft SQL Server) do Amazon RDS

Quando você usa uma fonte de dados Amazon RDS (Microsoft SQL Server), Amazon Kendra obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma fonte de dados de banco de dados Amazon RDS (Microsoft SQL Server) tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados Amazon RDS (MySQL)

Quando você usa uma fonte de dados Amazon RDS (MySQL), Amazon Kendra obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma Amazon RDS fonte de dados de banco de dados (MySQL) tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para Amazon RDS fontes de dados (Oracle)

Quando você usa uma fonte de dados Amazon RDS (Oracle), Amazon Kendra obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma fonte de dados de banco de dados Amazon RDS (Oracle) tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.

- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados (PostgreSQL) do Amazon RDS

Quando você usa uma fonte de dados Amazon RDS (PostgreSQL) Amazon Kendra , obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma fonte de dados de banco de dados Amazon RDS (PostgreSQL) tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados do Amazon S3

Você adiciona a filtragem de contexto do usuário a um documento em uma fonte de Amazon S3 dados usando um arquivo de metadados associado ao documento. Você adiciona as informações a campo `AccessControlList` no documento do JSON. Para obter mais informações sobre como adicionar metadados aos documentos indexados de uma fonte de dados do Amazon S3 , consulte Metadados do documento do S3.

Você fornece três informações:

- O acesso que a entidade deve ter. Você poderá dizer `ALLOW` ou `DENY`.
- O tipo de entidade. Você poderá dizer `USER` ou `GROUP`.
- O nome da entidade.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes Amazon WorkDocs de dados

Quando você usa uma fonte Amazon WorkDocs de dados, Amazon Kendra obtém informações do usuário e do grupo da Amazon WorkDocs instância.

As IDs do Amazon WorkDocs grupo e do usuário são mapeadas da seguinte forma:

- `_group_ids`—Os IDs de grupo existem Amazon WorkDocs em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos nomes dos grupos em Amazon WorkDocs.
- `_user_id`—As IDs de usuário existem Amazon WorkDocs em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos nomes de usuário em Amazon WorkDocs.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Box

Quando você usa uma fonte de dados do Box, Amazon Kendra obtém informações do usuário e do grupo da instância do Box.

O grupo Box e os IDs de usuário são mapeados da seguinte forma:

- `_group_ids` – Há IDs de grupo no Box em arquivos em que há permissões de acesso definidas. Eles são mapeados a partir dos nomes dos grupos no Box.
- `_user_id` – Há IDs de usuário no Box em arquivos em que há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como IDs de usuário no Box.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Confluence

Quando você usa uma fonte de dados do Confluence, Amazon Kendra obtém informações de usuários e grupos da instância do Confluence.

Você configura o acesso de usuários e grupos aos espaços usando a página de permissões de espaço. Para páginas e blogs, você usa a página de restrições. Para obter mais informações sobre permissões de espaço, consulte Visão geral das permissões de espaço no site de suporte do Confluence. Para obter mais informações sobre restrições de páginas e blogs, consulte Restrições de página no site de suporte do Confluence.

O grupo e os nomes de usuário do Confluence são mapeados da seguinte forma:

- `_group_ids` – Os nomes dos grupos estão presentes em espaços, páginas e blogs onde há restrições. Eles são mapeados a partir do nome do grupo no Confluence. Os nomes dos grupos estão sempre em minúsculas.

- `_user_id` – Os nomes de usuário estão presentes no espaço, na página ou no blog em que há restrições. Eles são mapeados de acordo com o tipo de instância do Confluence que você está usando.

  Para o conector do Confluence v1.0

  - Servidor – O `_user_id` é o nome do usuário. O nome de usuário é sempre minúsculo.
  - Nuvem – O `_user_id` é o ID da conta do usuário.

  Para o conector do Confluence v2.0

  - Servidor – O `_user_id` é o nome do usuário. O nome de usuário é sempre minúsculo.
  - Nuvem – O`_user_id` é o ID de e-mail do usuário.

  > **Important**
  >
  > Para que a filtragem de contexto do usuário funcione corretamente no seu conector do Confluence, você precisa garantir que a visibilidade de um usuário com acesso a uma página do Confluence esteja definida como Qualquer pessoa. Para obter mais informações, consulte Definir a visibilidade do seu e-mail na documentação do desenvolvedor da Atlassian.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Dropbox

Quando você usa uma fonte de dados do Dropbox, Amazon Kendra obtém as informações do usuário e do grupo da instância do Dropbox.

Os IDs do grupo e do usuário são mapeados da seguinte forma:

- `_group_ids` – Há IDs de grupo no Dropbox em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos nomes dos grupos no Dropbox.
- `_user_id` – Há IDs de usuário no Dropbox em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como IDs no Dropbox.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Drupal

Quando você usa uma fonte de dados do Drupal, Amazon Kendra obtém as informações do usuário e do grupo da instância do Drupal.

Os IDs do grupo e do usuário são mapeados da seguinte forma:

- `_group_ids` – Há IDs de grupo no Drupal em arquivos em que há permissões de acesso definidas. Eles são mapeados a partir dos nomes dos grupos no Drupal.
- `_user_id` – Há IDs de usuário no Drupal em arquivos onde há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como IDs no Drupal.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes GitHub de dados

Quando você usa uma fonte de GitHub dados, Amazon Kendra obtém informações do usuário da GitHub instância.

As IDs de GitHub usuário são mapeadas da seguinte forma:

- `_user_id`—As IDs de usuário existem GitHub em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como os IDs em GitHub.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Gmail

Quando você usa uma fonte de dados do Gmail, Amazon Kendra obtém as informações do usuário da instância do Gmail.

Os IDs de usuário são mapeadas da seguinte forma:

- `_user_id` – Há IDs de usuário no Gmail em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como IDs no Gmail.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Google Drive

Uma fonte de dados do Google Workspace Drive retorna informações de usuários e grupos para usuários e grupos do Google Drive. A associação ao grupo e ao domínio é mapeada para o campo de índice de `_group_ids`. O nome de usuário do Google Drive é mapeado para o campo do `_user_id`.

Ao fornecer um ou mais endereços de e-mail de usuário na API de `Query`, somente os documentos que foram compartilhados com esses endereços de e-mail são retornados. O parâmetro `AttributeFilter` a seguir retorna somente documentos compartilhados com "martha@example.com”.

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_user_id",
                   "Value": {
                       "StringValue": "martha@example.com"
                   }
               }
          }
```

Se fornecer um ou mais endereços de e-mail de grupo na consulta, somente os documentos compartilhados com os grupos serão retornados. O parâmetro `AttributeFilter` a seguir retorna somente documentos compartilhados com o grupo "hr@example.com”.

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_group_ids",
                   "Value": {
                       "StringListValue": ["hr@example.com"]
                   }
               }
          }
```

Se fornecer o domínio na consulta, todos os documentos compartilhados com o domínio serão retornados. O parâmetro `AttributeFilter` a seguir retorna documentos compartilhados com o domínio “exemplo.com”.

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_group_ids",
```

```
                "Value": {
                    "StringListValue": ["example.com"]
                }
            }
        }
```

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do IBM DB2

Quando você usa uma fonte de dados IBM DB2, Amazon Kendra obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma fonte de dados de banco de dados do IBM DB2 tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados do Jira

Quando você usa uma fonte de dados do Jira, Amazon Kendra obtém informações de usuários e grupos da instância do Jira.

Os IDs de usuário do Jira são mapeados da seguinte forma:

- `_user_id` – Há IDs de usuário no Jira em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como IDs de usuário no Jira.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto de usuário para fontes de dados do Microsoft Exchange

Quando você usa uma fonte de dados do Microsoft Exchange, Amazon Kendra obtém as informações do usuário da instância do Microsoft Exchange.

As IDs de usuário do Microsoft Exchange são mapeadas da seguinte forma:

- `_user_id`—Existem IDs de usuário nas permissões do Microsoft Exchange para que os usuários acessem determinados conteúdos. Eles são mapeados a partir dos nomes de usuário como IDs no Microsoft Exchange.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto de usuário para fontes de OneDrive dados da Microsoft

Amazon Kendra recupera informações de usuários e grupos da Microsoft OneDrive ao indexar os documentos no site. As informações do usuário e do grupo são obtidas do SharePoint site subjacente da Microsoft que hospeda OneDrive.

Ao usar um OneDrive usuário ou grupo para filtrar os resultados da pesquisa, calcule o ID da seguinte forma:

1. Obter o nome do site. Por exemplo, `https://host.onmicrosoft.com/sites/siteName.`.
2. Pegue o hash MD5 do nome do site. Por exemplo, `430a6b90503eef95c89295c8999c7981`.
3. Crie o e-mail do usuário ou o ID do grupo concatenando o hash MD5 com uma barra vertical (|) e o ID. Por exemplo, se o nome de um grupo for "localGroupName“, o ID do grupo seria:

   ```
   "430a6b90503eef95c89295c8999c7981 | localGroupName"
   ```

   > **Note**
   >
   > Inclua um espaço antes e depois da barra vertical. A barra vertical é usada para se identificar `localGroupName` com seu hash MD5.

   Para o nome de usuário “someone@host.onmicrosoft.com”, o ID do usuário seria o seguinte:

   ```
   "430a6b90503eef95c89295c8999c7981 | someone@host.onmicrosoft.com"
   ```

Envie o ID do usuário ou do grupo Amazon Kendra como o `_group_id` atributo `_user_id` or ao chamar a API de consulta. Por exemplo, o AWS CLI comando que usa um grupo para filtrar os resultados da pesquisa tem a seguinte aparência:

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "430a6b90503eef95c89295c8999c7981 |
 localGroupName"}
                  }}'
```

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto de usuário para fontes de dados Microsoft OneDrive v2.0

Uma fonte de dados Microsoft OneDrive v2.0 retorna informações de seção e página de entidades da lista de controle de OneDrive acesso (ACL). Amazon Kendra usa o domínio do OneDrive locatário para se conectar à OneDrive instância e, em seguida, pode filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo às seções e nomes dos arquivos.

Para objetos padrão, os `_user_id` e `_group_id` são usados da seguinte forma:

- `_user_id`— Seu ID de e-mail de OneDrive usuário da Microsoft é mapeado para o `_user_id` campo.
- `_group_id`— Seu e-mail de OneDrive grupo da Microsoft é mapeado para o `_group_id` campo.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto de usuário para fontes de SharePoint dados da Microsoft

Amazon Kendra recupera informações de usuários e grupos da Microsoft SharePoint ao indexar os documentos do site. Para filtrar os resultados da pesquisa com base no acesso do usuário ou do grupo, forneça informações do usuário e do grupo ao chamar a `Query` API.

Para filtrar usando um nome de usuário, use o endereço de e-mail do usuário. Por exemplo, johnstiles@example.com.

Ao usar um SharePoint grupo para filtrar os resultados da pesquisa, calcule a ID do grupo da seguinte forma:

Para grupos locais

1. Obter o nome do site. Por exemplo, `https://host.onmicrosoft.com/sites/siteName.`.
2. Pegue o hash SHA256 do nome do site. Por exemplo, `430a6b90503eef95c89295c8999c7981`.
3. Crie o ID do grupo concatenando o hash SHA256 com uma barra vertical (|) e o nome do grupo. Por exemplo, se o nome do grupo for "localGroupName“, o ID do grupo seria:

   `"430a6b90503eef95c89295c8999c7981 | localGroupName"`

> **Note**
>
> Inclua um espaço antes e depois da barra vertical. A barra vertical é usada para se identificar `localGroupName` com seu hash SHA256.

Envie o ID do grupo Amazon Kendra como `_group_id` atributo ao chamar a API de consulta. Por exemplo, o AWS CLI comando tem a seguinte aparência:

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "430a6b90503eef95c89295c8999c7981 |
 localGroupName"}
                  }}'
```

Para grupos do AD

1. Use o ID do grupo AD para configurar a filtragem dos resultados da pesquisa.

Envie o ID do grupo Amazon Kendra como `_group_id` atributo ao chamar a API de consulta. Por exemplo, o AWS CLI comando tem a seguinte aparência:

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
```

```
                    "Value": {"StringValue": "AD group"}
                }}'
```

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto de usuário para fontes de dados do Microsoft SQL Server

Quando você usa uma fonte de dados do Microsoft SQL Server, Amazon Kendra obtém informações de usuários e grupos de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma fonte de dados do banco de dados do Microsoft SQL Server tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados do Microsoft Teams

Amazon Kendra recupera informações do usuário do Microsoft Teams ao indexar os documentos. As informações do usuário são obtidas da instância subjacente do Microsoft Teams.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Microsoft Yammer

Amazon Kendra recupera informações do usuário do Microsoft Yammer ao indexar os documentos. As informações do usuário e do grupo são obtidas da instância subjacente do Microsoft Yammer.

Os IDs de usuário do Microsoft Yammer são mapeadas da seguinte forma:

- `_email_id`— O ID de e-mail da Microsoft mapeado para o `_user_id` campo.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados MySQL

Quando você usa uma fonte de dados MySQL, Amazon Kendra obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma fonte de dados de banco de dados do MySQL tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados do Oracle Database

Quando você usa uma fonte de dados do Oracle Database, Amazon Kendra obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma fonte de dados do banco de dados Oracle Database tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados do PostgreSQL

Quando você usa uma fonte de dados do PostgreSQL Amazon Kendra , obtém informações do usuário e do grupo de uma coluna na tabela de origem. Você especifica essa coluna no console ou usa o TemplateConfigurationobjeto como parte da CreateDataSourceAPI.

Uma fonte de dados de banco de dados do PostgreSQL tem as seguintes limitações:

- Você poderá apenas especificar uma lista de permissões para uma fonte de dados do banco de dados. Se desejar, especifique uma lista de negação.
- Você poderá apenas especificar grupos. Não é possível especificar usuários individuais para a lista de permissões.
- A coluna do banco de dados deve ser uma sequência de caracteres contendo uma lista de grupos delimitada por ponto e vírgula.

Filtragem de contexto do usuário para fontes de dados do Quip

Quando você usa uma fonte de dados do Quip, Amazon Kendra obtém as informações do usuário da instância do Quip.

Os IDs de usuário do Quip são mapeados da seguinte forma:

- `_user_id` – Os IDs de usuário existem no Quip em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como IDs no Quip.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Salesforce

Uma fonte de dados do Salesforce retorna informações de usuários e grupos de entidades da lista de controle de acesso (ACL) do Salesforce. Aplique a filtragem de contexto do usuário aos objetos padrão e aos feeds de bate-papo do Salesforce. A filtragem de contexto do usuário não está disponível para artigos de conhecimento do Salesforce.

Se você mapear qualquer campo do Salesforce para os campos de título e corpo do documento do Amazon Kendra, o Amazon Kendra usará dados dos campos de título e corpo do documento nas respostas de pesquisa.

Para objetos padrão, os `_user_id` e `_group_ids` são usados da seguinte forma:

- `_user_id` – O nome de usuário do Salesforce.
- `_group_ids`—
  - Nome do Salesforce `Profile`
  - Nome do Salesforce `Group`
  - Nome do Salesforce `UserRole`

- Nome do Salesforce `PermissionSet`

Para feeds de bate-papo, os `_user_id` e `_group_ids` são usados da seguinte forma:

- `_user_id` – O nome de usuário do Salesforce. Disponível somente se o item for publicado no feed do usuário.
- `_group_ids` – Os IDs de grupo são usados da seguinte forma: Disponível somente se o item do feed for publicado em um bate-papo ou em um grupo de colaboração.
  - O nome da conversa ou grupo de colaboração.
  - Se o grupo for público, `PUBLIC:ALL`.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do ServiceNow

A filtragem de contexto do usuário ServiceNow é compatível somente com a TemplateConfiguration API e o ServiceNow Connector v2.0. ServiceNowConfigurationA API e o ServiceNow Connector v1.0. não oferecem suporte à filtragem de contexto do usuário.

Quando você usa uma fonte de ServiceNow dados, Amazon Kendra obtém as informações do usuário e do grupo da ServiceNow instância.

Os IDs do grupo e do usuário são mapeados da seguinte forma:

- `_group_ids`—Os IDs de grupo existem ServiceNow em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos nomes das funções de `sys_ids` in ServiceNow.
- `_user_id`—As IDs de usuário existem ServiceNow em arquivos nos quais há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como os IDs em ServiceNow.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Slack

Quando você usa uma fonte de dados do Slack, Amazon Kendra obtém as informações do usuário da instância do Slack.

Os IDs de usuário do Slack são mapeados da seguinte forma:

- `_user_id` – Há IDs de usuário no Slack em mensagens e canais em que há permissões de acesso definidas. Eles são mapeados a partir dos e-mails do usuário como IDs no Slack.

Adicionar até 200 entradas no campo `AccessControlList`.

Filtragem de contexto do usuário para fontes de dados do Zendesk

Quando você usa uma fonte de dados do Zendesk, Amazon Kendra obtém as informações do usuário e do grupo da instância do Zendesk.

Os IDs do grupo e do usuário são mapeados da seguinte forma:

- `_group_ids` – Há IDs de grupo nos tickets e artigos do Zendesk em que há permissões de acesso definidas. Eles são mapeados a partir dos nomes dos grupos no Zendesk.
- `_user_id` – Há IDs de grupo nos tickets e artigos do Zendesk em que há permissões de acesso definidas. Eles são mapeados a partir dos emails do usuário como IDs no Zendesk.

Adicionar até 200 entradas no campo `AccessControlList`.

Respostas de consulta e tipos de resposta

Amazon Kendra oferece suporte a diferentes respostas de consulta e tipos de resposta.

Respostas de consulta

Uma chamada para a API de consulta retorna informações sobre os resultados de uma pesquisa. Os resultados estão em uma matriz de QueryResultItemobjetos (`ResultItems`). Cada um do `QueryResultItem` inclui um resumo do resultado. Os atributos do documento associados ao resultado da consulta estão incluídos.

Informações resumidas

As informações de resumo variam de acordo com o tipo de resultado. Em cada caso, ele inclui o texto do documento que corresponde ao termo de pesquisa. Também inclui informações de destaque que você poderá usar para destacar o texto de pesquisa na saída do seu aplicativo. Por exemplo, se o termo de pesquisa for qual é a altura do Space Needle? , as informações resumidas incluem a localização do texto para as palavras altura e agulha de espaço. Para obter mais informações sobre tipos de resposta, consulte Respostas de consulta e tipos de resposta.

Atributos do documento

Cada resultado contém atributos de documento para o documento que correspondem a uma consulta. Alguns dos atributos são predefinidos, como `DocumentId`, `DocumentTitle` e `DocumentUri`. Outros são atributos personalizados que você define. Use os atributos do documento para filtrar a resposta da API de `Query`. Por exemplo, talvez você queira somente os documentos escritos por um autor específico ou por uma versão específica de um documento. Para ter mais informações, consulte Filtragem e pesquisa de facetas. Você especifica os atributos do documento ao adicionar documentos a um índice. Para obter mais informações, consulte Campos ou atributos personalizados.

Veja a seguir um exemplo de código JSON para o resultado de uma consulta. Observe os atributos do documento em `DocumentAttributes` e `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 resposta

Amazon Kendra retorna três tipos de resposta de consulta.

- Resposta (inclui resposta da tabela)
- Documento
- Perguntas e respostas

O tipo da resposta é retornado no campo de `Type` resposta do QueryResultItemobjeto.

Resposta

Amazon Kendra detectou uma ou mais respostas de pergunta na resposta. Um factóide é a resposta a uma pergunta sobre quem, o quê, quando ou onde, como Onde fica o centro de serviço mais próximo de mim?, o Amazon Kendra retorna o texto no índice que melhor corresponde à consulta. O texto está no campo `AnswerText` e contém informações de destaque para o termo de pesquisa no texto da resposta. A `AnswerText` inclui o trecho completo do documento com texto destacado,

enquanto o `DocumentExcerpt` inclui o trecho do documento truncado (290 caracteres) com texto destacado.

Amazon Kendra retorna apenas uma resposta por documento, e essa é a resposta com a maior confiança. Para retornar várias respostas de um documento, divida o documento em vários 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 retorna documentos classificados para aqueles que correspondem ao termo de pesquisa. A classificação é baseada na confiança que Amazon Kendra se tem na precisão do resultado da pesquisa. As informações sobre o documento correspondente são retornadas no QueryResultItem. Inclui o título do documento. O trecho inclui informações de destaque para o texto de pesquisa e a seção de texto correspondente no documento. O URI para documentos correspondentes está no atributo `SourceURI` do documento. O exemplo de JSON a seguir mostra o resumo do documento correspondente.

```
{
    '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'
}
```

Perguntas e respostas

Uma resposta de pergunta e resposta é retornada quando uma pergunta Amazon Kendra corresponde a uma das perguntas mais frequentes em seu índice. A resposta inclui a pergunta e a resposta correspondentes no QueryResultItemcampo. Também inclui informações de destaque para termos de consulta detectados na sequência de caracteres de consulta. O JSON a seguir mostra uma resposta de pergunta e resposta. Observe que a resposta inclui o texto da pergunta.

```
{
    '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 obter informações sobre como adicionar texto de perguntas e respostas a um índice, consulte Criação de perguntas frequentes.

Ajuste e classificação de respostas

Modifique o efeito de um campo ou atributo na relevância da pesquisa por meio do ajuste de relevância. Você também poderá classificar os resultados da pesquisa por um determinado atributo ou campo.

Tópicos

- Ajuste de respostas
- Classificando respostas

Ajuste de respostas

Modifique o efeito de um campo ou atributo na relevância da pesquisa por meio do ajuste de relevância. Para testar rapidamente o ajuste de relevância, use a API de consulta para transmitir configurações de ajuste na consulta. Em seguida, visualize os diferentes resultados de pesquisa obtidos em diferentes configurações. O ajuste de relevância no nível da consulta não tem suporte no console. Você também poderá ajustar campos ou atributos que são do tipo `StringList`somente no nível do índice. Para obter mais informações, consulte Ajustar a relevância da pesquisa.

Por padrão, as respostas da consulta são classificadas pela pontuação de relevância que Amazon Kendra determina cada resultado na resposta.

Ajuste os resultados para qualquer atributo/campo incorporado ou personalizado dos seguintes tipos:

- Valor da data
- Valor longo
- Valor da sequência de caracteres

Não é possível classificar atributos do seguinte tipo:

- Valores da lista de sequência de caracteres

Classifique e ajuste os resultados do documento (AWS SDK)

Defina o parâmetro `Searchable` como verdadeiro para aumentar a configuração dos metadados do documento.

Para ajustar um atributo em uma consulta, defina o parâmetro `DocumentRelevanceOverrideConfigurations` da API de `Query` e especifique o nome do atributo a ser ajustado.

O exemplo de JSON a seguir mostra um objeto `DocumentRelevanceOverrideConfigurations` que substitui o ajuste do atributo chamado “departamento” no índice.

```
"DocumentRelevanceOverrideConfigurations" : [
    "Name": "department",
    "Relevance": {
        "Importance": 1,
```

```
        "ValueImportanceMap": {
            "IT": 3,
            "HR": 7
        }
    }
]
```

Classificando respostas

Amazon Kendra usa o atributo ou campo de classificação como parte dos critérios para os documentos retornados pela consulta. Por exemplo, os resultados retornados por uma consulta classificada por “_created_at” podem não conter os mesmos resultados de uma consulta classificada por “_version”.

Por padrão, as respostas da consulta são classificadas pela pontuação de relevância que Amazon Kendra determina cada resultado na resposta. Para alterar a ordem de classificação, torne um atributo do documento classificável e configure Amazon Kendra para usar esse atributo para classificar as respostas.

Classifique os resultados em qualquer atributo/campo incorporado ou personalizado dos seguintes tipos:

- Valor da data
- Valor longo
- Valor da sequência de caracteres

Não é possível classificar atributos do seguinte tipo:

- Valores da lista de sequência de caracteres

Classifique em um ou mais atributos do documento em cada consulta. As consultas retornam 100 resultados. Se houver menos de 100 documentos com o conjunto de atributos de classificação, os documentos sem um valor para o atributo de classificação serão retornados no final dos resultados, classificados por relevância para a consulta.

Para classificar os resultados do documento (AWS SDK)

1. Para usar a UpdateIndexAPI para tornar um atributo classificável, defina o `Sortable` parâmetro como. `true` O exemplo de JSON a seguir é usado o

`DocumentMetadataConfigurationUpdates`, para adicionar um atributo chamado “Departamento” ao índice e torná-lo classificável.

```
"DocumentMetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE",
       "Search": {
           "Sortable": "true"
       }
   }
]
```

2. Para usar um atributo classificável em uma consulta, defina o parâmetro `SortingConfiguration` da API de consulta. Especifique o nome do atributo a ser classificado e se a resposta deve ser classificada em ordem crescente ou decrescente.

   O exemplo de JSON a seguir mostra o parâmetro `SortingConfiguration` a ser usado para classificar os resultados de uma consulta pelo atributo “Departamento” em ordem crescente.

   ```
      "SortingConfiguration": {
         "DocumentAttributeKey": "Department",
         "SortOrder": "ASC"
      }
   ```

3. Para usar mais de um atributo classificável em uma consulta, defina o parâmetro `SortingConfigurations` da API de consulta. Configure até 3 campos que o Amazon Kendra deve classificar os resultados. Você também poderá especificar se os resultados devem ser classificados em ordem crescente ou decrescente. A cota do campo de classificação pode ser aumentada.

   Se você não fornecer uma configuração de classificação, os resultados serão classificados pela relevância que Amazon Kendra determina o resultado. No caso de empates na classificação dos resultados, os resultados são classificados por relevância.

   O exemplo de JSON a seguir mostra o parâmetro `SortingConfigurations` a ser usado para classificar os resultados de uma consulta pelos atributos “Nome” e “Preço” em ordem crescente.

   ```
   "CollapseConfiguration" : {
       "DocumentAttributeKey": "Name",
       "SortingConfigurations": [
   ```

```
        {
            "DocumentAttributeKey": "Price",
            "SortOrder": "ASC"
        }
    ],
    "MissingAttributeKeyStrategy": "IGNORE"
}
```

Para classificar resultados do documento (console)

> **Note**
>
> Atualmente, a classificação de vários atributos não tem suporte pelo AWS Management Console.

1. Para tornar um atributo classificável no console, escolha Classificável na definição do atributo. Torne um atributo classificável ao criar o atributo ou modificá-lo posteriormente.
2. Para classificar uma resposta de consulta no console, escolha o atributo para classificar a resposta no menu Classificar. Somente os atributos marcados como classificáveis durante a configuração da fonte de dados aparecem na lista.

Reduzir/expandir os resultados da consulta

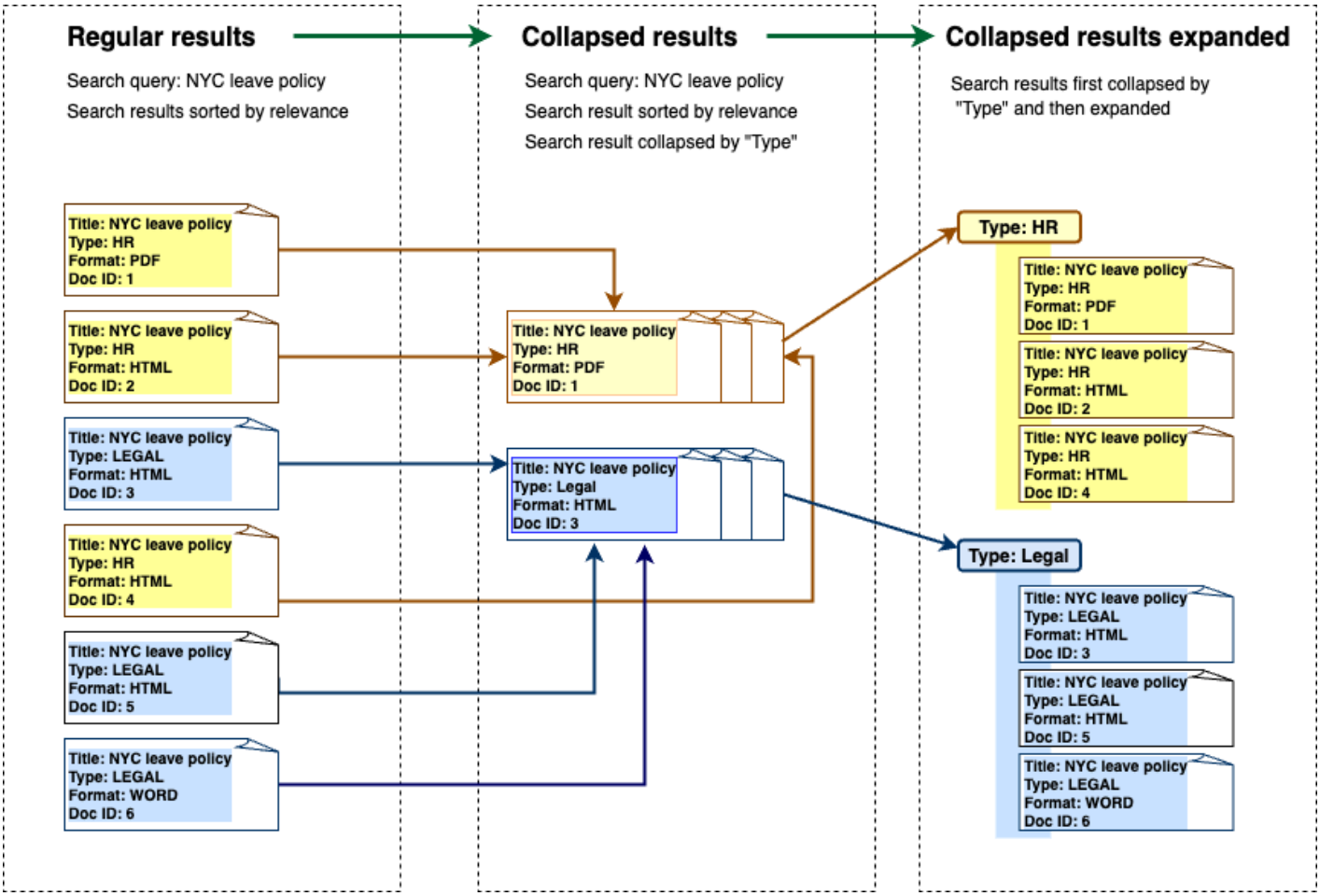
Quando você se conecta Amazon Kendra aos seus dados, ele rastreia os atributos de metadados do documento_`document_title`, como, e`_created_at`, e `_document_id` usa esses atributos ou campos para fornecer recursos avançados de pesquisa durante o tempo de consulta.

O recurso de Reduzir e expandir os resultados da consulta do Amazon Kendra. permite agrupar os resultados da pesquisa usando um atributo de documento comum e exibi-los, reduzidos ou parcialmente expandidos, em um documento primário designado.

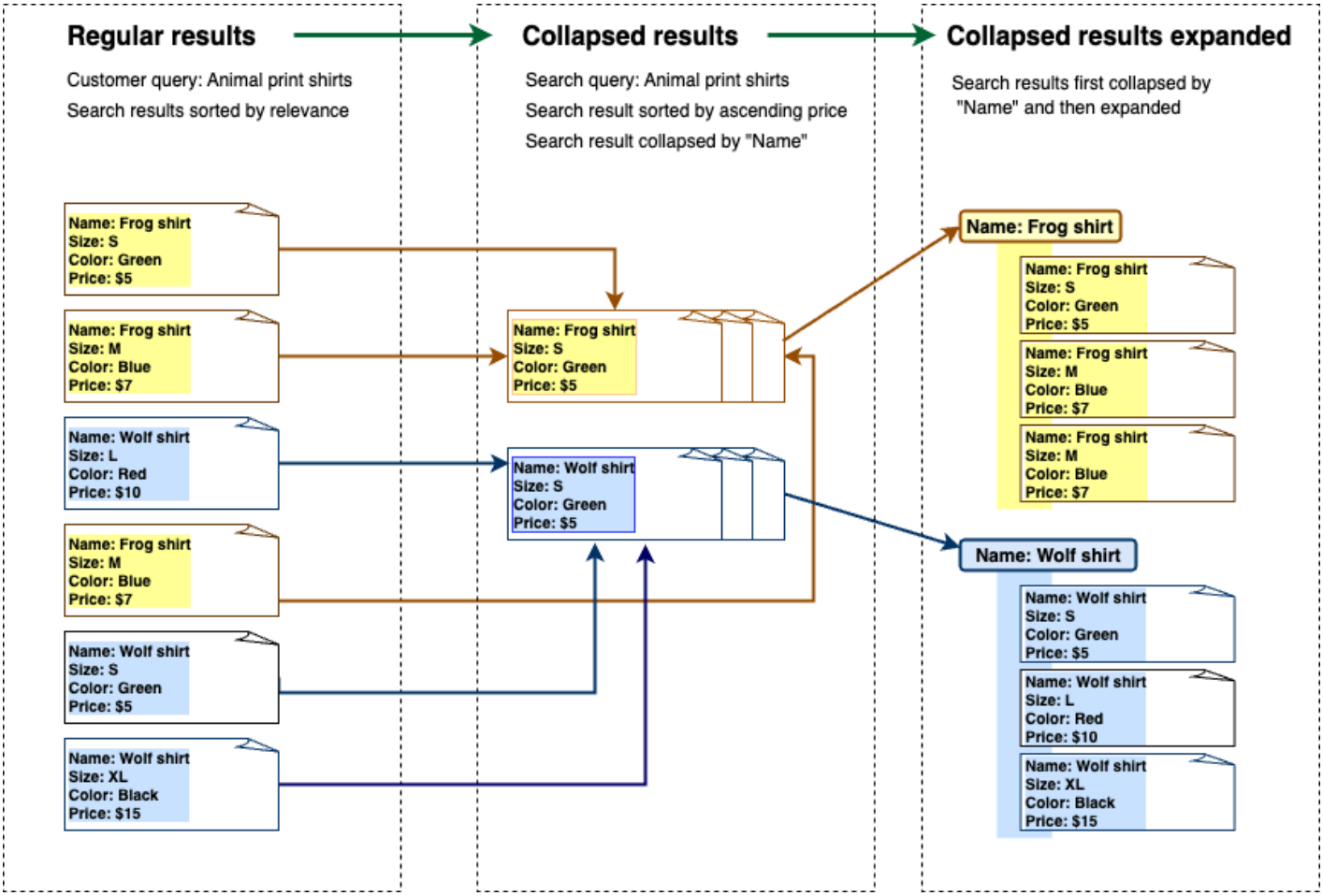
> **Note**
>
> Atualmente, o recurso de reduzir e expandir os resultados da consulta está disponível somente por meio da API do Amazon Kendra.

Isso é útil nos seguintes tipos de situações de pesquisa:

- Existem várias versões do conteúdo nos documentos do seu índice. Quando seu usuário final consulta o índice, você quer que ele veja a versão mais relevante do documento com duplicatas ocultas/reduzidas. Por exemplo, se seu índice contém várias versões de um documento chamado “Política de licença de Nova York”, você pode optar por recolher os documentos dos grupos específicos “RH” e “Jurídico” usando o atributo/campo “Tipo”.



- Seu índice contém vários documentos com informações exclusivas sobre um tipo de item ou objeto, como um inventário de produtos, por exemplo. Para capturar e classificar as informações do item de forma conveniente, você deseja que os usuários finais acessem todos os documentos vinculados por um item ou objeto como um resultado de pesquisa. No exemplo abaixo, uma pesquisa de um cliente por “camisetas com estampa animal” retorna resultados agrupados por nome e classificados por ordem de preço crescente.



Reduzir os resultados

Para agrupar documentos semelhantes ou relacionados, você deve especificar o atributo para recolher (por exemplo, você pode recolher/agrupar documentos por). `_category` Para fazer isso, chame a API de consulta e use o CollapseConfigurationobjeto para especificar o `DocumentAttributeKey` a ser retraído. Os controles de `DocumentAttributeKey` sobre os quais os resultados da pesquisa de campo serão reduzidos. Os campos de chave de atributo com suporte incluem `String` e `Number`. Os tipos `String list` e `Date` não têm suporte.

Escolher um documento primário usando a ordem de classificação

Para configurar o documento principal para ser exibido em um grupo reduzido, use o `SortingConfigurations` parâmetro abaixo CollapseConfiguration. Por exemplo, para obter a versão mais recente de um documento, você classificaria cada grupo reduzido por`_version`. Especifique até 3 atributos/campos para classificar e uma ordem de classificação para cada atributo/

campo usando as `SortingConfigurations`. Solicite um aumento de cota para o número de atributos de classificação.

Por padrão, Amazon Kendra classifica as respostas da consulta pela pontuação de relevância que ela determina para cada resultado na resposta. Para alterar a ordem de classificação padrão, torne os atributos do documento classificáveis e configure Amazon Kendra para usar esses atributos para classificar as respostas. Para obter mais informações, consulte Classificar respostas.

Estratégia-chave do documento ausente

Se seu documento não tiver um valor de atributo de colapso, o Amazon Kendra oferece três opções de personalização:

- Escolha entre `COLLAPSE` todos os documentos com valores nulos ou ausentes em um grupo. Essa é a configuração padrão.
- Escolha `IGNORE` os documentos com valores nulos ou ausentes. Documentos ignorados não aparecerão nos resultados da consulta.
- Escolha `EXPAND` cada documento com um valor nulo ou ausente em um grupo próprio.

Expandir os resultados

Você pode escolher se os grupos de resultados de pesquisa reduzidos se expandem usando o `Expand` parâmetro no CollapseConfigurationobjeto. Os resultados expandidos mantêm a mesma ordem de classificação usada para selecionar o documento principal para o grupo.

Para configurar o número de grupos de resultados de pesquisa reduzidos a serem expandidos, use o `MaxResultItemstoExpand` parâmetro no ExpandConfigurationobjeto. Se definir esse valor como 10, por exemplo, somente os primeiros 10 dos 100 grupos de resultados terão a funcionalidade de expansão.

Para configurar o número de resultados expandidos a serem exibidos por documento primário reduzido, use o parâmetro `MaxExpandResultsPerItem`. Por exemplo, se definir esse valor como 3, no máximo 3 resultados por grupo reduzido serão exibidos.

Interações com outros Amazon Kendra recursos

- Reduzir e expandir os resultados não altera o número de facetas nem afeta o número total de resultados exibidos.

- Os resultados da pesquisa em destaque no Amazon Kendra não serão reduzidos, mesmo que tenham o mesmo valor de campo que o campo de recolhimento configurado.
- A redução e a expansão dos resultados só se aplicam a resultados do tipo`DOCUMENT`.

Ajuste de consulta

Amazon Kendra as consultas produzem resultados de pesquisa classificados por sua relevância. Todos os campos ou atributos pesquisáveis no índice contribuem para essa classificação.

Você poderá modificar o efeito de um campo ou atributo na relevância da pesquisa por meio do ajuste de relevância. O ajuste da relevância da pesquisa pode ser feito manualmente no nível do índice, onde você define as configurações de ajuste para o índice, ou no nível da consulta, substituindo as configurações definidas no nível do índice.

Ao usar o ajuste de relevância, um resultado recebe um impulso na resposta quando a consulta inclui termos que correspondem ao campo ou ao atributo. especifique também quanto impulso o documento recebe quando há uma correspondência. O ajuste de relevância não faz Amazon Kendra com que um documento seja incluído na resposta da consulta, é apenas um dos fatores Amazon Kendra usados para determinar a relevância de um documento.

Aumente campos ou atributos específicos no índice para atribuir mais importância a respostas específicas. Por exemplo, quando alguém pesquisa por “Quando é re:Inventar?” você poderia aumentar a relevância da atualização de documentos no `_last_update_at` campo. Ou, em um índice de relatórios de pesquisa, você poderá impulsionar uma fonte de dados específica no campo “fonte”.

Você também poderá melhorar documentos com base em votos ou contagens de visualizações, o que é comum em fóruns e outras bases de conhecimento de suporte. Você poderá combinar melhorias, por exemplo, para aprimorar documentos que são mais vistos e mais recentes.

Você define a quantidade de impulso que um documento recebe usando o parâmetro `Importance`. Quanto maior a `Importance`, mais o campo ou atributo aumenta a relevância de um documento. Ao ajustar seu índice ou ajustar no nível da consulta, aumente o valor do parâmetro `Importance` em pequenos incrementos até obter o efeito desejado. Para determinar se você está melhorando os resultados da pesquisa, realize a pesquisa e compare os resultados com as consultas anteriores.

Você poderá especificar atributos de data, número ou sequência de caracteres para ajustar um índice ou ajuste no nível da consulta. Você poderá ajustar campos ou atributos que são do tipo `StringList` somente no nível do índice. Cada campo ou atributo tem critérios específicos para quando ele impulsiona um resultado.

- Campos ou atributos de data – Há três critérios específicos para campos de data, `Duration`, `Freshness` e. `RankOrder`

- A `Duration` especifica o período ao qual o impulso se aplica. Por exemplo, se definir o período de tempo para 86400 segundos (ou seja, um dia), o aumento começará a diminuir após um dia. Quanto maior a importância, mais rápido o efeito de impulso diminui.
- O `Freshness` determina o quão recente um documento é quando aplicado a um campo ou atributo. Se aplicar o `Freshness` para o campo da data de criação ou da data da última atualização, um documento criado ou atualizado mais recentemente será considerado “mais recente” do que um documento mais antigo. Por exemplo, se o documento 1 foi criado em 14 de novembro e o documento 2 foi criado em 5 de novembro, o documento 1 é “mais novo” que o documento 2. E se o documento 1 foi atualizado pela última vez em 14 de novembro e o documento 2 foi atualizado pela última vez em 20 de novembro, o documento 2 é “mais novo” que o documento 1. Quanto mais recente o documento, mais esse impulso é aplicado. Você poderá ter somente um campo `Freshness` em seu índice.
- O `RankOrder` aplica o impulso em ordem crescente ou decrescente. Se você especificar `ASCENDING`, datas posteriores terão precedência. Se você especificar `DESCENDING`, datas anteriores têm precedência.

- Campos ou atributos numéricos — Para campos ou atributos numéricos, você pode especificar a ordem de classificação que Amazon Kendra deve ser usada ao determinar a relevância do campo ou atributo. Se você especificar `ASCENDING`, os números mais altos terão precedência. Se você especificar `DESCENDING`, os números menores terão precedência.
- Campos ou atributos de sequência de caracteres – Para campos ou atributos de sequência de caracteres, crie categorias de um campo para dar a cada categoria um impulso diferente. Por exemplo, se você melhorar um campo ou atributo chamado “Departamento”, poderá dar um impulso diferente aos documentos de “RH” do que aos documentos de “Jurídico”. Você poderá aumentar um campo ou atributo do tipo `String`. Você poderá aumentar a `StringList` os campos somente no nível do índice.

Ajuste de relevância no nível do índice

Você ajusta a relevância de um campo ou atributo no nível do índice usando o console para definir o ajuste nos detalhes do índice ou a UpdateIndexAPI.

O exemplo a seguir define o `_last_updated_at` campo como o `Freshness` campo de um documento.

```
"DocumentMetadataConfigurationUpdates" : [
    {
```

```
        "Name": "_last_updated_at",
        "Type": "DATE_VALUE",
        "Relevance": {
            "Freshness": TRUE,
            "Importance": 2
        }
    }
]
```

O exemplo a seguir dá importância diferente às diferentes categorias no campo “departamento”.

```
"DocumentMetadataConfigurationUpdates" : [
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 3,
                "Legal": 1
            }
        }
    }
]
```

Ajuste de relevância no nível da consulta

Você ajusta a relevância de um campo ou atributo no nível da consulta usando a API de consulta.

O ajuste de relevância no nível da consulta não tem suporte no console.

O ajuste no nível da consulta pode acelerar o processo de teste do ajuste de relevância porque você não precisa atualizar manualmente as configurações de ajuste no índice para cada teste. Você poderá ajustar a relevância de um documento passando configurações de ajuste na consulta. Em seguida, você poderá visualizar os diferentes resultados obtidos em diferentes configurações. Uma configuração passada na consulta substitui a configuração definida no nível do índice.

O exemplo a seguir substitui a importância aplicada ao campo “departamento” e a cada categoria de departamento definida no nível do índice, mostrado no exemplo acima. Quando um usuário insere sua consulta de pesquisa, o campo “departamento” tem um nível razoável de importância e o departamento jurídico tem mais importância do que o departamento de RH.

```
"DocumentRelevanceOverrideConfigurations" : [
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 2,
                "Legal": 8
            }
        }
    }
]
```

Obter informações com a análise de pesquisa

Você pode usar o Amazon Kendra Search Analytics para obter informações sobre como seu aplicativo de pesquisa está ajudando ou não seus usuários a encontrar informações.

Amazon Kendra O Analytics fornece uma visão geral de como seus usuários interagem com seu aplicativo de pesquisa e da eficácia da configuração do aplicativo de pesquisa. Você pode visualizar os dados de métricas usando a GetSnapshotsAPI ou selecionando Analytics no painel de navegação no console.

É possível renderizar os dados gerados do `GetSnapshots` em seu próprio painel personalizado. Ou use o painel de métricas fornecido no console, que inclui gráficos visuais. Com um painel visual, é possível procurar por tendências ou padrões no comportamento do usuário ao longo do tempo ou detectar problemas com a configuração do seu aplicativo de pesquisa. Por exemplo, um gráfico de linhas que mostra um número consistente de consultas por dia e um aumento constante pode indicar maior adoção e uso. Por outro lado, uma queda abrupta pode indicar que há um problema que precisa ser investigado.

Use as métricas para fazer conexões entre diferentes pontos de dados para resolver problemas com a forma como os usuários consultam informações ou descobrem oportunidades de negócios. Por exemplo, o documento “Como a IA funciona?” é o documento mais clicado nos resultados da pesquisa, e a consulta mais pesquisada é “Como funciona o machine learning?”. Isso informa sobre os termos e o idioma preferidos que seus usuários usam. Integre esses termos em seus documentos ou use sinônimos personalizados para esses termos para tornar seus documentos mais pesquisáveis para seus usuários.

Métricas para pesquisa

Há 10 métricas para analisar o desempenho do seu aplicativo de pesquisa ou quais informações seus usuários estão procurando. Para recuperar os dados de métricas, você especifica o nome da cadeia de caracteres dos dados métricos a ser recuperado ao chamar o `GetSnapshots`.

Forneça também um intervalo de tempo ou uma janela de tempo para visualizar os dados das métricas. O intervalo de tempo usa o fuso horário do seu índice. Visualize os dados nas seguintes janelas de tempo:

- `THIS_WEEK`: a semana atual, começando no domingo e terminando no dia anterior à data atual.
- `ONE_WEEK_AGO`: na semana anterior, começando no domingo e terminando no sábado seguinte.

- `TWO_WEEKS_AGO`: na semana anterior à semana anterior, começando no domingo e terminando no sábado seguinte.
- `THIS_MONTH`: no mês atual, começando no primeiro dia do mês e terminando no dia anterior à data atual.
- `ONE_MONTH_AGO`: no mês anterior, começando no primeiro dia do mês e terminando no último dia do mês.
- `TWO_MONTHS_AGO`: no mês anterior ao mês anterior, começando no primeiro dia do mês e terminando no último dia do mês.

No console, as janelas de tempo com suporte são Esta semana, Semana anterior, Este mês, Mês anterior.

Taxa de cliques

A proporção de consultas que levam ao clique em um documento nos resultados da pesquisa. Isso ajuda você a entender se a configuração do aplicativo de pesquisa ajuda seus usuários a encontrar informações relevantes para suas consultas. Para consultas que retornam respostas instantâneas, talvez os usuários não precisem clicar em um documento para obter mais informações. Para ter mais informações, consulte the section called “Taxa de resposta instantânea”. Você deve ligar SubmitFeedbackpara garantir que o feedback clicado seja coletado.

Para recuperar dados sobre a taxa de cliques usando a API de `GetSnapshots`, especifique o `metricType` como `AGG_QUERY_DOC_METRICS`. Visualize também essa métrica no console selecionando Análise no painel de navegação.

Taxa de cliques zero

A proporção de consultas que levam a zero cliques nos resultados da pesquisa. Isso ajuda você a entender as lacunas no conteúdo, fornecendo resultados de pesquisa irrelevantes. Para consultas que retornam respostas instantâneas, talvez os usuários não precisem clicar em um documento para obter mais informações. Para ter mais informações, consulte the section called “Taxa de resposta instantânea”. Além disso, suas configurações de pesquisa, como configurações de ajuste, podem ter um impacto na forma como os documentos são retornados nos resultados da pesquisa.

Para recuperar dados com taxa de clique zero usando a API de `GetSnapshots`, especifique o `metricType` como `AGG_QUERY_DOC_METRICS`. Visualize também essa métrica no console selecionando Análise no painel de navegação.

Taxa de resultados da pesquisa

A proporção de consultas que levam a zero resultados de pesquisa. Isso ajuda você a entender as lacunas em seu conteúdo, não fornecendo resultados de pesquisa relevantes.

Para recuperar dados com taxa zero de resultados de pesquisa usando a API de `GetSnapshots`, especifique o `metricType` como `AGG_QUERY_DOC_METRICS`. Visualize também essa métrica no console selecionando Análise no painel de navegação.

Taxa de resposta instantânea

A proporção de consultas com uma resposta instantânea ou perguntas frequentes retornadas. Isso ajuda você a entender o papel das respostas instantâneas no fornecimento de informações.

Para recuperar dados sobre a taxa de resposta instantânea usando a API de `GetSnapshots`, especifique o `metricType` como `AGG_QUERY_DOC_METRICS`. Visualize também essa métrica no console selecionando Análise no painel de navegação.

Principais consultas

As 100 principais consultas pesquisadas por seus usuários. Isso ajuda você a entender quais consultas são populares e o tipo de informação em que seus usuários estão mais interessados.

As métricas incluem o número de vezes que a consulta é pesquisada, a proporção de cliques em um documento, a proporção de nenhum clique em um documento, a profundidade média de cliques nos resultados da pesquisa, a proporção de respostas instantâneas para a consulta e a confiança média dos 10 primeiros resultados de pesquisa de uma consulta.

Para recuperar dados das principais consultas usando a API de `GetSnapshots`, especifique o `metricType` como `QUERIES_BY_COUNT`. Visualize também essa métrica no console selecionando Análise no painel de navegação no console e selecionando Principais consultas em Listas de consultas.

Principais consultas com zero cliques

As 100 principais consultas que levam a zero cliques nos resultados da pesquisa. Isso ajuda você a entender quaisquer lacunas no conteúdo, onde há falta de documentos relevantes para algumas consultas ou a configuração do aplicativo de pesquisa está retornando resultados de pesquisa irrelevantes. Para consultas que retornam respostas instantâneas, talvez os usuários não precisem

clicar em um documento para obter mais informações. Para ter mais informações, consulte the section called “Taxa de resposta instantânea”.

As métricas incluem o número de vezes que a consulta leva a zero cliques, a proporção de zero cliques para a consulta, a proporção de respostas instantâneas para a consulta e a confiança média dos primeiros 10 resultados de pesquisa de uma consulta.

Para recuperar dados das principais consultas com zero cliques usando a API de `GetSnapshots`, especifique o `metricType` como `QUERIES_BY_ZERO_CLICK_RATE`. Visualize também essa métrica no console selecionando Análise no painel de navegação no console e selecionando Principais consultas com zero clique em Listas de consultas.

Principais consultas com zero resultados de pesquisa

As 100 principais consultas que levam a zero resultados de pesquisa. Isso ajuda você a entender quaisquer lacunas no conteúdo, onde não há documentos relevantes para algumas consultas. Ou os usuários podem consultar com termos especializados que possivelmente não resultam em resultados de pesquisa, solicitando que você crie sinônimos personalizados para resolver isso.

As métricas incluem o número de vezes que a consulta leva a zero resultados de pesquisa, a proporção de zero resultados de pesquisa para a consulta e a proporção de vezes que a consulta é pesquisada em comparação com todas as consultas.

Para recuperar dados das principais consultas com zero resultados de pesquisa usando a API de `GetSnapshots`, especifique o `metricType` como `QUERIES_BY_ZERO_RESULT_RATE`. Visualize também essa métrica no console selecionando Análise no painel de navegação no console e selecionando Principais consultas de zero resultados em Listas de consultas.

Documentos mais clicados

Os 100 documentos mais clicados nos resultados da pesquisa. Isso ajuda você a entender quais documentos ou resultados de pesquisa são mais relevantes para os usuários quando eles consultam informações.

As métricas incluem o número de vezes que o documento é clicado, o número de curtidas que um documento recebe dos usuários (polegares para cima) e o número de não curtidas que um documento recebe dos usuários (polegares para baixo).

Para recuperar dados nos documentos mais clicados usando a API de `GetSnapshots`, especifique o `metricType` como `DOCS_BY_CLICK_COUNT`. Visualize também essa métrica no console

selecionando Análise no painel de navegação no console e selecionando Documentos mais clicados em Listas de consultas.

Total de consultas

O número total de consultas pesquisadas pelos usuários. Isso ajuda você a entender o quanto os usuários estão engajados com o aplicativo de pesquisa.

Para recuperar dados sobre o total de consultas usando a API de `GetSnapshots`, especifique o `metricType` como `AGG_QUERY_DOC_METRICS`. Visualize também essa métrica no console selecionando Análise no painel de navegação.

Total de documentos

O número total de documentos em seu índice. Isso ajuda você a comparar o tamanho do índice com o número total de consultas para verificar se há um número adequado de documentos para o volume de consultas.

Para recuperar dados sobre o total de documentos usando a API de `GetSnapshots`, especifique o `metricType` como `AGG_QUERY_DOC_METRICS`. Visualize também essa métrica no console selecionando Análise no painel de navegação.

Exemplo de recuperação de dados métricos

O código a seguir é um exemplo de recuperação de dados nas principais consultas do mês anterior.

Console

Para recuperar as principais consultas do mês anterior

1. No painel de navegação esquerdo, em Índices, selecione o índice e, em seguida, selecione Análise.
2. Na página Análise, selecione o botão Esta semana para alterar a janela de tempo para recuperar os dados para o Mês anterior.
3. Na página Análise, em Listas de consultas, selecione Principais consultas.

CLI

Para recuperar as principais consultas do mês anterior

```
aws kendra get-snapshots \
--index-id index-id \
--interval "ONE_MONTH_AGO" \
--metric-type "QUERIES_BY_COUNT"
```

Python

Para recuperar as principais consultas do mês 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 as principais consultas do mês 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()))
```

De métricas a informações acionáveis

Informações acionáveis são informações significativas extraídas de dados brutos e usadas para orientar suas ações ou decisões. Para extrair significado das métricas e usá-las para gerar informações acionáveis, é importante não apenas analisar as métricas isoladamente, mas também fazer conexões entre elas.

Por exemplo, a consulta principal com zero cliques é “Quais regiões estão disponíveis atualmente?”. No entanto, ele também tem uma taxa de resposta instantânea de 100%. Isso sugere que seus usuários recebam a resposta para essa pergunta sem precisar clicar em um resultado de pesquisa ou documento que forneça informações sobre as regiões disponíveis. Se você analisasse apenas zero cliques, não entenderia a história completa e possivelmente tiraria conclusões erradas sobre o sucesso da configuração do aplicativo de pesquisa ao tratar com essa consulta.

Outro exemplo de uma visão acionável é descobrir uma oportunidade de negócio. As empresas geralmente buscam oportunidades de aumentar seus clientes analisando as métricas de pesquisa. O documento mais clicado é “Regiões disponíveis”. Além disso, a maioria das consultas mais pesquisadas está relacionada a perguntas sobre a disponibilidade de produtos na região oceânica, com taxas de resposta instantânea de 100% e uma alta taxa de cliques para obter mais informações sobre as regiões disponíveis como parte da resposta. Isso sugere que há interesse e demanda por seu produto ou serviço nessa região.

Visualizar e relatar as análises de pesquisa

Há cinco métricas que incluem dados de tendências para você visualizar e procurar tendências ou padrões ao longo do tempo. Se usa o console, gráficos dos dados de tendências são fornecidos.

Se usar as APIs, poderá recuperar os dados de tendências para criar seus próprios gráficos ou visualizações. A maioria dos gráficos no console traça os pontos de dados diários na janela de tempo escolhida.

O console fornece um painel de métricas em que é possível selecionar um gráfico e uma lista principal que você tem interesse em visualizar. Exporte as métricas mostradas em seu painel no formato CSV selecionando Exportar na página inicial da Análise. Inclua esses relatórios em seus documentos comerciais ou apresentações.

Você poderá visualizar as seguintes métricas:

Gráfico de consultas totais

Um gráfico de linhas do número de consultas emitidas por dia. O gráfico ajuda a visualizar padrões no engajamento diário do usuário. Alguns exemplos incluem um aumento ou diminuição constante no engajamento do usuário ou uma queda drástica para 0 consultas devido a uma falha no aplicativo de pesquisa ou problemas com o site.

Se usar a API, poderá recuperar esses dados especificando o `TREND_QUERY_DOC_METRICS`. Use os dados para criar seus próprios gráficos ou use os gráficos fornecidos no console.

Gráfico de taxas de cliques

Um gráfico de linhas das proporções de cliques por dia. O gráfico ajuda você a visualizar padrões na taxa de cliques diária. Alguns exemplos incluem um aumento ou diminuição constante na taxa de cliques ou uma diminuição nas respostas instantâneas que possivelmente influenciam um aumento no número de cliques.

Se usar a API, poderá recuperar esses dados especificando o `TREND_QUERY_DOC_METRICS`. Use os dados para criar seus próprios gráficos ou use os gráficos fornecidos no console.

Gráfico de taxa de zero cliques

Um gráfico de linhas da proporção de zero cliques por dia. O gráfico ajuda a visualizar padrões na taxa diária de zero cliques. Alguns exemplos incluem um aumento ou diminuição constante na taxa de zero cliques ou um aumento nas respostas instantâneas possivelmente influenciando um aumento em zero cliques.

Se usar a API, poderá recuperar esses dados especificando o `TREND_QUERY_DOC_METRICS`. Use os dados para criar seus próprios gráficos ou use os gráficos fornecidos no console.

Gráfico de taxa de resultados de pesquisa zero

Um gráfico de linhas da proporção de zero resultados de pesquisa por dia. O gráfico ajuda você a visualizar padrões na taxa diária de resultados de pesquisa zero. Alguns exemplos incluem um aumento ou diminuição constante na taxa de zero resultados de pesquisa ou uma diminuição acentuada no número de documentos em seu índice, possivelmente influenciando um aumento em zero resultados de pesquisa.

Se usar a API, poderá recuperar esses dados especificando o `TREND_QUERY_DOC_METRICS`. Use os dados para criar seus próprios gráficos ou use os gráficos fornecidos no console.

Gráfico de taxa de resposta instantânea

Um gráfico de linhas da proporção de consultas com uma resposta instantânea ou perguntas frequentes retornadas. O gráfico ajuda a visualizar padrões na taxa de resposta instantânea diária. Alguns exemplos incluem aumento ou diminuição constante nas consultas do tipo pergunta-resposta ou uma diminuição nos cliques, possivelmente influenciando um aumento nas respostas instantâneas.

Se usar a API, poderá recuperar esses dados especificando o `TREND_QUERY_DOC_METRICS`. Use os dados para criar seus próprios gráficos ou use os gráficos fornecidos no console.

Envio de feedback para aprendizado incremental

Amazon Kendra usa aprendizado incremental para melhorar os resultados da pesquisa. Usando o feedback das consultas, o aprendizado incremental melhora os algoritmos de classificação e otimiza os resultados da pesquisa para obter maior precisão.

Por exemplo, suponha que os usuários pesquisam a frase “benefícios de assistência médica”. Se os usuários escolherem consistentemente o segundo resultado da lista, com o tempo, o Amazon Kendra aumentará esse resultado para o primeiro lugar. O aumento diminui com o tempo, portanto, se os usuários pararem de selecionar um resultado, Amazon Kendra eventualmente o removerá e mostrará outro resultado mais popular. Isso ajuda a Amazon Kendra priorizar os resultados com base na relevância, idade e conteúdo.

O aprendizado incremental é ativado para todos os índices e para todos os tipos de documentos com suporte.

Amazon Kendra começa a aprender assim que você fornece feedback, embora possa levar mais de 24 horas para ver os resultados do feedback. Amazon Kendra fornece três métodos para você enviar feedback: o AWS console, uma JavaScript biblioteca que você pode incluir na sua página de resultados de pesquisa e uma API que você pode usar.

Amazon Kendra aceita dois tipos de feedback do usuário:

- Cliques – Informações sobre quais resultados de consulta o usuário escolheu. O feedback inclui o ID do resultado e a marcação de tempo do UNIX, da data e hora em que o resultado da pesquisa foi escolhido.

  Para enviar o feedback de cliques, o aplicativo deve coletar informações de cliques das atividades dos usuários e, em seguida, enviar essas informações para o Amazon Kendra. Você pode coletar informações de cliques com o console, a JavaScript biblioteca e a Amazon Kendra API.
- Relevância – Informações sobre a relevância de um resultado de pesquisa, que o usuário normalmente fornece. O feedback contém o ID do resultado e um indicador de relevância (`RELEVANT` ou`NOT_RELEVANT`). O usuário determina as informações relevantes.

  Para enviar o feedback de relevância, o aplicativo deve fornecer um mecanismo de feedback que permita ao usuário escolher a relevância apropriada para o resultado de uma consulta e, em seguida, enviar essas informações para o Amazon Kendra. Você só pode coletar informações relevantes com o console e a Amazon Kendra API.

O feedback é usado enquanto o índice está ativo. O feedback afeta apenas o índice ao qual é enviado, não pode ser usado em vários índices ou em contas diferentes.

Você deve fornecer contexto de usuário adicional ao consultar seu Amazon Kendra índice. Quando você fornece o contexto do usuário, Amazon Kendra é capaz de saber se o feedback é fornecido por um único usuário ou por vários usuários e ajustar os resultados da pesquisa adequadamente.

Ao fornecer o contexto do usuário, o feedback da consulta é associado ao usuário específico fornecido no contexto. Se não especificar o contexto do usuário, poderá fornecer um ID de visitante usado para agrupar e agregar consultas.

Se não fornecer o contexto do usuário ou um ID de visitante, o feedback será anônimo e agregado a outros comentários anônimos.

O código a seguir mostra como incluir o contexto do usuário como token ou ID do visitante.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserToken = {
        Token = "token"
    })

    OR

    response = kendra.query(
    QueryText = query,
    IndexId = index,
    VisitorId = "visitor-id")
```

Para aplicativos da web, use cookies, localizações ou usuários do navegador para gerar um ID de visitante para cada usuário.

Para consultas principais, o maior volume de consultas, fornecer feedback por clique fornece informações suficientes para melhorar a precisão geral. Para consultas finais, aquelas que são raras, os especialistas no assunto devem enviar feedback relevante e não relevante para melhorar a precisão dessas consultas.

Além do console, você pode usar um dos dois métodos: uma JavaScript biblioteca ou a SubmitFeedbackAPI. Você só deverá usar um método de coleta de feedback. Para obter melhores resultados, envie feedback dentro de 24 horas após fazer a consulta.

Tópicos

- Usando a Amazon Kendra JavaScript biblioteca para enviar feedback
- Usando a Amazon Kendra API para enviar feedback

Usando a Amazon Kendra JavaScript biblioteca para enviar feedback

Amazon Kendra fornece uma JavaScript biblioteca que você pode usar para adicionar feedback de cliques à sua página de resultados de pesquisa. Para usar a biblioteca, você insere uma tag de script no código do cliente que exibe o resultado da pesquisa e, em seguida, adiciona informações a cada um dos links do documento na sua lista de resultados. Quando um usuário escolhe um link para visualizar um documento, as informações de clique são enviadas para o Amazon Kendra.

A biblioteca funciona com navegadores compatíveis com a JavaScript versão ES6/ES2015.

Etapa 1: inserir uma tag de script em seu aplicativo Amazon Kendra de pesquisa

No código do cliente que renderiza os resultados da Amazon Kendra pesquisa, insira uma <script>tag e adicione uma referência à 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>
```

O script baixa de forma assíncrona a JavaScript biblioteca de uma CDN Amazon Kendra hospedada e inicializa uma variável global chamada `kendraFeedback` que permite definir parâmetros opcionais.

Substitua *`o URL de download da biblioteca`* e o *`endpoint de feedback`* por um identificador da tabela a seguir com base na região que hospeda seu Amazon Kendra índice.

Região	Faça download do URL	Endpoint de feedback
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
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

Região	Faça download do URL	Endpoint de feedback
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 exemplo, se seu índice estiver no Leste dos EUA (Norte da Virgínia), o *URL de download da biblioteca* é `https://d2zm0lpns956f8.cloudfront.net/ksf-v1.js` e o *endpoint de feedback* é `https://ujxwp5s92h.execute-api.us-east-1.amazonaws.com/prod/submit`.

Há duas configurações opcionais que você pode fazer para a Amazon Kendra JavaScript biblioteca:

- `disableCookies`— Por padrão, Amazon Kendra define um cookie que identifica o usuário de forma exclusiva. Defina isso como `true` para desativar o cookie.

  ```
  kendraFeedback('disableCookie', 'true | false');
  ```

  `searchDivClassName` – Por padrão, o Amazon Kendra monitora todos os links na página de resultados de pesquisa em busca de cliques. Defina isso como um nome de classe `<div>` para monitorar somente os links na classe especificada.

  ```
  kendraFeedback('searchDivClassName', 'class name');
  ```

Etapa 2: adicionar o token de comentários aos resultados da pesquisa

Na sua página de resultados, adicione um atributo do HTML chamado `data-kendra-token` à tag âncora ou à tag div primária imediata que contém um link para o documento a partir da resposta da consulta. Por exemplo: .

```
<a href="document location" data-kendra-token="feedback token value"></a>
OR
<div data-url="document location" data-kendra-token="feedback token value"></div>
```

Uma resposta de consulta contém um token no campo de `feedbackToken`. O token identifica a resposta de forma exclusiva se o usuário a escolher. Atribua o valor do token ao atributo de `data-kendra-token`. A Amazon Kendra JavaScript biblioteca procura esse token quando o usuário escolhe o resultado e o envia para um Amazon Kendra endpoint como feedback.

A Amazon Kendra JavaScript biblioteca envia apenas o token de feedback e outros metadados, como a hora em que o resultado foi escolhido e um ID de visitante exclusivo.

Etapa 3: testar o script de comentários

Para garantir que a JavaScript biblioteca esteja configurada corretamente e enviando feedback para o endpoint correto, faça o seguinte. Este exemplo usa o navegador do Chrome.

1. Abra as ferramentas para desenvolvedores da Web no navegador. No Chrome, abra o menu do Chrome no canto superior direito do navegador, escolha Mais ferramentas e escolha Ferramentas para desenvolvedores.
2. Verifique se não há erros relacionados à Amazon Kendra JavaScript biblioteca na guia do console.
3. Faça uma pesquisa e escolha qualquer resultado. Na guia Rede das ferramentas do desenvolvedor. Visualize uma solicitação enviada ao endpoint de comentários, o token do resultado e um status de 200 OK.

Usando a Amazon Kendra API para enviar feedback

Para usar a Amazon Kendra API para enviar feedback de consulta, use a SubmitFeedbackAPI. Para identificar a consulta, você fornece o ID do índice ao qual a consulta se aplica e o ID da consulta retornado na resposta da API de consulta.

O exemplo a seguir mostra como enviar feedback de cliques e relevância usando a API Amazon Kendra . Envie vários conjuntos de comentários por meio das matrizes `RelevanceFeedbackItems` e `ClickFeedbackItems`. Este exemplo envia um único clique e um único item de comentários de relevância. O envio dos comentários usa a hora atual.

Para enviar feedback para uma pesquisa (AWS SDK)

1. Você pode usar o código de exemplo a seguir com os valores necessários:

   a. `index id`— O ID do índice ao qual a consulta se aplica.

   b. `query id`— A consulta sobre a qual você deseja fornecer feedback.

   c. `result id`— O ID do resultado da consulta sobre o qual você deseja fornecer feedback. A resposta da consulta contém o ID do resultado.

   d. `relevance value`—Ou `RELEVANT` (o resultado da consulta é relevante) ou `NOT_RELEVANT` (o resultado da consulta não é 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. Execute o código. Depois que os comentários forem enviados, o código exibirá uma mensagem.

Adicionar sinônimos personalizados a um índice

Para adicionar sinônimos personalizados a um índice, você os especifica em um arquivo de dicionário de sinônimos. Você pode incluir termos específicos ou especializados da empresa ao Amazon Kendra usar sinônimos. Sinônimos genéricos em inglês, como`leader, head`, são incorporados Amazon Kendra e não devem ser incluídos em um arquivo de dicionário de sinônimos, incluindo sinônimos genéricos que usam hífens. Amazon Kendra suporta sinônimos para todos os tipos de resposta, que incluem tipos de `DOCUMENT` resposta `QUESTION_ANSWER` e/ou tipos de `ANSWER` resposta. Amazon Kendra atualmente não suporta a adição de sinônimos marcados como palavras irrelevantes. Esse sinalizador será incluído em um release futuro.

Amazon Kendra faz correlações entre sinônimos. Por exemplo, usando o par sinônimo`Dynamo, Amazon DynamoDB`, Amazon Kendra correlaciona o Dynamo com. Amazon DynamoDB A consulta “O que é dínamo?” em seguida, retorna um documento como “O que é Amazon DynamoDB?”. Com sinônimos, Amazon Kendra é mais fácil captar a correlação.

O arquivo de dicionário de sinônimos é um arquivo de texto armazenado em um Amazon S3 bucket. Consulte Adicionar um dicionário de sinônimos a um índice.

O arquivo de dicionário de sinônimos usa o formato de sinônimo Solr. Amazon Kendra tem um limite no número de dicionários de sinônimos por índice. Consulte Cotas.

Os sinônimos podem ser úteis nos seguintes cenários:

- Termos especializados que não são sinônimos tradicionais do idioma inglês, como `NLP, Natural Language Processing`.
- Substantivos próprios com associações semânticas complexas. Esses são substantivos que são improváveis que o público em geral entenda como, por exemplo, no machine learning, `cost, loss, model performance`.
- Diferentes formas de nomes de produtos como, por exemplo, `Elastic Compute Cloud, EC2`.
- Termos específicos do domínio ou da empresa, como nomes de produtos. Por exemplo, `Route53, DNS`.

Não use sinônimos nos seguintes cenários:

- Sinônimos genéricos do idioma inglês, como `leader, head`. Esses sinônimos não são específicos do domínio, e o uso de sinônimos nesses cenários pode ter efeitos indesejados.

- Erros tipográficos, como `teh => the`.
- Variantes morfológicas como os plurais e os possessivos dos substantivos, a forma comparativa e superlativa dos adjetivos e o pretérito, o particípio passado e a forma progressiva dos verbos. Um exemplo de adjetivos comparativos e superlativos é `good, better, best`.
- Unigram (palavra única) interrompe palavras como `WHO`. Palavras-limite do Unigram não são permitidas no dicionário de sinônimos e são excluídas da pesquisa. Por exemplo, `WHO => World Health Organization` é rejeitado. No entanto, você pode usar `W.H.O.` como um termo sinônimo e pode usar palavras irregulares como parte de um sinônimo de várias palavras. Por exemplo, `of` não é permitido, mas `United States of America` é.

Os sinônimos personalizados facilitam a compreensão Amazon Kendra da terminologia específica da sua empresa, expandindo suas consultas para abranger os sinônimos específicos da sua empresa. Embora os sinônimos possam melhorar a precisão da pesquisa, é importante entender como os sinônimos afetam a latência para que você possa otimizar isso.

Uma regra geral para sinônimos é: quanto mais termos em sua consulta forem combinados e expandidos com sinônimos, maior será o impacto potencial na latência. Outros fatores que afetam a latência incluem o tamanho médio dos documentos indexados, o tamanho do seu índice, qualquer filtragem nos resultados da pesquisa e a carga geral do seu índice. Amazon Kendra As consultas que não correspondem a nenhum sinônimo não são afetadas.

Uma diretriz geral sobre como os sinônimos afetam a latência:

Caso de uso	Aumento na latência*
Consultas típicas de linguagem natural ou palavra-chave de 3 a 5 palavras cada	Menor que 15%
Um termo de consulta se expande para 3 sinônimos	
Índice de cerca de 500 mil documentos (média de 10,48 KB de texto extraído por documento) ou 30 mil pares de perguntas frequentes/ perguntas	

*O desempenho varia com base no uso específico de sinônimos e configurações no índice. É melhor testar o desempenho da pesquisa para obter referências mais precisas para o caso de uso específico.

Se o dicionário de sinônimos for grande, tiver uma alta taxa de expansão de prazo e o aumento de latência não estiver dentro dos limites aceitáveis, você pode tentar uma ou as duas opções a seguir:

- Corte o dicionário de sinônimos para reduzir a taxa de expansão (número de sinônimos por termo).
- Reduza a cobertura geral dos termos (número de linhas no dicionário de sinônimos).

Como alternativa, você pode aumentar a capacidade de provisionamento (unidades de armazenamento virtual) para compensar o aumento da latência.

Tópicos

- Criando um arquivo de dicionário de sinônimos
- Adicionar um dicionário de sinônimos a um índice
- Atualizando um dicionário de sinônimos
- Atualizando um dicionário de sinônimos
- Destaques nos resultados da pesquisa

Criando um arquivo de dicionário de sinônimos

Um arquivo de Amazon Kendra dicionário de sinônimos é um arquivo codificado em UTF-8 contendo uma lista de sinônimos no formato de lista de sinônimos Solr. O arquivo .zip deve ter menos de 5 MB.

Há duas maneiras de especificar mapeamentos de sinônimos:

- Os sinônimos bidirecionais são especificados como uma lista de termos separados por vírgulas. Se o usuário consultar qualquer um dos termos, todos os termos da lista serão usados para pesquisar documentos, o que inclui o termo original consultado.
- Sinônimos unidirecionais são especificados como termos separados pelo símbolo “=>” entre eles para mapear termos para seus sinônimos. Se o usuário consultar um termo à esquerda do símbolo “=>”, ele será mapeado para um termo à direita para pesquisar documentos usando o sinônimo. Não é mapeado vice-versa, o que o torna unidirecional.

Os sinônimos em si diferenciam maiúsculas de minúsculas, mas os termos para os quais eles são mapeados não fazem distinção entre maiúsculas e minúsculas. Por exemplo, `ML => Machine Learning` significa que se so usuário consultar “ML”, “ml” ou usar algum outro caso, ele será mapeado para “Machine Learning”. Se você mapeasse isso vice-versa, `Machine Learning => ML`, “Machine Learning”, “machine learning” ou algum outro caso seria mapeado para “ML”.

Um sinônimo não busca uma correspondência exata em caracteres especiais. Por exemplo, se você pesquisar por "dead-letter-queue“, Amazon Kendra poderá retornar documentos que correspondam à “fila de letras mortas” (sem hífen). Se seus documentos contiverem hífens, como "dead-letter-queue“, Amazon Kendra processará os documentos durante a pesquisa para remover hífens. Para termos de sinônimos genéricos em inglês que estão incorporados Amazon Kendra e não devem ser incluídos em um arquivo de dicionário de sinônimos, é Amazon Kendra possível pesquisar tanto a versão com hífen do termo quanto a versão sem hífen do termo. Por exemplo, se você pesquisar “terceiros” e “terceiros”, Amazon Kendra retornará documentos que correspondam a qualquer uma das versões desses termos.

Para sinônimos que contêm palavras irrelevantes ou palavras comumente usadas, Amazon Kendra retorna documentos que correspondem a termos, incluindo palavras irrelevantes. Por exemplo, você pode criar uma regra de sinônimo para mapear “integração” e “integração”. Você não pode usar apenas palavras irrelevantes para sinônimos. Por exemplo, se você pesquisar por “ativado”, Amazon Kendra não poderá retornar todos os documentos que contenham “ativado”.

Algumas regras de sinônimos são ignoradas. Por exemplo, `a => b` é uma regra, mas `a => a` é ignorada e não conta como regra.

A contagem de termos é o número de termos exclusivos no arquivo de sinônimos. O arquivo de exemplo abaixo inclui termos`AWS CodeStar`,`ML`,`Machine Learning`, `autoscaling group`ASG, e muito mais.

Há uma quantidade máxima de regras de sinônimos por dicionário de sinônimos e uma quantidade máxima de sinônimos por termo. Para ter mais informações, consulte Cotas para Amazon Kendra.

O exemplo a seguir mostra um arquivo de dicionário de sinônimos com regras de sinônimos. Cada linha contém uma única regra de sinônimo. Linhas em branco e comentários são ignorados.

```
# 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.
```

Adicionar um dicionário de sinônimos a um índice

Os procedimentos a seguir mostram como adicionar um arquivo de dicionário de sinônimos a um índice. Pode levar até 30 minutos para ver os efeitos do arquivo de sinônimos atualizado. Para

obter mais informações sobre o arquivo de sinônimos, consulte Criando um arquivo de dicionário de sinônimos:

Console

Para adicionar um dicionário de sinônimos

1. No painel de navegação esquerdo, abaixo do índice em que você deseja adicionar uma lista de sinônimos, o dicionário de sinônimos, escolha Sinônimos.
2. Na página Sinônimo, escolha Adicionar dicionário de sinônimos.
3. Em Definir dicionário de sinônimos, dê um nome ao seu dicionário de sinônimos e uma descrição opcional.
4. Nas configurações do dicionário de sinônimos, forneça o Amazon S3 caminho para o arquivo do dicionário de sinônimos. O arquivo deve ter menos de 5 MB.
5. Para a função do IAM, selecione uma função ou selecione Criar uma nova função e especifique um nome de função para criar uma nova função. Amazon Kendra usa essa função para acessar o Amazon S3 recurso em seu nome. A função do IAM tem o prefixo "AmazonKendra-".
6. Escolha Salvar para salvar a configuração e adicionar o dicionário de sinônimos. Depois que o dicionário de sinônimos é ingerido, ele fica ativo e os sinônimos são destacados nos resultados. Pode levar até 30 minutos para ver os efeitos do arquivo de sinônimos atualizado.

CLI

Para adicionar um dicionário de títulos a um índice com o AWS CLI, chame`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
```

Chame `list-thesauri` para ver uma lista de dicionários de sinônimos:

```
aws kendra list-thesauri \
--index-id index-id
```

Para ver os detalhes de um dicionário de sinônimos, chame `describe-thesaurus`:

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

Pode levar até 30 minutos para ver os efeitos do arquivo de sinônimos atualizado.

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.");
```

```
  }
}
```

Atualizando um dicionário de sinônimos

Não é possível alterar a configuração de um dicionário de sinônimos depois que ele é criado. Você pode alterar detalhes como nome do dicionário de sinônimos e informações do IAM. Você também pode alterar a localização do caminho do arquivo de dicionário de sinônimos no Amazon S3. Se você alterar o caminho para o arquivo do dicionário de sinônimos, Amazon Kendra substituirá o dicionário de sinônimos existente pelo dicionário de sinônimos especificado no caminho atualizado.

Pode levar até 30 minutos para ver os efeitos do arquivo de sinônimos atualizado.

> **Note**
>
> Se houver erros de validação ou de sintaxe no arquivo de dicionário de sinônimos, o arquivo de dicionário de sinônimos carregado anteriormente será retido.

Os procedimentos a seguir mostram como modificar os detalhes do dicionário de sinônimos.

Console

Como modificar detalhes do dicionário de sinônimos

1. No painel de navegação à esquerda, no índice que deseja modificar, escolha Sinônimos.
2. Na página Sinônimo, selecione o dicionário de sinônimos que você deseja modificar e escolha Editar.
3. Na página Atualizar dicionário de sinônimos, atualize os detalhes do dicionário de sinônimos.
4. (Opcional) Escolha Alterar o caminho do arquivo do dicionário de sinônimos e, em seguida, especifique um Amazon S3 caminho para o novo arquivo do dicionário de sinônimos. O arquivo de dicionário de sinônimos existente é substituído pelo arquivo que você especificar. Se você não alterar o caminho, Amazon Kendra recarrega o dicionário de sinônimos a partir do caminho existente.

   Se você selecionar Manter o arquivo de dicionário de sinônimos atual, Amazon Kendra não recarregará o arquivo de dicionário de sinônimos.
5. Escolha Salvar para salvar a nova configuração.

Você também pode recarregar o dicionário de sinônimos a partir do caminho do dicionário de sinônimos existente.

Para recarregar um dicionário de sinônimos a partir de um caminho existente

1. No painel de navegação à esquerda, no índice que deseja modificar, escolha Sinônimos.
2. Na página Sinônimo, selecione o dicionário de sinônimos que você deseja recarregar e escolha Atualizar.
3. Na página Recarregar arquivo de dicionário de sinônimos, confirme que você deseja atualizar o arquivo do dicionário de sinônimos.

CLI

Para atualizar um dicionário de sinônimos, chame `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.");
    }
}
```

Atualizando um dicionário de sinônimos

Os procedimentos a seguir mostram como excluir um dicionário de sinônimos.

Console

1. No painel de navegação à esquerda, no índice que deseja modificar, escolha Sinônimos.
2. Na página Sinônimo, selecione o dicionário de sinônimos que deseja excluir.
3. Na página de Detalhes do dicionário de sinônimos, selecione Excluir e, em seguida, confirme para excluir.

CLI

Para excluir um dicionário de sinônimos de um índice com o AWS CLI, chame`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);
  }
}
```

Destaques nos resultados da pesquisa

O destaque de sinônimos está ativado por padrão. As informações de destaque estão incluídas nos resultados da consulta do Amazon Kendra SDK e da CLI. Se você interagir Amazon Kendra usando o SDK ou a CLI, determinará como exibir os resultados.

Os destaques de sinônimos terão o tipo de destaque `THESAURUS_SYNONYM`. Para obter mais informações sobre destaques, consulte o objeto Destaque.

Tutorial: criando uma solução de pesquisa inteligente e enriquecida com metadados com o Amazon Kendra

Este tutorial mostra como criar uma solução de pesquisa inteligente enriquecida com metadados, baseada em linguagem natural, para os dados corporativos usando o Amazon Kendra, o Amazon Comprehend, o Amazon Simple Storage Service (S3) e. AWS CloudShell

O Amazon Kendra é um serviço de pesquisa inteligente que pode criar um índice de pesquisa para seus repositórios de dados não estruturados em linguagem natural. Para facilitar que o clientes encontrem e filtrem respostas relevantes, use o Amazon Comprehend para extrair metadados dosdados e inseri-los no índice de pesquisa do Amazon Kendra.

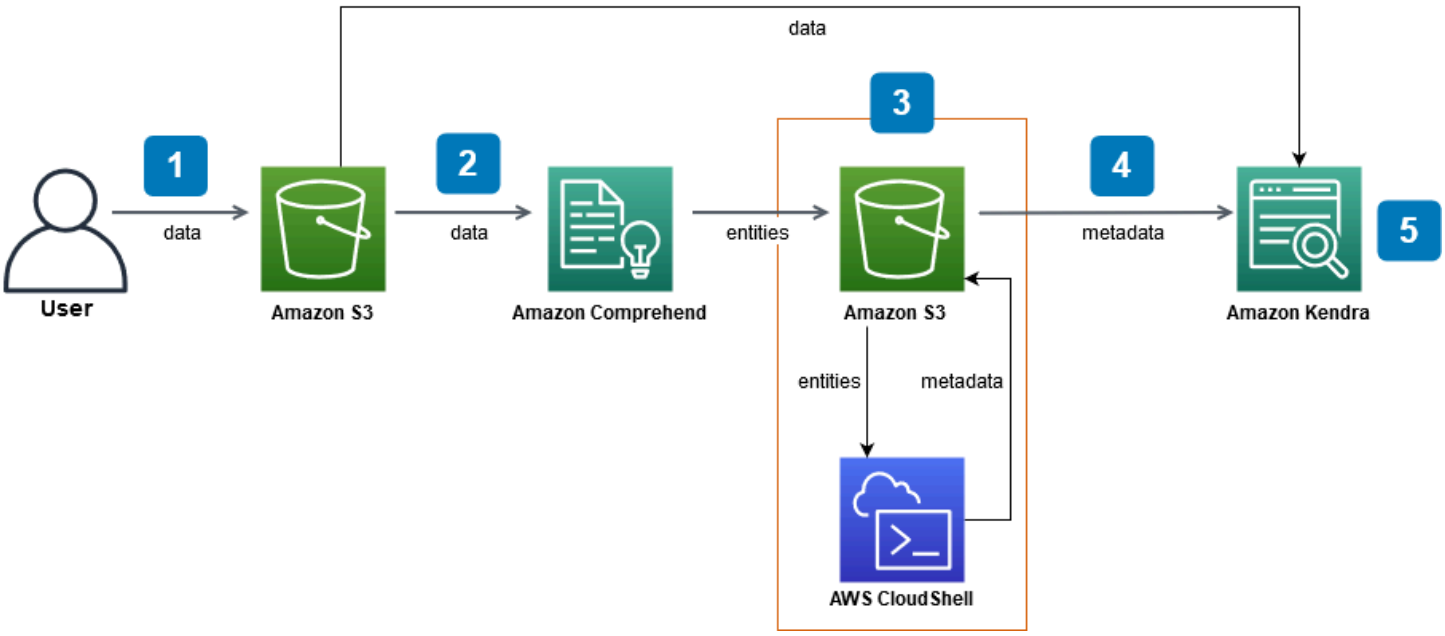
O Amazon Comprehend é um serviço gerenciado de processamento de linguagem natural (PLN) capaz de identificar entidades. Entidades são referências a pessoas, lugares, locais, organizações e objetos nos dados.

Este tutorial usa um conjunto de dados de amostra de artigos de notícias para extrair entidades, convertê-las em metadados e inseri-las no índice do Amazon Kendra para realizar pesquisas. Os metadados adicionados permitem filtrar os resultados da pesquisa usando qualquer subconjunto dessas entidades e melhoram a precisão da pesquisa. Ao seguir este tutorial, você aprenderá como criar uma solução de pesquisa para seus dados corporativos sem nenhum conhecimento especializado em machine learning.

Este tutorial mostra como criar sua solução de pesquisa usando as seguintes etapas:

1. Armazenar um conjunto de dados de amostra de artigos de notícias no Amazon S3.
2. Usar o Amazon Comprehend para extrair entidades dos dados.
3. Executar um script Python 3 para converter as entidades no formato de metadados do índice Amazon Kendra e armazenar esses metadados no S3.
4. Criar um índice de pesquisa do Amazon Kendra e ingerir os dados e os metadados.
5. O índice de pesquisa da consulta.

O diagrama mostra o seguinte fluxo de trabalho:



Tempo estimado para concluir este tutorial: 1 hora

Custo estimado: algumas das ações deste tutorial geram cobranças em sua AWS conta. Para obter mais informações sobre o custo de cada serviço, consulte as páginas de preços do Amazon S3, do Amazon Comprehend e do Amazon Kendra AWS CloudShell.

Tópicos

- Pré-requisitos
- Etapa 1: adicionando documentos ao Amazon S3
- Etapa 2: executar um trabalho de análise de entidades no Amazon Comprehend
- Etapa 3: formatar a saída da análise de entidades como metadados do Amazon Kendra
- Etapa 4: criar um índice de pesquisa do Amazon Kendra e ingerir os metadados.
- Etapa 5: consultar o índice do Amazon Kendra
- Etapa 5: limpar

Pré-requisitos

Para concluir este tutorial, você precisará dos seguintes recursos:

- Uma AWS conta. Se você não tiver uma AWS conta, siga as etapas em Configurar o Amazon Kendra para configurar sua conta. AWS

- Um computador de desenvolvimento que executa Windows, macOS, Linux ou Unix para acessar o console de gerenciamento da AWS . Para obter mais informações, consulte Configurando o console AWS de gerenciamento.
- Um usuário do IAM AWS Identity and Access Management Para saber como configurar usuários e grupos para sua conta, consulte o tutorial de Conceitos básicos no Guia do usuário do IAM.

  Se você estiver usando o AWS Command Line Interface, também precisará anexar a política a seguir ao seu usuário do IAM para conceder a ele as permissões básicas necessárias para concluir este tutorial.

  Para obter mais informações, consulte Criar políticas do IAM em Adicionar e remover permissões de identidade do IAM..
- A Lista de serviços regionais da AWS. Para reduzir a latência, escolha a região AWS mais próxima da sua localização geográfica que seja compatível com o Amazon Comprehend e o Amazon Kendra.
- (Opcional) Um AWS Key Management Service. Embora este tutorial não use criptografia, talvez você queira usar as melhores práticas de criptografia para o caso de uso específico.
- (Opcional) Uma Amazon Virtual Private Cloud. Embora este tutorial não use VCP, talvez você queira usar as melhores práticas de VCP para garantir a segurança dos dados do caso de uso específico.

Etapa 1: adicionando documentos ao Amazon S3

Antes de executar um trabalho de análise de entidades do Amazon Comprehend no conjunto de dados, você cria um bucket do Amazon S3 para hospedar os dados, os metadados e a saída da análise de entidades do Amazon Comprehend.

Tópicos

- Baixar o conjunto de dados de amostra
- Como criar um bucket do Amazon S3
- Criação de pastas de dados e metadados no bucket do S3
- Carregue os dados de entrada

Baixar o conjunto de dados de amostra

Antes que o Amazon Comprehend possa executar um trabalho de análise de entidades em seus dados, você deve baixar e extrair o conjunto de dados e carregá-lo em um bucket do S3.

Para baixar e extrair o conjunto de dados (console)

1. Baixe a pasta tutorial-dataset.zip em seu dispositivo.
2. Extraia a pasta `tutorial-dataset` para acessar a pasta `data`.

Para baixar e extrair o conjunto de dados (terminal)

1. Faça o download de `tutorial-dataset`, execute o seguinte comando na janela do terminal:

   Linux

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   Em que:

   - *path/* é o caminho do arquivo local para o local em que você deseja salvar a pasta zip.

   macOS

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   Em que:

   - *path/* é o caminho do arquivo local para o local em que você deseja salvar a pasta zip.

   Windows

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   Em que:

- *path/* é o caminho do arquivo local para o local em que você deseja salvar a pasta zip.

2. Para extrair os dados da pasta zip, execute o seguinte comando na janela do terminal:

   Linux

   ```
   unzip path/tutorial-dataset.zip -d path/
   ```

   Em que:

   - *path/* é o caminho do arquivo local para sua pasta zip salva.

   macOS

   ```
   unzip path/tutorial-dataset.zip -d path/
   ```

   Em que:

   - *path/* é o caminho do arquivo local para sua pasta zip salva.

   Windows

   ```
   tar -xf path/tutorial-dataset.zip -C path/
   ```

   Em que:

   - *path/* é o caminho do arquivo local para sua pasta zip salva.

No final desta etapa, você deve ter os arquivos extraídos em uma pasta descompactada chamada `tutorial-dataset`. Essa pasta contém um arquivo `README` com uma atribuição de código aberto do Apache 2.0 e uma pasta chamada `data` contendo o conjunto de dados deste tutorial. O conjunto de dados consiste em 100 arquivos com `.story` extensões.

Como criar um bucket do Amazon S3

Depois de baixar e extrair a pasta de dados de amostra, você a armazena em um bucket do Amazon S3.

> **Important**
>
> O nome do bucket do Amazon S3 deve ser exclusivo em todos os AWS.

Para criar um bucket do S3 (console)

1. Faça login AWS Management Console e abra o console do Amazon S3 em https://console.aws.amazon.com/s3/.
2. Em Buckets, escolha Criar bucket.
3. Em Bucket name (Nome do bucket), insira um nome exclusivo.
4. Em Região, escolha a AWS região em que você deseja criar o bucket.

   > **Note**
   >
   > Você deve escolher uma região que ofereça suporte ao Amazon Comprehend e ao Amazon Kendra. Não é possível alterar a região de um bucket após sua criação.

5. Deixe as configurações padrão para Propriedade do objeto, Configurações de bucket para bloquear acesso público, Versionamento de bucket e Tags.
6. Em Criptografia padrão, escolha Desabilitar.
7. Mantenha as configurações padrão para as Configurações avançadas.
8. Revise as configurações do bucket e escolha Criar bucket.

Para criar um bucket do S3 (AWS CLI)

1. Para criar um bucket do S3 com a , use o comando create-bucket no AWS CLI:

   Linux

   ```
   aws s3api create-bucket \
           --bucket DOC-EXAMPLE-BUCKET \
           --region aws-region \
           --create-bucket-configuration LocationConstraint=aws-region
   ```

   Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.
- *aws-region* é a região na qual você deseja criar o bucket.

macOS

```
aws s3api create-bucket \
        --bucket DOC-EXAMPLE-BUCKET \
        --region aws-region \
        --create-bucket-configuration LocationConstraint=aws-region
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.
- *aws-region* é a região na qual você deseja criar o bucket.

Windows

```
aws s3api create-bucket ^
        --bucket DOC-EXAMPLE-BUCKET ^
        --region aws-region ^
        --create-bucket-configuration LocationConstraint=aws-region
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.
- *aws-region* é a região na qual você deseja criar o bucket.

> **Note**
>
> Você deve escolher uma região que ofereça suporte ao Amazon Comprehend e ao Amazon Kendra. Não é possível alterar a região de um bucket após sua criação.

2. Para garantir que o bucket foi criado com êxito, execute o comando lista:

Linux

```
aws s3 ls
```

macOS

```
aws s3 ls
```

Windows

```
aws s3 ls
```

Criação de pastas de dados e metadados no bucket do S3

Depois de criar o bucket do S3, crie pastas de dados e metadados nele.

Para criar pastas no bucket do S3 (console)

1. Abra o console do Amazon S3 em https://console.aws.amazon.com/s3/.
2. Em Buckets, clique no nome do bucket na lista de buckets.
3. Na guia Objetos, escolha Criar pasta.
4. Para o novo nome da pasta, insira **`data`**.
5. Para a configuração de criptografia, escolha Desabilitar.
6. Selecione Criar pasta.
7. Repita as etapas 3 a 6 para criar outra pasta para armazenar os metadados do Amazon Kendra e nomeie a pasta criada na etapa 4. **`metadata`**

Para criar pastas no bucket do S3 (AWS CLI)

1. Para criar a pasta `data` n bucket do S3, use o comando put-object no AWS CLI:

   Linux

   ```
   aws s3api put-object \
           --bucket DOC-EXAMPLE-BUCKET \
   ```

```
        --key data/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

macOS

```
aws s3api put-object \
        --bucket DOC-EXAMPLE-BUCKET \
        --key data/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

Windows

```
aws s3api put-object ^
        --bucket DOC-EXAMPLE-BUCKET ^
        --key data/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

2. Para criar a pasta `metadata` n bucket do S3, use o comando put-object no AWS CLI:

Linux

```
aws s3api put-object \
        --bucket DOC-EXAMPLE-BUCKET \
        --key metadata/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

macOS

```
aws s3api put-object \
        --bucket DOC-EXAMPLE-BUCKET \
        --key metadata/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

Windows

```
aws s3api put-object ^
        --bucket DOC-EXAMPLE-BUCKET ^
        --key metadata/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

3. Para garantir que as pastas tenham sido criadas com sucesso, verifique o conteúdo do bucket usando o comando lista:

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

Carregue os dados de entrada

Depois de criar pastas de dados e metadados, carregue o conjunto de dados de amostra na pasta `data`.

Para carregar o conjunto de dados de amostra na pasta de dados (Console)

1. Abra o console do Amazon S3 em https://console.aws.amazon.com/s3/.
2. Em Buckets, clique no nome do bucket na lista de buckets e, em seguida, escolha.
3. Selecione Adicionar arquivo e clique em Carregar arquivo.
4. Na caixa de diálogo, navegue até a pasta `data` dentro da pasta `tutorial-dataset` em seu dispositivo local, selecione todos os arquivos e escolha Abrir.
5. Mantenha as configurações padrão para Destino, Permissões e Propriedades.
6. Escolha Carregar.

Para carregar o conjunto de dados de amostra na pasta de dados (AWS CLI)

1. Para carregar os dados de amostra na pasta, use o comando `data` copiar em AWS CLI:

   Linux

   ```
   aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
   ```

   Em que:

   - *path/* é o caminho do arquivo para a pasta `tutorial-dataset` no seu dispositivo,
   - DOC-EXAMPLE-BUCKET é o nome do bucket.

macOS

```
aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
```

Em que:

- *path/* é o caminho do arquivo para a pasta `tutorial-dataset` no seu dispositivo,
- DOC-EXAMPLE-BUCKET é o nome do bucket.

Windows

```
aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
```

Em que:

- *path/* é o caminho do arquivo para a pasta `tutorial-dataset` no seu dispositivo,
- DOC-EXAMPLE-BUCKET é o nome do bucket.

2. Para garantir que os arquivos do conjunto de dados tenham sido enviados com sucesso para a pasta `data`, use o comando list na AWS CLI:

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

Ao final desta etapa, você terá um bucket do S3 com o conjunto de dados armazenado dentro da pasta `data` e uma pasta `metadata` vazia, que armazenará os metadados do Amazon Kendra.

Etapa 2: executar um trabalho de análise de entidades no Amazon Comprehend

Depois de armazenar o conjunto de dados de amostra no bucket do S3, execute um trabalho de análise de entidades do Amazon Comprehend para extrair entidades dos documentos. Essas entidades formarão atributos personalizados do Amazon Kendra e ajudarão você a filtrar os resultados da pesquisa no índice. Para obter mais informações, consulte Detectar eventos.

Tópicos

- executando um trabalho de análise de entidades no Amazon Comprehend

executando um trabalho de análise de entidades no Amazon Comprehend

Depois de armazenar o conjunto de dados,, execute um trabalho de análise de entidades do Amazon Comprehend.

Se você estiver usando a AWS CLI nesta etapa, primeiro crie e anexe uma função e uma política AWS do IAM para o Amazon Comprehend e, em seguida, execute um trabalho de análise de entidades. Para executar um trabalho de análise de entidades dos dados de amostra, o Amazon Comprehend precisa de:

- uma função AWS Identity and Access Management (IAM) que a reconhece como uma entidade confiável

- uma política AWS do IAM anexada à função do IAM que lhe dá permissões para acessar seu bucket do S3

Para obter mais informações, consulte Como o Amazon Comprehend funciona com o IAM e Políticas baseadas em identidade para o Amazon Comprehend.

Para executar um trabalho de análise de entidades no Amazon Comprehend (console)

1. Abra o console do Amazon Comprehend em https://console.aws.amazon.com/comprehend/.

   > **Important**
   >
   > Certifique-se de que você esteja na mesma região em que você criou o bucket do Amazon S3. Se você estiver em outra região, escolha a AWS região em que criou seu bucket do S3 no seletor de regiões na barra de navegação superior.

2. Escolha Executar o Amazon Comprehend).
3. No painel de navegação à esquerda, escolha Trabalhos de análise.
4. Escolha Criar trabalho.
5. Na seção Configurações de trabalho, faça o seguinte:
   a. Em Nome, insira **`data-entities-analysis`**.
   b. Em Tipo de análise, escolha Entidades.
   c. Em Idioma, escolha Inglês.
   d. Mantenha a Criptografia do trabalho desativada.
6. Na seção Dados de entrada, faça o seguinte:
   a. Em Fonte de dados, escolha Meus documentos.
   b. Para a Localização do S3, escolha Procurar no S3.
   c. Em Escolher recursos, clique no nome do bucket na lista de buckets.
   d. Em Objetos, selecione o botão de opção para `data` e escolha Escolher.
   e. Em Formato de entrada, escolha Um documento por linha.
7. Na seção Dados de saída, faça o seguinte:
   a. Para a Localização do S3, escolha Procurar no S3 e, em seguida, escolha a caixa de opção para o bucket na lista de buckets e escolha Escolher.

b. Mantenha a Criptografia desativada.

8. Na seção Permissões de acesso, faça o seguinte:

   a. Em Perfil do IAM), escolha Criar um perfil do IAM).

   b. Em Permissões para acessar, escolha Buckets do S3 de entrada e saída.

   c. Em Sufixo do nome, insira **`comprehend-role`**. Essa função fornece acesso ao bucket do Amazon S3.

9. Mantenha a Configuração da VPC padrão.

10. Escolha Criar trabalho.

Para executar um trabalho de análise de entidades no Amazon Comprehend (AWS CLI)

1. Para criar e anexar um perfil do IAM para a Amazon Comprehend que a reconheça como uma entidade confiável, faça o seguinte:

   a. Salve a política de confiança a seguir como um arquivo JSON chamado `comprehend-trust-policy.json` em um editor ou texto em seu computador.

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

   b. Para criar um perfil do IAM chamado `comprehend-role` e anexar o arquivo `comprehend-trust-policy.json` salvo para ele, use o comando create-role:

      Linux

```
aws iam create-role \
          --role-name comprehend-role \
```

```
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `comprehend-trust-policy.json` no dispositivo local.

macOS

```
aws iam create-role \
          --role-name comprehend-role \
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `comprehend-trust-policy.json` no dispositivo local.

Windows

```
aws iam create-role ^
          --role-name comprehend-role ^
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `comprehend-trust-policy.json` no dispositivo local.

c. Copie o nome do recurso da Amazon (ARN) no editor de texto e salve-o localmente como `comprehend-role-arn`.

> **Note**
>
> O ARN tem um formato semelhante a *arn:aws:iam::123456789012:role/comprehend-role*. Você precisa do ARN no qual você salvou `comprehend-role-arn` para executar o trabalho de análise do Amazon Comprehend.

2. Para criar e anexar uma política do IAM ao seu perfil do IAM do IAM que conceda permissões para acessar seu bucket do S3, faça o seguinte:

   a. Salve a política de confiança a seguir como um arquivo JSON chamado `comprehend-S3-access-policy.json` em um editor ou texto em seu computador.

```
{
    "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 criar uma política do IAM chamada `comprehend-S3-access-policy` para acessar o bucket do S3, use o comando create-policy:

Linux

```
aws iam create-policy \
          --policy-name comprehend-S3-access-policy \
          --policy-document file://path/comprehend-S3-access-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `comprehend-S3-access-policy.json` no dispositivo local.

macOS

```
aws iam create-policy \
          --policy-name comprehend-S3-access-policy \
          --policy-document file://path/comprehend-S3-access-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `comprehend-S3-access-policy.json` no dispositivo local.

Windows

```
aws iam create-policy ^
          --policy-name comprehend-S3-access-policy ^
          --policy-document file://path/comprehend-S3-access-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `comprehend-S3-access-policy.json` no dispositivo local.

c. Copie o nome do recurso da Amazon (ARN) no editor de texto e salve-o localmente como `comprehend-S3-access-arn`.

> **Note**
>
> O ARN tem um formato semelhante a *`arn:aws:iam::123456789012:role/comprehend-S3-access-policy`*. Você precisa do ARN em que salvou `comprehend-S3-access-arn` para anexar `comprehend-S3-access-policy` ao perfil do IAM.

d. Para anexar o `comprehend-S3-access-policy` à sua função do IAM, use o attach-role-policycomando:

Linux

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name comprehend-role
```

Em que:

- *`policy-arn`* é o ARN com o qual você salvou `comprehend-S3-access-arn`.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name comprehend-role
```

Em que:

- *`policy-arn`* é o ARN com o qual você salvou `comprehend-S3-access-arn`.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name comprehend-role
```

Em que:

- *policy-arn* é o ARN com o qual você salvou `comprehend-S3-access-arn`.

3. Para executar um trabalho de análise de entidades do Amazon Comprehend, use o comando: start-entities-detection-job

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
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket do S3.
- *policy-arn* é o ARN com o qual você salvou `comprehend-role-arn`.
- *aws-region* é sua região. 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
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket do S3.
- *policy-arn* é o ARN com o qual você salvou `comprehend-role-arn`.

- *aws-region* é sua região. 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
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket do S3.
- *policy-arn* é o ARN com o qual você salvou `comprehend-role-arn`.
- *aws-region* é sua região. AWS

4. Copie a análise das entidades `JobId` e salve-a em um editor de texto como `comprehend-job-id`. `JobId` ajuda a rastrear o status do trabalho de análise de entidades.
5. Para acompanhar o progresso do seu trabalho de análise de entidades, use o describe-entities-detection-jobcomando:

   Linux

   ```
   aws comprehend describe-entities-detection-job \
           --job-id entities-job-id \
           --region aws-region
   ```

   Em que:

   - *entities-job-id*é seu salvo`comprehend-job-id`,
   - *aws-region* é sua região. AWS

   macOS

   ```
   aws comprehend describe-entities-detection-job \
   ```

```
        --job-id entities-job-id \
        --region aws-region
```

Em que:

- *entities-job-id*é seu salvocomprehend-job-id,
- *aws-region* é sua região. AWS

Windows

```
aws comprehend describe-entities-detection-job ^
        --job-id entities-job-id ^
        --region aws-region
```

Em que:

- *entities-job-id*é seu salvocomprehend-job-id,
- *aws-region* é sua região. AWS

Pode levar vários minutos para que `JobStatus` seja alterado para `COMPLETED`.

Ao final dessa etapa, o Amazon Comprehend armazena os resultados da análise de entidades como um arquivo `output.tar.gz` compactado dentro de uma pasta dentro de `output` uma pasta gerada automaticamente no bucket do S3. O status do trabalho de análise deve estar concluído antes de passar para a próxima etapa.

Etapa 3: formatar a saída da análise de entidades como metadados do Amazon Kendra

Para converter as entidades extraídas pelo Amazon Comprehend para o formato de metadados exigido por um índice do Amazon Kendra, execute um script Python 3. Os resultados da conversão são armazenados na pasta `metadata` do bucket da Amazon S3.

Para obter mais informações sobre o formato e a estrutura dos metadados do Amazon Kendra, consulte Metadados do documento do S3.

Tópicos

- Baixando e extraindo a saída do Amazon Comprehend
- Carregando a saída no bucket do S3
- Conversão da saída para o formato de metadados do Amazon Kendra
- Como limpar o bucket do Amazon S3

Baixando e extraindo a saída do Amazon Comprehend

Para formatar a saída da análise de entidades do Amazon Comprehend, você deve primeiro baixar o arquivo de análise de entidades do Amazon Comprehend do `output.tar.gz` e extrair o arquivo de análise de entidades.

Para baixar e extrair os arquivos de saída (console)

1. No console do Amazon Comprehend, no painel de navegação, acesse às Tarefas de análise..
2. Escolha sua tarefa de análise de entidades `data-entities-analysis`.
3. Em Saída, escolha o link exibido ao lado do Local dos dados de saída. Isso redireciona você para o arquivo de `output.tar.gz` em seu bucket do S3.
4. Na página Visão geral selecione Fazer download.

> **Tip**
>
> A saída de todos os trabalhos de análise do Amazon Comprehend tem o mesmo nome. Renomear p arquivo ajudará você a rastreá-lo com mais facilidade.

5. Descompacte e extraia o arquivo do Amazon Comprehend baixado para o seu dispositivo.

Para baixar e extrair os arquivos de saída (AWS CLI)

1. Para acessar o nome da pasta gerada automaticamente pelo Amazon Comprehend em seu bucket do S3 que contém os resultados do trabalho de análise de entidades, use o comando: describe-entities-detection-job

   Linux

   ```
   aws comprehend describe-entities-detection-job \
             --job-id entities-job-id \
             --region aws-region
   ```

Em que:

- *entities-job-id*você está salvo comprehend-job-id dethe section called “Etapa 2: detectar entidades”,
- *aws-region* é sua região. AWS

macOS

```
aws comprehend describe-entities-detection-job \
          --job-id entities-job-id \
          --region aws-region
```

Em que:

- *entities-job-id*você está salvo comprehend-job-id dethe section called “Etapa 2: detectar entidades”,
- *aws-region* é sua região. AWS

Windows

```
aws comprehend describe-entities-detection-job ^
          --job-id entities-job-id ^
          --region aws-region
```

Em que:

- *entities-job-id*você está salvo comprehend-job-id dethe section called “Etapa 2: detectar entidades”,
- *aws-region* é sua região. AWS

2. Do objeto OutputDataConfig na descrição do cargo de sua entidade, copie e salve o valor S3Uri como comprehend-S3uri em um editor de texto.

> **Note**
>
> O S3Uri valor tem um formato semelhante a *s3://DOC-EXAMPLE-BUCKET/...* / output/output.tar.gz.

3. Para baixar o arquivo de saída das entidades, use o comando copiar:

   Linux

   ```
   aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
   ```

   Em que:

   - *s3://DOC-EXAMPLE-BUCKET/...* /output/output.tar.gz é o `S3Uri` valor que você salvou como `comprehend-S3uri`
   - *path/* é o diretório local em que você deseja salvar a saída.

   macOS

   ```
   aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
   ```

   Em que:

   - *s3://DOC-EXAMPLE-BUCKET/...* /output/output.tar.gz é o `S3Uri` valor que você salvou como `comprehend-S3uri`
   - *path/* é o diretório local em que você deseja salvar a saída.

   Windows

   ```
   aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
   ```

   Em que:

   - *s3://DOC-EXAMPLE-BUCKET/...* /output/output.tar.gz é o `S3Uri` valor que você salvou como `comprehend-S3uri`
   - *path/* é o diretório local em que você deseja salvar a saída.

4. Para extrair a saída das entidades, execute o seguinte comando em uma janela de terminal:

   Linux

   ```
   tar -xf path/output.tar.gz -C path/
   ```

Em que:

- *path/* é o caminho do arquivo para o arquivo `output.tar.gz` baixado no dispositivo local.

macOS

```
tar -xf path/output.tar.gz -C path/
```

Em que:

- *path/* é o caminho do arquivo para o arquivo `output.tar.gz` baixado no dispositivo local.

Windows

```
tar -xf path/output.tar.gz -C path/
```

Em que:

- *path/* é o caminho do arquivo para o arquivo `output.tar.gz` baixado no dispositivo local.

Ao final desta etapa, você deve ter um arquivo no dispositivo chamado `output` com uma lista de entidades identificadas pelo Amazon Comprehend.

Carregando a saída no bucket do S3

Depois de baixar e extrair o arquivo de análise de entidades do Amazon Comprehend, carregue o arquivo extraído `output` no bucket do Amazon S3.

Faça upload de arquivos de saída para o Amazon Comprehend (console)

1. Abra o console do Amazon S3 em https://console.aws.amazon.com/s3/.
2. Em Buckets, escolha o nome do bucket e, em seguida, escolha Carregar.
3. Em Arquivos e pastas, escolha Adicionar arquivos.

4. Na caixa de diálogo, navegue até o arquivo `output` extraído no dispositivo, selecione-o e escolha Abrir.
5. Mantenha as configurações padrão para Destino, Permissões e Propriedades.
6. Escolha Carregar.

Faça upload de arquivos de saída para o Amazon Comprehend (AWS CLI)

1. Para fazer o upload do arquivo extraído `output` para o bucket, use o comando copiar:

   Linux

   ```
   aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
   ```

   Em que:

   - *path/* é o caminho do arquivo local para o arquivo extraído `output`,
   - DOC-EXAMPLE-BUCKET é o nome do bucket.

   macOS

   ```
   aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
   ```

   Em que:

   - *path/* é o caminho do arquivo local para o arquivo extraído `output`,
   - DOC-EXAMPLE-BUCKET é o nome do bucket.

   Windows

   ```
   aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
   ```

   Em que:

   - *path/* é o caminho do arquivo local para o arquivo extraído `output`,
   - DOC-EXAMPLE-BUCKET é o nome do bucket.

2. Para garantir que o arquivo `output` tenha sido carregado com sucesso no bucket do S3, verifique o conteúdo usando o comando list:

   Linux

   ```
   aws s3 ls s3://DOC-EXAMPLE-BUCKET/
   ```

   Em que:

   - DOC-EXAMPLE-BUCKET é o nome do bucket.

   macOS

   ```
   aws s3 ls s3://DOC-EXAMPLE-BUCKET/
   ```

   Em que:

   - DOC-EXAMPLE-BUCKET é o nome do bucket.

   Windows

   ```
   aws s3 ls s3://DOC-EXAMPLE-BUCKET/
   ```

   Em que:

   - DOC-EXAMPLE-BUCKET é o nome do bucket.

Conversão da saída para o formato de metadados do Amazon Kendra

Para converter a saída do Amazon Comprehend em metadados do Amazon Kendra, execute um script Python 3. Se você estiver usando o console, use AWS CloudShell para esta etapa.

Para executar o script Python 3 (Console)

1. Baixe o arquivo compactado converter.py.zip em seu dispositivo.
2. Extraia o arquivo Python 3 `converter.py`.

3. Faça login no AWS Management Console e certifique-se de que sua AWS região esteja configurada para a mesma região do bucket do S3 e do trabalho de análise do Amazon Comprehend.
4. Escolha o AWS CloudShell ícone ou digite AWS CloudShellna caixa Pesquisar na barra de navegação superior para iniciar um ambiente.

   > **Note**
   >
   > Quando AWS CloudShell é iniciado em uma nova janela do navegador pela primeira vez, um painel de boas-vindas é exibido e lista os principais recursos. O shell estará pronto para interação após você fechar esse painel e o prompt de comando for exibido.

5. Depois que o terminal estiver preparado, escolha Ações no painel de navegação e escolha Carregar arquivo no menu.
6. Na caixa de diálogo que se abre, escolha Selecionar arquivo e, em seguida, escolha o arquivo Python 3 baixado `converter.py` do dispositivo. Escolha Carregar.
7. No AWS CloudShell ambiente, insira o seguinte comando:

   ```
   python3 converter.py
   ```

8. Quando a interface do shell solicitar que você Insira o nome do bucket do S3, insira o nome do bucket do S3 e pressione enter.
9. Quando a interface do shell solicitar que você Insira o caminho completo do arquivo de saída do Comprehend, digite e pressione enter **`output`**.
10. Quando a interface do shell solicitar que você Insira o caminho completo do arquivo de metadados, digite e pressione enter **`metadata/`**.

> **Important**
>
> Para que os metadados sejam formatados corretamente, os valores de entrada nas etapas 8 a 10 devem ser exatos.

Para executar o script Python 3 (AWS CLI)

1. Faça o download do arquivo Python `converter.py`, execute o seguinte comando na janela do terminal:

Linux

```
curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/
samples/converter.py.zip
```

Em que:

- *path/* é o caminho do arquivo para o local em que você deseja salvar a pasta zip.

macOS

```
curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/
samples/converter.py.zip
```

Em que:

- *path/* é o caminho do arquivo para o local em que você deseja salvar a pasta zip.

Windows

```
curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/
samples/converter.py.zip
```

Em que:

- *path/* é o caminho do arquivo para o local em que você deseja salvar a pasta zip.

2. Para extrair o arquivo Python 3, execute o seguinte comando na janela do terminal:

   Linux

   ```
   unzip path/converter.py.zip -d path/
   ```

   Em que:

   - *path/* é o caminho do arquivo salvo `converter.py.zip`.

macOS

```
unzip path/converter.py.zip -d path/
```

Em que:

- *path/* é o caminho do arquivo salvo `converter.py.zip`.

Windows

```
tar -xf path/converter.py.zip -C path/
```

Em que:

- *path/* é o caminho do arquivo salvo `converter.py.zip`.

3. Certifique-se de que o Boto3 esteja instalado no dispositivo executando o seguinte comando:

Linux

```
pip3 show boto3
```

macOS

```
pip3 show boto3
```

Windows

```
pip3 show boto3
```

> **Note**
>
> Se você não tiver o Boto3 instalado, execute `pip3 install boto3` para instalá-lo.

4. Para executar o script Python 3 para converter o `output` arquivo, execute o comando a seguir.

Linux

```
python path/converter.py
```

Em que:

- *path/* é o caminho do arquivo salvo `converter.py.zip`.

macOS

```
python path/converter.py
```

Em que:

- *path/* é o caminho do arquivo salvo `converter.py.zip`.

Windows

```
python path/converter.py
```

Em que:

- *path/* é o caminho do arquivo salvo `converter.py.zip`.

5. Quando AWS CLI solicitado`Enter the name of your S3 bucket`, insira o nome do seu bucket do S3 e pressione enter.
6. Quando AWS CLI solicitado`Enter the full filepath to your Comprehend output file`, insira **`output`** e pressione enter.
7. Quando AWS CLI solicitado`Enter the full filepath to your metadata folder`, insira **`metadata/`** e pressione enter.

> **Important**
>
> Para que os metadados sejam formatados corretamente, os valores de entrada nas etapas 5 a 7 devem ser exatos.

No final dessa etapa, os metadados formatados são depositados dentro da pasta `metadata` no bucket do S3.

Como limpar o bucket do Amazon S3

Como o índice do Amazon Kendra sincroniza todos os arquivos armazenados em um bucket, recomendamos que você limpe o bucket do Amazon S3 para evitar resultados de pesquisa redundantes.

Limpe o bucket do Amazon S3 (Console)

1. Abra o console do Amazon S3 em https://console.aws.amazon.com/s3/.
2. Em Buckets, escolha o bucket e, em seguida, selecione a pasta de saída da análise de entidades do Amazon Comprehend, o arquivo de análise de entidades `.temp` do Amazon Comprehend e o arquivo `output` extraído do Amazon Comprehend.
3. Na guia Visão geral, escolha Excluir.
4. Em Excluir objetos, escolha Excluir objetos permanentemente? e insira **`permanently delete`** no campo de entrada de texto.
5. Escolha Delete objects (Excluir objetos).

Como limpar o bucket do Amazon S3 (AWS CLI)

1. Para excluir todos os arquivos e as pastas no bucket do S3, exceto as pastas `data` e `metadata` use o comando remover no AWS CLI:

   Linux

   ```
   aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude
    "metadata/*"
   ```

   Em que:

   - DOC-EXAMPLE-BUCKET é o nome do bucket.

macOS

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude
 "metadata/*"
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

Windows

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude
 "metadata/*"
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

2. Para garantir que os objetos tenham sido carregados com sucesso no bucket do S3, verifique o conteúdo usando o comando list:

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Em que:

- DOC-EXAMPLE-BUCKET é o nome do bucket.

Ao final desta etapa, você converteu a saída da análise de entidades do Amazon Comprehend em metadados do Amazon Kendra. Agora, você está pronto para criar um índice do Amazon Kendra.

Etapa 4: criar um índice de pesquisa do Amazon Kendra e ingerir os metadados.

Para implementar a solução de pesquisa inteligente, você cria um índice do Amazon Kendra e ingere seus dados e metadados do S3 nele.

Antes de adicionar metadados ao índice do Amazon Kendra, você cria campos de índice personalizados correspondentes aos atributos personalizados do documento, que, por sua vez, correspondem aos tipos de entidade do Amazon Comprehend. O Amazon Kendra usa os campos de índice e os atributos personalizados do documento que você cria para pesquisar e filtrar os documentos.

Para obter mais informações, consulte Índice e criação de atributos de documentos personalizados.

Tópicos

- Criar um índice do Amazon Kendra
- Atualizar o perfil do IAM para acessar o Amazon S3
- Criação de campos de índice de pesquisa personalizados do Amazon Kendra
- Adicionar um bucket do Amazon S3 como fonte de dados para o índice
- Sincronizar o índice do Amazon Kendra

Criar um índice do Amazon Kendra

Para consultar os documentos de origem, crie um índice do Amazon Kendra.

Se você estiver usando o AWS CLI nesta etapa, você cria e anexa uma função e uma política AWS do IAM que permitem que a Amazon Kendra acesse CloudWatch seus registros antes de criar um índice. Para obter mais informações, consulte Pré-requisitos.

Para criar um índice do Amazon Kendra (console)

1. Abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/.

   > **Important**
   >
   > Certifique-se de estar na mesma região em que você criou o trabalho de análise de entidades do Amazon Comprehend e o bucket do Amazon S3. Se você estiver em outra região, escolha a AWS região em que você criou seu bucket do Amazon S3 no seletor de regiões na barra de navegação superior.

2. Escolha Criar índice.
3. Para Detalhes do índice na página Especificar detalhes do índice, faça o seguinte:
   a. Em Nome do índice, insira **`kendra-index`**.
   b. Mantenha o campo Descrição em branco.
   c. Em Perfil do IAM, selecione Criar uma função. Essa função fornece acesso ao bucket do Amazon S3.
   d. Em Nome do perfil, insira **`kendra-role`**. O perfil do IAM terá o prefixo `AmazonKendra-`.
   e. Mantenha as configurações padrão para Criptografia e Tags e escolha Avançar.
4. Para Configurações de controle de acesso na página Configurar controle de acesso do usuário, escolha Não e, em seguida, escolha Avançar.
5. Para Edições de aprovisionamento na página de Detalhes de aprovisionamento, escolha Developer Edition e escolha Criar.

Para criar um índice do Amazon Kendra (AWS CLI)

1. Para criar e anexar um perfil do IAM para a Amazon Kendra que a reconheça como uma entidade confiável, faça o seguinte:
   a. Salve a política de confiança a seguir como um arquivo JSON chamado `kendra-trust-policy.json` em um editor ou texto em seu computador.

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

b. Para criar um perfil do IAM chamado `kendra-role` e anexar o arquivo `kendra-trust-policy.json` salvo para ele, use o comando create-role:

Linux

```
aws iam create-role \
          --role-name kendra-role \
          --assume-role-policy-document file://path/kendra-trust-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `kendra-trust-policy.json` no dispositivo local.

macOS

```
aws iam create-role \
          --role-name kendra-role \
          --assume-role-policy-document file://path/kendra-trust-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `kendra-trust-policy.json` no dispositivo local.

Windows

```
aws iam create-role ^
```

```
        --role-name kendra-role ^
        --assume-role-policy-document file://path/kendra-trust-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `kendra-trust-policy.json` no dispositivo local.

c. Copie o nome do recurso da Amazon (ARN) no editor de texto e salve-o localmente como `kendra-role-arn`.

> **Note**
>
> O ARN tem um formato semelhante a *arn:aws:iam::123456789012:role/kendra-role*. Você precisa do ARN no qual você salvou `kendra-role-arn` para executar os trabalhos do Amazon Kendra.

2. Antes de criar um índice, você deve fornecer permissão para gravar `kendra-role` no CloudWatch Logs. Para fazer isso, conclua as seguintes etapas:

a. Salve a política de confiança a seguir como um arquivo JSON chamado `kendra-cloudwatch-policy.json` em um editor ou texto em seu computador.

```
{
   "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:*"
      }
   ]
}
```

Substitua *aws-region* pela sua região *aws-account-id*e pelo ID da AWS sua conta de 12 dígitos. AWS

b. Para criar uma política do IAM para acessar CloudWatch os registros, use o comando create-policy:

Linux

```
aws iam create-policy \
          --policy-name kendra-cloudwatch-policy \
          --policy-document file://path/kendra-cloudwatch-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `kendra-cloudwatch-policy.json` no dispositivo local.

macOS

```
aws iam create-policy \
          --policy-name kendra-cloudwatch-policy \
          --policy-document file://path/kendra-cloudwatch-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `kendra-cloudwatch-policy.json` no dispositivo local.

Windows

```
aws iam create-policy ^
          --policy-name kendra-cloudwatch-policy ^
          --policy-document file://path/kendra-cloudwatch-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `kendra-cloudwatch-policy.json` no dispositivo local.

c. Copie o nome do recurso da Amazon (ARN) no editor de texto e salve-o localmente como `kendra-cloudwatch-arn`.

> **Note**
>
> O ARN tem um formato semelhante ao arn:aws:iam: *:123456789012:role/*. kendra-cloudwatch-policy Você precisa do ARN em que salvou `kendra-cloudwatch-arn` para anexar `kendra-cloudwatch-policy` ao perfil do IAM.

d. Para anexar o `kendra-cloudwatch-policy` à sua função do IAM, use o attach-role-policycomando:

Linux

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

Em que:

- *policy-arn* é seu salvo `kendra-cloudwatch-arn`.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

Em que:

- *policy-arn* é seu salvo `kendra-cloudwatch-arn`.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name kendra-role
```

Em que:

- *policy-arn* é seu salvo `kendra-cloudwatch-arn`.

3. Para criar um índice, use o comando create-index:

Linux

```
aws kendra create-index \
        --name kendra-index \
        --edition DEVELOPER_EDITION \
        --role-arn role-arn \
        --region aws-region
```

Em que:

- *role-arn* é seu salvo `kendra-role-arn`,
- *aws-region* é sua região. AWS

macOS

```
aws kendra create-index \
```

```
        --name kendra-index \
        --edition DEVELOPER_EDITION \
        --role-arn role-arn \
        --region aws-region
```

Em que:

- *role-arn* é seu salvo `kendra-role-arn`,
- *aws-region* é sua região. AWS

Windows

```
aws kendra create-index ^
        --name kendra-index ^
        --edition DEVELOPER_EDITION ^
        --role-arn role-arn ^
        --region aws-region
```

Em que:

- *role-arn* é seu salvo `kendra-role-arn`,
- *aws-region* é sua região. AWS

4. Copie o índice `Id` e salve-o em um editor de texto como `kendra-index-id`. `Id` ajuda a rastrear o status da criação do índice.
5. Para acompanhar o progresso do trabalho de criação de índice, use o comando describe-index:

   Linux

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

   Em que:

   - *kendra-index-id*é seu salvo`kendra-index-id`,
   - *aws-region* é sua região. AWS

macOS

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

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Windows

```
aws kendra describe-index ^
        --id kendra-index-id ^
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

O processo de criação do índice leva, em média, 15 minutos, mas pode levar mais tempo. Quando o status do índice é ativo, o índice está pronto para uso. Enquanto o índice está sendo criado, você pode começar a próxima etapa.

Se você estiver usando o AWS CLI nesta etapa, você cria e anexa uma política do IAM à sua função do Amazon Kendra IAM que concede ao seu índice permissões para acessar seu bucket do S3.

Atualizar o perfil do IAM para acessar o Amazon S3

Enquanto o índice está sendo criado, você atualiza o perfil do Amazon Kendra IAM para permitir que o índice que você criou leia dados do bucket do Amazon S3. Para obter mais informações, consulte Perfis do IAM para o Amazon Kendra.

Para atualizar seu perfil do IAM (console)

1. Abra o console IAM em https://console.aws.amazon.com/iam/.
2. No painel de navegação esquerdo, escolha Funções e insira **`kendra-role`** na caixa Pesquisar acima do Nome da função.
3. Nas opções sugeridas, clique em `kendra-role`.
4. Em Resumo, escolha Anexar políticas.
5. Em Anexar permissões, na caixa Pesquisar, insira **`S3`** e selecione a caixa de seleção ao lado da ReadOnlyAccess política do AmazonS3 nas opções sugeridas.
6. Escolha Anexar política. Na página de Resumo, agora você verá duas políticas anexadas ao perfil do IAM.
7. Retorne ao console do Amazon Kendra em https://console.aws.amazon.com/kendra/ e aguarde até que o status do índice mude de Criando para Ativo antes de continuar com a próxima etapa.

Para atualizar o perfil do IAM (console)

1. Salve a política de confiança a seguir como um arquivo JSON chamado `kendra-S3-access-policy.json` em um editor ou texto em seu computador.

```
{
   "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"
         ]
      }
   ]
}
```

Substitua DOC-EXAMPLE-BUCKET pelo nome do bucket do S3, *aws-region pela sua região*, pelo ID da conta de 12 dígitos e pelo que AWS você *aws-account-id*salvou. AWS *kendra-index-id*kendra-index-id

2. Para criar uma política do IAM chamada para acessar o bucket do S3, use o comando create-policy:

   Linux

   ```
   aws iam create-policy \
             --policy-name kendra-S3-access-policy \
             --policy-document file://path/kendra-S3-access-policy.json
   ```

   Em que:

   - *path/* é o caminho do arquivo para a pasta kendra-S3-access-policy.json no dispositivo local.

   macOS

   ```
   aws iam create-policy \
             --policy-name kendra-S3-access-policy \
             --policy-document file://path/kendra-S3-access-policy.json
   ```

   Em que:

- *path/* é o caminho do arquivo para a pasta `kendra-S3-access-policy.json` no dispositivo local.

Windows

```
aws iam create-policy ^
          --policy-name kendra-S3-access-policy ^
          --policy-document file://path/kendra-S3-access-policy.json
```

Em que:

- *path/* é o caminho do arquivo para a pasta `kendra-S3-access-policy.json` no dispositivo local.

3. Copie o nome do recurso da Amazon (ARN) no editor de texto e salve-o localmente como `kendra-S3-access-arn`.

> **Note**
>
> O ARN tem um formato semelhante a *arn:aws:iam::123456789012:role/kendra-S3-access-policy*. Você precisa do ARN em que salvou `kendra-S3-access-arn` para anexar `kendra-S3-access-policy` ao perfil do IAM.

4. Para anexar o `kendra-S3-access-policy` à sua função do Amazon Kendra IAM, use o comando: attach-role-policy

Linux

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

Em que:

- *policy-arn* é seu salvo `kendra-S3-access-arn`.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

Em que:

- *policy-arn* é seu salvo `kendra-S3-access-arn`.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name kendra-role
```

Em que:

- *policy-arn* é seu salvo `kendra-S3-access-arn`.

Criação de campos de índice de pesquisa personalizados do Amazon Kendra

Para preparar o Amazon Kendra para reconhecer os metadados como atributos personalizados do documento, crie campos personalizados correspondentes aos tipos de entidade do Amazon Comprehend. Insira os nove tipos de entidade do Amazon Comprehend a seguir como campos personalizados:

- COMMERCIAL_ITEM
- DATA
- EVENTO
- LOCALIZAÇÃO
- ORGANIZAÇÃO
- OUTRO
- PESSOA

- QUANTIDADE
- TITLE

> **Important**
>
> Tipos de entidade com erros ortográficos não serão reconhecidos pelo índice.

Para criar campos personalizados para seu índice do Amazon Kendra (console)

1. Abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/.
2. Na lista de Índices, clique em `kendra-index`.
3. No painel de navegação esquerdo, em Gerenciamento de dados, escolha Definição de faceta.
4. No menu Campos do índice, escolha Adicionar campo.
5. Na caixa de diálogo Adicionar campo de índice, faça o seguinte:
   a. No campo Nome, insira **`COMMERCIAL_ITEM`**.
   b. Em Tipo de dados, escolha Lista de cadeias de caracteres.
   c. Em Tipos de uso, selecione Facetável, Pesquisável e Exibível e, em seguida, escolha Adicionar.
   d. Repita as etapas de a a c para cada tipo de entidade do Amazon Comprehend: COMMERCIAL_ITEM, DATE, EVENT, LOCATION, ORGANIZATION, OTHER, PERSON, QUANTITY, TITLE.

O console exibe mensagens de adição de campo bem-sucedida. Você pode optar por fechá-los antes de prosseguir para a próxima etapa.

Para criar campos personalizados para o índice do Amazon Kendra (AWS CLI)

1. Salve a política de confiança a seguir como um arquivo JSON chamado `custom-attributes.json` em um editor ou texto em seu computador.

```
[
   {
       "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 criar campos personalizados no índice, use o comando update-index:

   Linux

   ```
   aws kendra update-index \
           --id kendra-index-id \
           --document-metadata-configuration-updates file://path/custom-
   attributes.json \
   ```

```
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *path/* é o caminho do arquivo para custom-attributes.json no dispositivo local,
- *aws-region* é sua região. AWS

macOS

```
aws kendra update-index \
        --id kendra-index-id \
        --document-metadata-configuration-updates file://path/custom-
attributes.json \
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *path/* é o caminho do arquivo para custom-attributes.json no dispositivo local,
- *aws-region* é sua região. AWS

Windows

```
aws kendra update-index ^
        --id kendra-index-id ^
        --document-metadata-configuration-updates file://path/custom-
attributes.json ^
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *path/* é o caminho do arquivo para custom-attributes.json no dispositivo local,
- *aws-region* é sua região. AWS

3. Para verificar se os atributos personalizados foram adicionados ao índice, use o comando describe-index:

Linux

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

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

macOS

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

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Windows

```
aws kendra describe-index ^
        --id kendra-index-id ^
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Adicionar um bucket do Amazon S3 como fonte de dados para o índice

Antes de sincronizar o índice, conecte a fonte de dados do S3 a ele.

Para conectar um bucket do S3 ao índice do Amazon Kendra (console)

1. Abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/.
2. Na lista de Índices, clique em `kendra-index`.
3. No menu de navegação à esquerda, em Gerenciamento de dados, escolha Fontes de dados.
4. Na seção Selecionar tipo de conector da fonte de dados, navegue até Amazon S3 e escolha Adicionar conector.
5. Na página Especificar detalhes da fonte de dados, faça o seguinte:
   a. Em Nome e descrição, para Nome da fonte de dados, insira um **`S3-data-source`**.
   b. Mantenha a seção Descrição em branco.
   c. Mantenha a configuração padrão para Tags.
   d. Escolha Próximo.
6. Na página Definir configurações, na seção Escopo de sincronização, faça o seguinte:
   a. Em Inserir o local da fonte de dados, escolha Procurar no S3.
   b. Em Escolher recursos, selecione so bucket do S3 e escolha Escolher.
   c. Em Localização da pasta de prefixo de arquivos de metadados, escolha Procurar S3.
   d. Em Escolher recursos, clique no nome do bucket na lista de buckets.
   e. Em Objetos, selecione o botão de opção para `metadata` e escolha Escolher. O campo de localização agora deve dizer `metadata/`.
   f. Mantenha as configurações padrão para a Localização do arquivo de configuração da lista de controle de acesso, Selecionar chave de descriptografia e Configuração adicional.
7. Para o perfil do IAM, na página Definir configurações de sincronização, escolha `kendra-role`.
8. Na página Definir configurações de sincronização, em Agenda de execução da sincronização, em Frequência, escolha Executar sob demanda e, em seguida, escolha Avançar.
9. Na página Revisar e criar, analise os detalhes da fonte de dados e escolha Criar fonte de dados.

Para conectar um bucket do S3 ao índice do Amazon Kendra (AWS CLI)

1. Salve a política de confiança a seguir como um arquivo JSON chamado `S3-data-connector.json` em um editor ou texto em seu computador.

```
{
   "S3Configuration":{
```

```
      "BucketName":"DOC-EXAMPLE-BUCKET",
      "DocumentsMetadataConfiguration":{
         "S3Prefix":"metadata"
      }
   }
}
```

Substitua DOC-EXAMPLE-BUCKET pelo nome do bucket do S3.

2. Para conectar seu bucket do S3 ao seu índice, use o 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
   ```

   Em que:

   - *kendra-index-id*é seu salvokendra-index-id,
   - *path/* é o caminho do arquivo para S3-data-connector.json no dispositivo local,
   - *role-arn* é seu salvo kendra-role-arn,
   - *aws-region* é sua região. 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
   ```

   Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *path/* é o caminho do arquivo para S3-data-connector.json no dispositivo local,
- *role-arn* é seu salvo kendra-role-arn,
- *aws-region* é sua região. 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
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *path/* é o caminho do arquivo para S3-data-connector.json no dispositivo local,
- *role-arn* é seu salvo kendra-role-arn,
- *aws-region* é sua região. AWS

3. Copie o índice Id e salve-o em um editor de texto como S3-connector-id. Id ajuda você a rastrear o status do processo de conexão de dados.
4. Para garantir que sua fonte de dados do S3 tenha sido conectada com êxito, use o describe-data-sourcecomando:

   Linux

   ```
   aws kendra describe-data-source \
           --id S3-connector-id \
           --index-id kendra-index-id \
           --region aws-region
   ```

   Em que:

   - *O S3-Connector-ID* é seu S3-connector-id salvo,

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

macOS

```
aws kendra describe-data-source \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

Em que:

- *O S3-Connector-ID* é seu S3-connector-id salvo,
- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Windows

```
aws kendra describe-data-source ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

Em que:

- *O S3-Connector-ID* é seu S3-connector-id salvo,
- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Ao final dessa etapa, a fonte de dados do Amazon S3 é conectada ao índice.

Sincronizar o índice do Amazon Kendra

Com a fonte de dados do Amazon S3 adicionada, agora você sincroniza o índice Amazon Kendra com ela.

Para sincronizar um índice do Amazon Kendra (console)

1. Abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/.
2. Na lista de Índices, clique em `kendra-index`.
3. No menu de navegação à esquerda, escolha Fontes de dados.
4. Em Fontes de dados, selecione `S3-data-source`.
5. Na barra de navegação superior, escolha Sincronizar agora.

Para sincronizar um índice do Amazon Kendra (AWS CLI)

1. Para sincronizar seu índice, use o 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
   ```

   Em que:

   - *O S3-Connector-ID* é seu `S3-connector-id` salvo,
   - *kendra-index-id*é seu salvo`kendra-index-id`,
   - *aws-region* é sua região. AWS

   macOS

   ```
   aws kendra start-data-source-sync-job \
           --id S3-connector-id \
           --index-id kendra-index-id \
           --region aws-region
   ```

   Em que:

   - *O S3-Connector-ID* é seu `S3-connector-id` salvo,
   - *kendra-index-id*é seu salvo`kendra-index-id`,
   - *aws-region* é sua região. AWS

Windows

```
aws kendra start-data-source-sync-job ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

Em que:

- *O S3-Connector-ID* é seu S3-connector-id salvo,
- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

2. Para verificar o status da sincronização do índice, use o 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
```

Em que:

- *O S3-Connector-ID* é seu S3-connector-id salvo,
- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

macOS

```
aws kendra list-data-source-sync-jobs \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

Em que:

- *O S3-Connector-ID* é seu S3-connector-id salvo,

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Windows

```
aws kendra list-data-source-sync-jobs ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

Em que:

- *O S3-Connector-ID* é seu S3-connector-id salvo,
- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Ao final desta etapa, você criou um índice Amazon Kendra pesquisável e filtrável para o conjunto de dados.

Etapa 5: consultar o índice do Amazon Kendra

O índice do Amazon Kendra agora está pronto para consultas em linguagem natural. Ao pesquisar o índice, o Amazon Kendra usa todos os dados e metadados fornecidos para retornar as respostas mais precisas à consulta de pesquisa.

Há três tipos de consultas que a Amazon Kendra pode responder:

- Consultas factóides (perguntas sobre “quem”, “o quê”, “quando” ou “onde”)
- Consultas descritivas (perguntas do tipo “como”)
- Pesquisas por palavra-chave (perguntas cuja intenção e escopo não são claros)

Tópicos

- Consulte o índice do Amazon Kendra
- Filtrar os resultados de pesquisa

Consulte o índice do Amazon Kendra

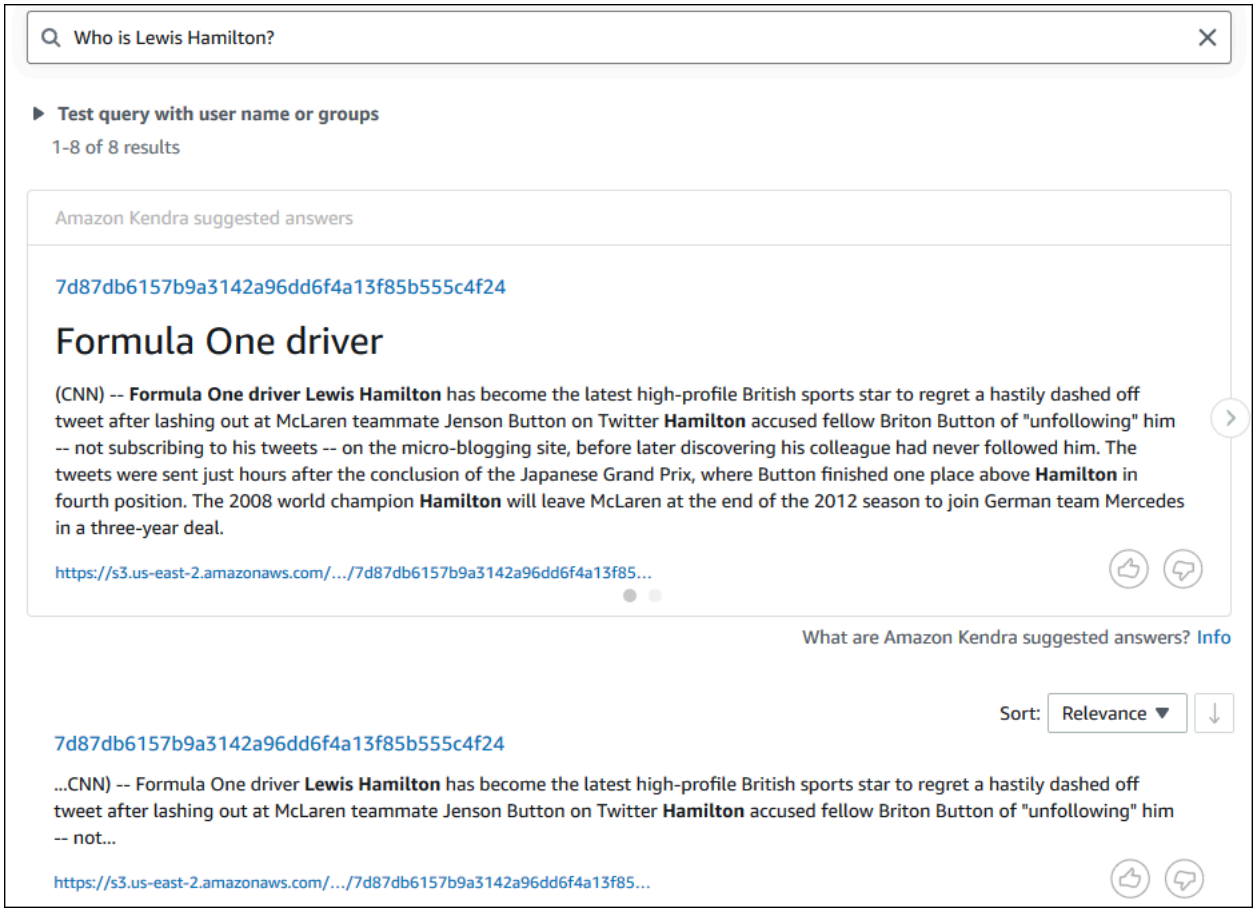
Você pode consultar o índice do Amazon Kendra usando perguntas que correspondem aos três tipos de consultas compatíveis com o Amazon Kendra. Para obter mais informações consulte Consultas.

As perguntas de exemplo nesta seção foram escolhidas com base no conjunto de dados de amostra.

Para consultar um índice do Amazon Kendra (console)

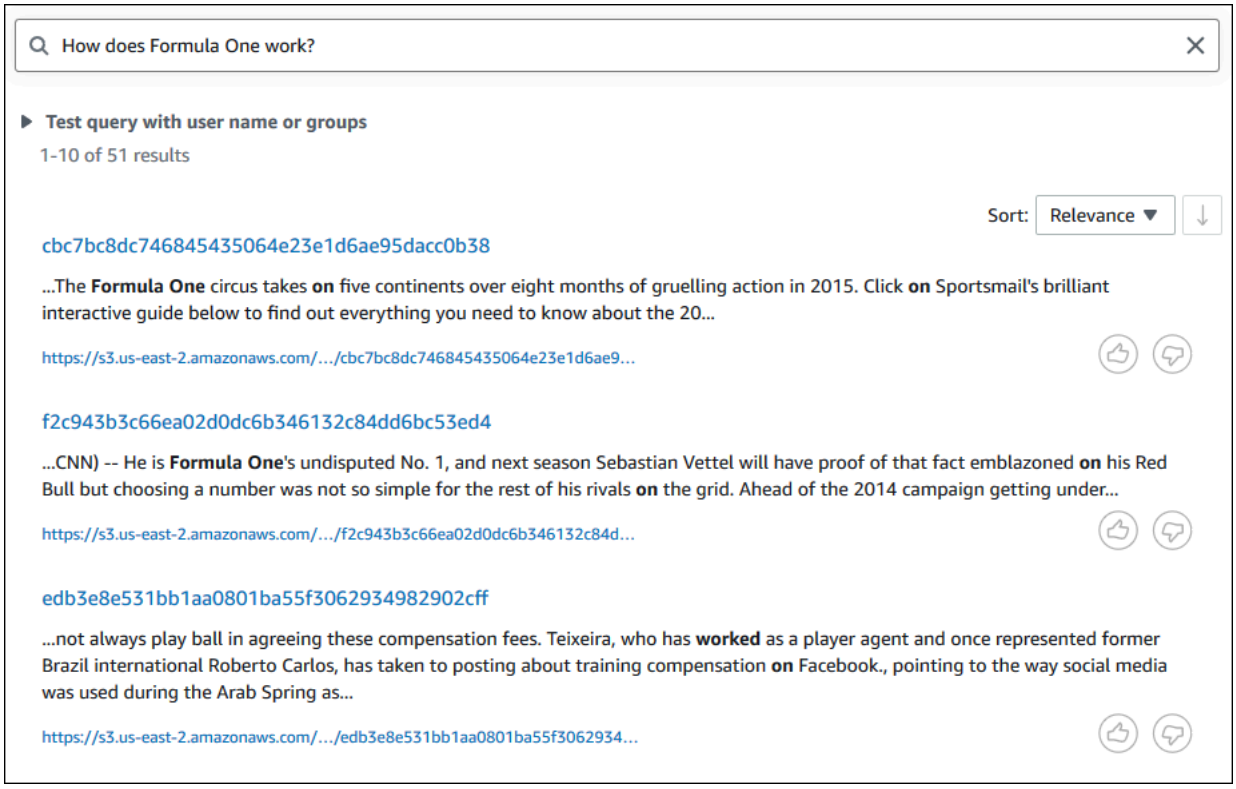
1. Abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/.
2. Na lista de Índices, clique em `kendra-index`.
3. No menu de navegação à esquerda, escolha a opção de pesquisar no índice.
4. Para executar um exemplo de consulta de factoide, insira **`Who is Lewis Hamilton?`** na caixa de pesquisa e pressione enter.

   O primeiro resultado retornado é a resposta sugerida pela Amazon Kendra, junto com o arquivo de dados contendo a resposta. O restante dos resultados forma o conjunto de documentos recomendados.



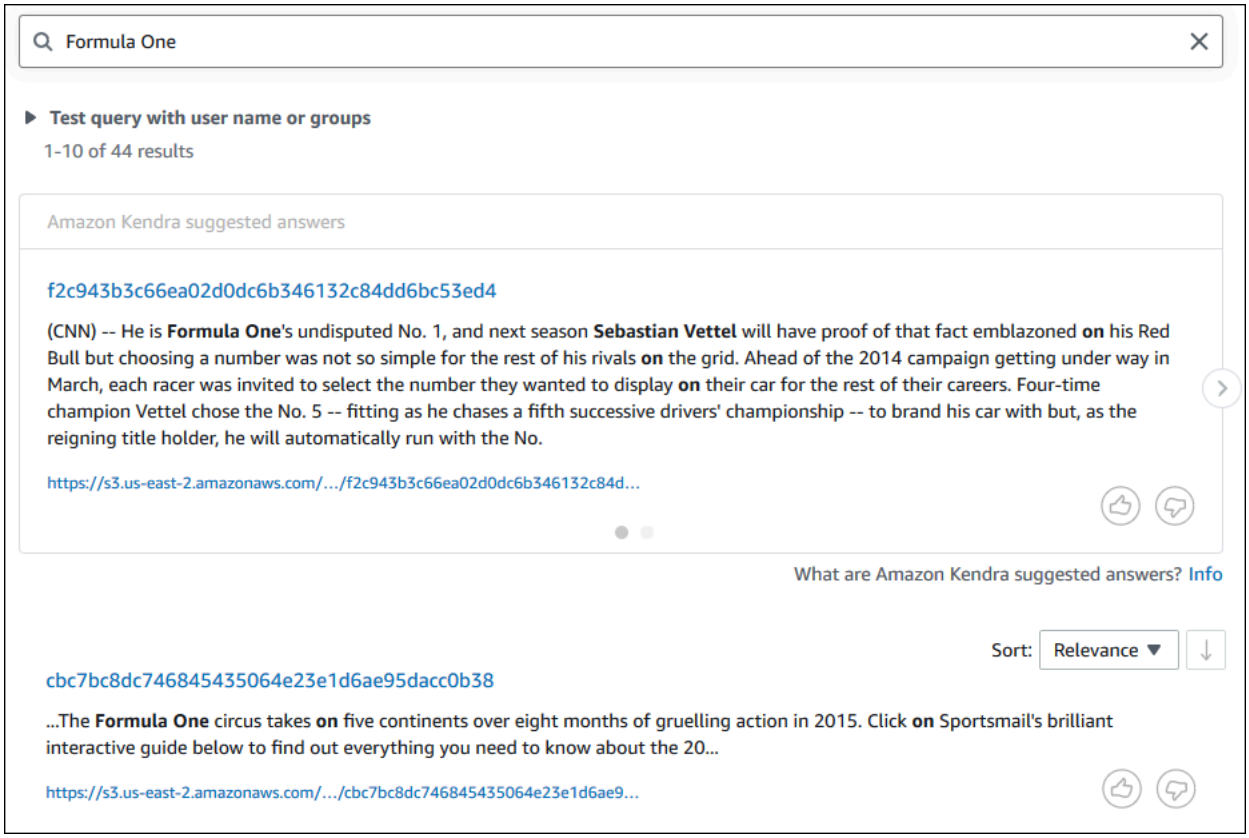
5. Para executar um exemplo de consulta, insira **`How does Formula One work?`** na caixa de pesquisa e pressione enter.

   Você verá outro resultado retornado pelo console Amazon Kendra, desta vez com a frase relevante destacada.



6. Para executar uma pesquisa por palavra-chave, insira **`Formula One`** na caixa de pesquisa e pressione enter.

   Você verá outro resultado retornado pelo console do Amazon Kendra, seguido pelos resultados de todas as outras menções da frase no conjunto de dados.



Para consultar um índice do Amazon Kendra (AWS CLI)

1. Para executar um exemplo de consulta factoide, use o comando consulta:

   Linux

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Who is Lewis Hamilton?" \
           --region aws-region
   ```

   Em que:

   - *kendra-index-id*é seu salvokendra-index-id,
   - *aws-region* é sua região. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Who is Lewis Hamilton?" \
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Who is Lewis Hamilton?" ^
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

AWS CLI Exibe os resultados da sua consulta.

2. Para executar um exemplo de consulta descritiva, use o comando consulta:

Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "How does Formula One work?" \
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,

- *aws-region* é sua região. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "How does Formula One work?" \
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "How does Formula One work?" ^
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

AWS CLI Exibe os resultados da sua consulta.

3. Para executar uma amostra de pesquisa por palavra-chave, use o comando consulta:

Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Formula One" \
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Formula One" \
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Formula One" ^
        --region aws-region
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

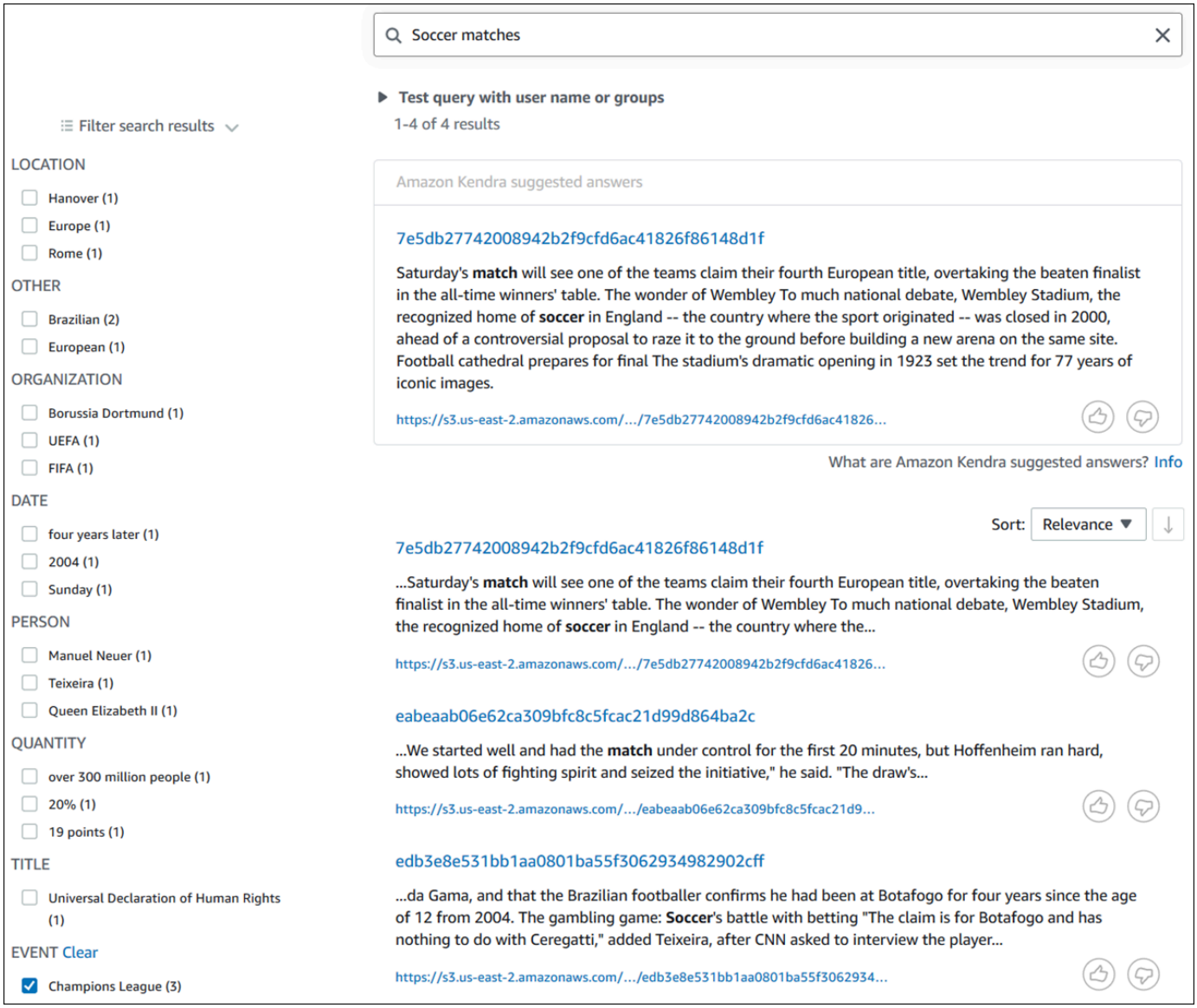
AWS CLI Exibe as respostas retornadas à sua consulta.

Filtrar os resultados de pesquisa

Você pode filtrar e classificar os resultados de pesquisa usando atributos de documentos personalizados no console do Amazon Kendra. Para obter mais informações sobre como o Amazon Kendra processa consultas, consulte Filtragem de consultas.

Para filtrar os resultados da pesquisa (console)

1. Abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/.
2. Na lista de Índices, clique em `kendra-index`.
3. No menu de navegação à esquerda, escolha a opção de pesquisar no índice.
4. Na caixa de pesquisa, insira **`Soccer matches`** como uma consulta e pressione enter.
5. No menu de navegação à esquerda, escolha Filtrar resultados da pesquisa para ver uma lista das facetas que você pode usar para filtrar a pesquisa.
6. Marque a caixa de seleção “Liga dos Campeões” sob o subtítulo EVENTO, para ver os resultados da pesquisa filtrados somente pelos resultados que contêm “Liga dos Campeões”.



Para filtrar os resultados da pesquisa (console)

1. Para ver as entidades de um tipo específico (como EVENT) que estão disponíveis para uma pesquisa, use o comando consulta:

   Linux

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Soccer matches" \
           --facets '[{"DocumentAttributeKey":"EVENT"}]' \
           --region aws-region
   ```

   Em que:

   - *kendra-index-id*é seu salvokendra-index-id,
   - *aws-region* é sua região. AWS

   macOS

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Soccer matches" \
           --facets '[{"DocumentAttributeKey":"EVENT"}]' \
           --region aws-region
   ```

   Em que:

   - *kendra-index-id*é seu salvokendra-index-id,
   - *aws-region* é sua região. AWS

   Windows

   ```
   aws kendra query ^
           --index-id kendra-index-id ^
           --query-text "Soccer matches" ^
           --facets '[{"DocumentAttributeKey":"EVENT"}]' ^
           --region aws-region
   ```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

AWS CLI Exibe os resultados da pesquisa. Para obter uma lista de facetas do tipoEVENT, navegue até a seção "FacetResults" da AWS CLI saída para ver uma lista de facetas filtráveis com suas contagens. Por exemplo, uma das facetas é a “Liga dos Campeões”.

> **Note**
>
> Em vez de EVENT, você pode escolher qualquer um dos campos de índice que você criou em the section called “Criar um índice do Amazon Kendra” para o valor DocumentAttributeKey.

2. Para executar a mesma pesquisa, mas filtrar somente pelos resultados que contêm “Liga dos Campeões”, use o comando de consulta:

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
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. 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
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. 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
```

Em que:

- *kendra-index-id*é seu salvokendra-index-id,
- *aws-region* é sua região. AWS

AWS CLI Exibe os resultados da pesquisa filtrada.

Etapa 5: limpar

Como limpar os arquivos

Para parar de incorrer em cobranças em sua AWS conta depois de concluir este tutorial, você pode seguir as seguintes etapas:

1. Exclua o bucket do Amazon S3

   Para obter informações sobre como excluir um bucket, consulte Excluir um bucket.

2. Exclua o índice do Amazon Kendra

Para obter informações sobre como excluir um índice do Amazon Kendra, consulte Excluir um índice.

3. Excluir **`converter.py`**
   - Para console: acesse e verifique se a região está definida como sua AWS região. AWS CloudShell Depois que o shell bash for carregado, digite o seguinte comando no ambiente e pressione enter:

     ```
     rm converter.py
     ```

   - Para AWS CLI: Execute o comando a seguir em uma janela de terminal.

     Linux

     ```
     rm file/converter.py
     ```

     Em que:

     - *`path/`* é o caminho do arquivo para `converter.py` no dispositivo local.

     macOS

     ```
     rm file/converter.py
     ```

     Em que:

     - *`path/`* é o caminho do arquivo para `converter.py` no dispositivo local.

     Windows

     ```
     rm file/converter.py
     ```

     Em que:

     - *`path/`* é o caminho do arquivo para `converter.py` no dispositivo local.

Saiba mais

Para saber mais sobre a integração do Amazon Kendra no fluxo de trabalho, você pode conferir as seguintes publicações do blog:

- Marcação de metadados de conteúdo para pesquisa aprimorada

- Crie uma solução de pesquisa inteligente com enriquecimento automatizado de conteúdo

Para saber mais, consulte O que é o Amazon Comprehend? no Guia do desenvolvedor do Amazon Comprehend.

Monitorar e registro em log para Amazon Kendra

Tópicos

- Como monitorar seu índice (console)
- Registro de chamadas de API do Amazon Kendra com os logs do AWS CloudTrail
- Registro de chamadas da API do Amazon Kendra Intelligent Ranking com logs do AWS CloudTrail
- Monitoramento do Amazon Kendra com o Amazon CloudWatch
- Monitoramento do Amazon Kendra com o Amazon CloudWatch Logs

Como monitorar seu índice (console)

Use o console do Amazon Kendra para monitorar o estado dos índices e das fontes de dados. Você pode usar essas informações para rastrear o tamanho e os requisitos de armazenamento do índice e monitorar o progresso e o sucesso da sincronização entre o índice e as fontes de dados.

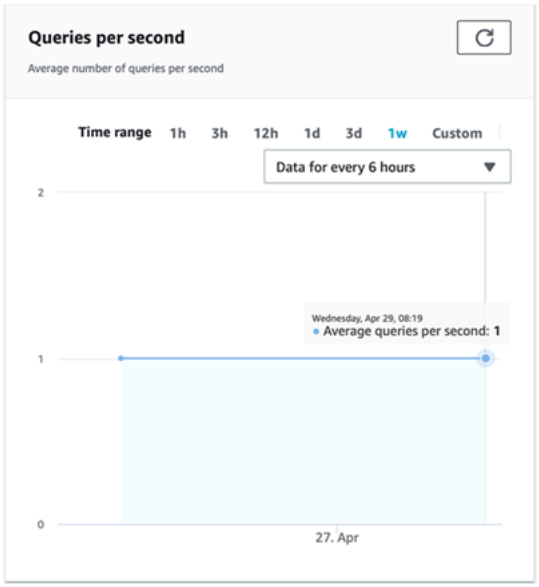
Para visualizar métricas do índice (console)

1. Faça login no AWS Management Console e abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/home.
2. Na lista de índices, escolha o índice a ser visualizado.
3. Role a tela para ver as métricas do índice.

Você pode ver as seguintes métricas sobre o índice.

- Contagem de documentos: o número total de documentos indexados. Isso inclui todos os documentos de todas as fontes de dados. Use essa métrica para determinar se você precisa comprar mais ou menos unidades de armazenamento para o índice.

- Consultas por segundo: o número de consultas de índice que são solicitadas a cada segundo. Use essa métrica para determinar se você precisa comprar mais ou menos unidades de consulta para o índice.



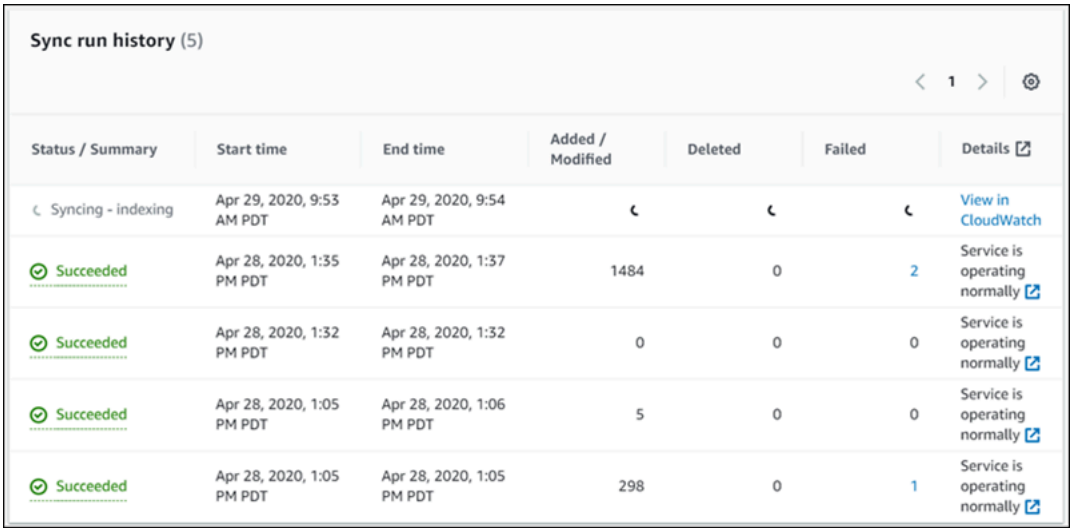
Para monitorar o progresso e o sucesso da sincronização entre o índice e uma fonte de dados, use o console Amazon Kendra. Use essas informações para ajudar a determinar a integridade da fonte de dados.

Para visualizar métricas de sincronização (console)

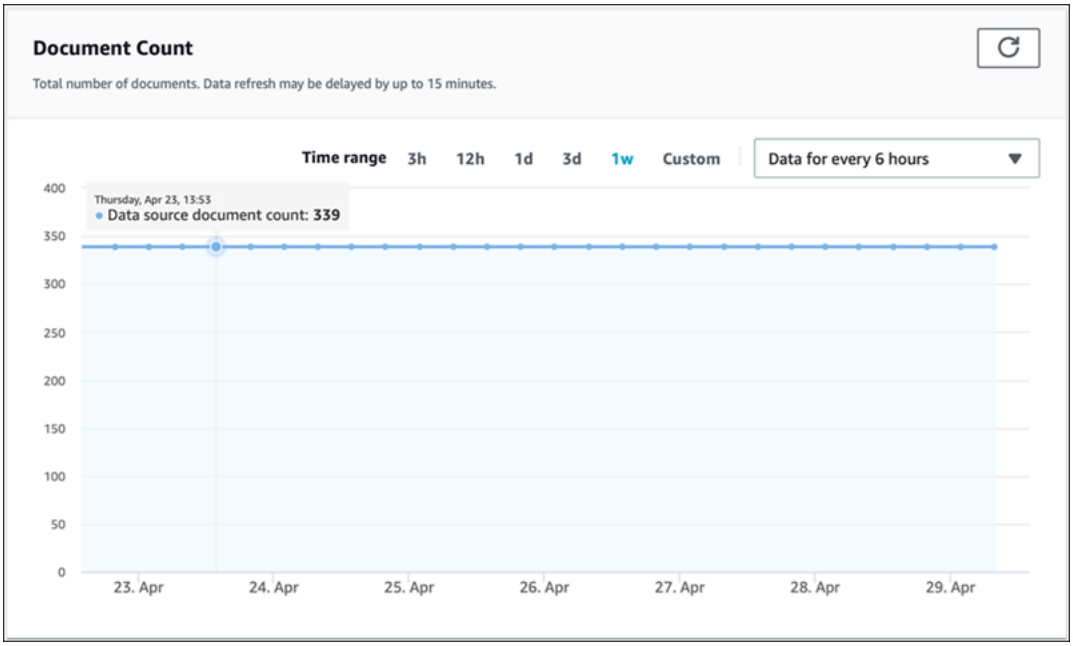
1. Faça login no AWS Management Console e abra o console do Amazon Kendra em https://console.aws.amazon.com/kendra/home.
2. Na lista de índices, escolha o índice para visualizar métricas de sincronização.
3. No menu á esquerda, escolha Fontes de dados.
4. Na lista de fontes dos dados, escolha o nome da fonte de dados que deseja exibir.
5. Role a tela para ver as métricas de execução de sincronização.

Você pode editar as seguintes informações.

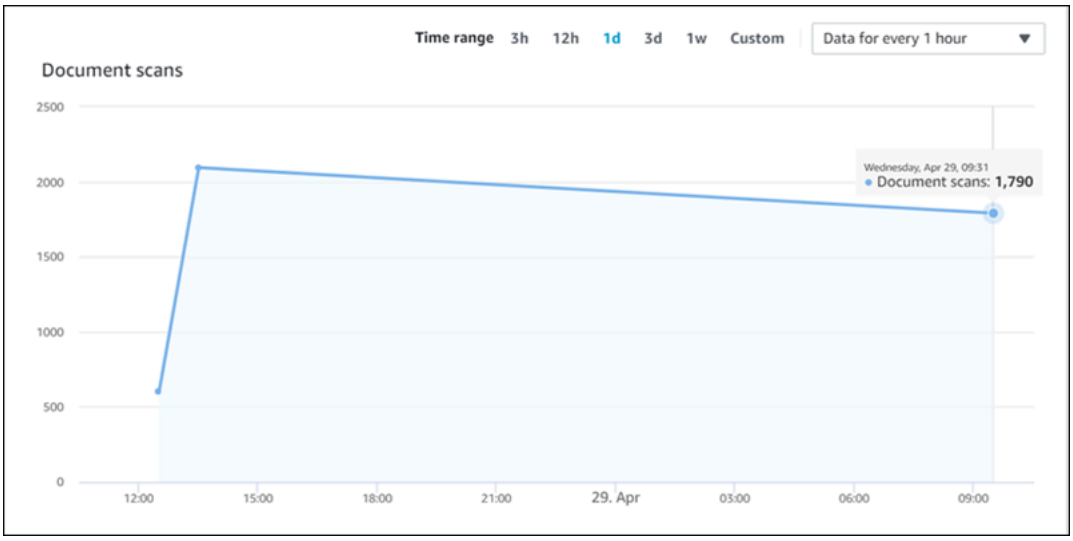
- Histórico de execução da sincronização: estatísticas sobre a execução da sincronização, incluindo a hora de início e término, o número de documentos adicionados, excluídos e que falharam. Se a execução da sincronização falhar, há um link para o CloudWatch Logs com mais informações. Escolha o ícone de configurações no canto superior esquerdo para alterar as colunas que são exibidas no histórico. Use essas informações para determinar a integridade da fonte de dados.



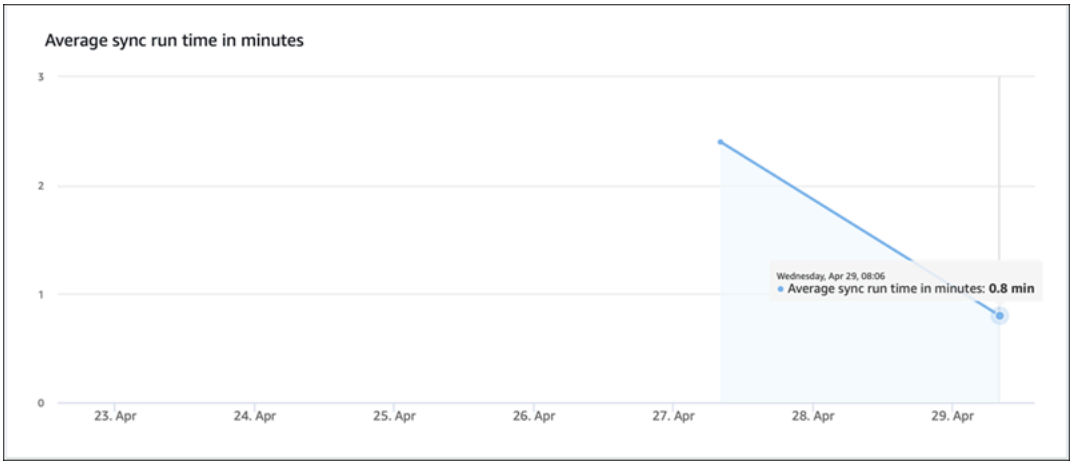
- Contagem de documentos: o número total de documentos indexados dessa fonte de dados. Esse é o total de todos os documentos adicionados à fonte de dados menos o total de todos os documentos excluídos da fonte de dados. Use essas informações para determinar quantos documentos dessa fonte de dados estão incluídos no índice.



- Digitalizações de documentos: o número total de documentos digitalizados durante a execução da sincronização. Isso inclui todos os documentos na fonte de dados, incluindo aqueles adicionados, atualizados, excluídos ou inalterados. Use essas informações para determinar se o Amazon Kendra está digitalizando todos os documentos na fonte de dados. O número de documentos digitalizados afeta o valor cobrado pelo serviço.



- Tempo médio de execução da sincronização em minutos: o tempo médio necessário para que uma execução de sincronização seja concluída. O tempo necessário para sincronizar uma fonte de dados afeta o valor cobrado pelo serviço.



Registro de chamadas de API do Amazon Kendra com os logs do AWS CloudTrail

O Amazon Kendra é integrado ao AWS CloudTrail, um serviço que fornece um registro das ações realizadas por um usuário, por uma função ou por um serviço da AWS no Amazon Kendra. O CloudTrail captura todas as chamadas de API do Amazon Kendra como eventos, inclusive as chamadas do console do Amazon Kendra e de chamadas de código para as APIs do Amazon Kendra. Se você criar uma trilha, poderá ativar a entrega contínua de eventos do CloudTrail para um bucket do Amazon S3, inclusive eventos para o Amazon Kendra. Se você não configurar uma trilha, ainda poderá visualizar os eventos mais recentes no console do CloudTrail em Histórico de eventos. Usando as informações coletadas pelo CloudTrail, é possível determinar a solicitação feita para o Amazon Kendra, o endereço IP do qual a solicitação foi feita, quem fez a solicitação, quando ela foi feita e detalhes adicionais.

Para saber mais sobre o CloudTrail, incluindo como configurá-lo e ativá-lo, consulte o Guia do usuário do AWS CloudTrail.

Informações sobre o Amazon Kendra no CloudTrail

O CloudTrail é ativado em sua conta da AWS quando ela é criada. Quando ocorre uma atividade no Amazon Kendra, essa atividade é registrada em um evento do CloudTrail junto com outros eventos de serviços da AWS em Histórico de eventos do CloudTrail. Você pode visualizar, pesquisar e baixar eventos recentes em sua conta da AWS. Para obter mais informações, consulte Como visualizar eventos com o histórico de eventos do CloudTrail.

Para obter um registro de eventos em andamento na sua conta da AWS, incluindo eventos do Amazon Kendra, crie uma trilha. Uma trilha é uma configuração que permite ao CloudTrail entregar eventos como arquivos de log a um bucket do S3 especificado. Por padrão, quando você cria uma trilha no console, ela é aplicada a todas as regiões da AWS. A trilha registra eventos de todas as regiões na partição da AWS e fornece os arquivos de log ao bucket do S3 que você especificar. Além disso, é possível configurar outros serviços da AWS para analisar e agir melhor com base nos dados de eventos coletados nos logs do CloudTrail. Para obter mais informações, consulte:

- Visão geral da criação de uma trilha
- Serviços e integrações compatíveis com o CloudTrail
- Configurar notificações do Amazon SNS para o CloudTrail
- Receber arquivos de log do CloudTrail de várias regiões e receber arquivos de log do CloudTrail de várias contas

O CloudTrail registra todas as ações do Amazon Kendra, que estão documentadas na Referência da API. Por exemplo, as chamadas para as operações `CreateIndex`, `CreateDataSource` e `Query` geram entradas nos arquivos de log do CloudTrail.

Cada entrada de log ou evento contém informações sobre quem gerou a solicitação. Para obter mais informações, consulte o Elemento de identidade do usuário do CloudTrail.

Exemplo: entradas de arquivo de log do Amazon Kendra

Uma trilha é uma configuração que permite a entrega de eventos como arquivos de log a um bucket do S3 especificado. Os arquivos de log do CloudTrail contêm uma ou mais entradas de log. Um evento representa uma única solicitação de qualquer origem e inclui informações sobre a ação solicitada, a data e a hora da ação, os parâmetros de solicitação e assim por diante. Os arquivos de log do CloudTrail não são um rastreamento de pilha ordenada de chamadas de API pública. Dessa forma, eles não são exibidos em uma ordem específica.

As chamadas para a operação de `Query` criam a seguinte 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 chamadas da API do Amazon Kendra Intelligent Ranking com logs do AWS CloudTrail

O AAmazon Kendra Intelligent Ranking é integrado ao AWS CloudTrail, serviço que fornece um registro das ações executadas por um usuário, uma função ou um serviço da AWS no Amazon Kendra Intelligent Ranking. O CloudTrail captura todas as chamadas de API do Amazon Kendra

Intelligent Ranking como eventos, incluindo as chamadas de código para as APIs do Amazon Kendra Intelligent Ranking. Se você criar uma trilha, poderá ativar a entrega contínua de eventos do CloudTrail para um bucket do Amazon S3, incluindo eventos para o Amazon Kendra Intelligent Ranking. Se você não configurar uma trilha, ainda poderá visualizar os eventos mais recentes no console do CloudTrail em Histórico de eventos. Usando as informações coletadas pelo CloudTrail, é possível determinar a solicitação feita para o Amazon Kendra Intelligent Ranking, o endereço IP do qual a solicitação foi feita, quem fez a solicitação, quando ela foi feita e detalhes adicionais.

Para saber mais sobre o CloudTrail, incluindo como configurá-lo e ativá-lo, consulte o Guia do usuário do AWS CloudTrail.

Informações do Amazon Kendra Intelligent Ranking no CloudTrail

O CloudTrail é ativado em sua conta da AWS quando ela é criada. Quando ocorre qualquer atividade no Amazon Kendra Intelligent Ranking, essa atividade é registrada em um evento do CloudTrail junto com outros eventos de serviços da AWS no Histórico de eventos do CloudTrail. Você pode visualizar, pesquisar e baixar eventos recentes em sua conta da AWS. Para obter mais informações, consulte Como visualizar eventos com o histórico de eventos do CloudTrail.

Para obter um registro contínuo de eventos da conta da AWS, incluindo eventos do Amazon Kendra Intelligent Ranking, crie uma trilha. Uma trilha é uma configuração que permite ao CloudTrail entregar eventos como arquivos de log a um bucket do S3 especificado. Por padrão, quando você cria uma trilha no console, ela é aplicada a todas as regiões da AWS. A trilha registra eventos de todas as regiões na partição da AWS e fornece os arquivos de log ao bucket do S3 que você especificar. Além disso, é possível configurar outros serviços da AWS para analisar e agir melhor com base nos dados de eventos coletados nos logs do CloudTrail. Para obter mais informações, consulte:

- Visão geral da criação de uma trilha
- Serviços e integrações compatíveis com o CloudTrail
- Configurar notificações do Amazon SNS para o CloudTrail
- Receber arquivos de log do CloudTrail de várias regiões e receber arquivos de log do CloudTrail de várias contas

O CloudTrail registra todas as ações do Amazon Kendra Intelligent Ranking, que estão documentadas na Referência da API. Por exemplo, chamadas para o `CreateRescoreExecutionPlan` gerar entradas nos arquivos de log do CloudTrail.

Cada entrada de log ou evento contém informações sobre quem gerou a solicitação. Para obter mais informações, consulte o Elemento de identidade do usuário do CloudTrail.

Exemplo: entradas de arquivos de log do Amazon Kendra Intelligent Ranking

Uma trilha é uma configuração que permite a entrega de eventos como arquivos de log a um bucket do S3 especificado. Os arquivos de log do CloudTrail contêm uma ou mais entradas de log. Um evento representa uma única solicitação de qualquer origem e inclui informações sobre a ação solicitada, a data e a hora da ação, os parâmetros de solicitação e assim por diante. Os arquivos de log do CloudTrail não são um rastreamento de pilha ordenada de chamadas de API pública. Dessa forma, eles não são exibidos em uma ordem específica.

As chamadas para a operação de `CreateRescoreExecutionPlan` criam a seguinte 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"
            }
        }
```

Monitoramento do Amazon Kendra com o Amazon CloudWatch

Para monitorar a integridade de seus índices, use o Amazon CloudWatch. Com o CloudWatch, obtenha métricas para sincronização de documentos para seu índice. Você também pode configurar os alarmes do CloudWatch para ser notificado quando uma ou mais métricas excederem um limite definido por você. Por exemplo, monitore o número de documentos enviados para serem indexados ou o número de documentos que não foram indexados.

Você deve ter as permissões do apropriadas para monitorar Amazon Kendra com o CloudWatch. Para obter mais informações, consulte Autenticação e controle de acesso para o Amazon CloudWatch no Guia do usuário do Amazon CloudWatch.

Visualizar métricas do Amazon Kendra

Visualize métricas do Amazon Kendra usando o console do CloudWatch.

Para visualizar métricas (console do CloudWatch)

1. Faça login no AWS Management Console e abra o console do CloudWatch em https://console.aws.amazon.com/cloudwatch/.
2. Selecione Métricas, escolha Todas as métricas e, em seguida, selecione Kendra.
3. Escolha a dimensão, informe um nome de métrica e selecione Adicionar ao gráfico.
4. Escolha um valor para o intervalo de datas. A contagem da métrica para o intervalo de datas selecionado é exibida no gráfico.

Criar um alarme

Um alarme do CloudWatch monitora uma única métrica durante um período de tempo especificado e executa uma ou mais ações: envio de uma notificação para uma parte superior do Amazon Simple Notification Service (Amazon SNS) ou política de ajuste de escala automático. As ações são baseadas no valor da métrica relativa a um limite determinado durante um número de períodos de tempo que você especificar. O CloudWatch também pode enviar uma mensagem do Amazon SNS quando o alarme muda de estado.

Os alarmes do invocam ações somente quando o estado mudar e tiver persistido pelo período que você especificada.

Para definir um alarme

1. Faça login no AWS Management Console e abra o console do CloudWatch em https://console.aws.amazon.com/cloudwatch/.
2. Escolha Alarmes e depois Criar alarme.
3. Selecionar uma métrica. Escolha uma métrica do Kendra para seu índice e fonte de dados. Defina também a hora como um número definido de horas, dias, semanas ou personalizado.
4. Escolha sua estatística. Por exemplo, Média. Escolha também o período de ativação do alarme como um número definido de minutos, horas por dia ou personalizado.
5. Escolha seu limite para acionar o alarme, seja para usar um valor estático ou uma banda e a condição a ser cumprida para o limite.
6. Escolha o estado do alarme para o gatilho, independentemente de a métrica estar fora do limite definido ou de outro estado. Selecione quem/para qual e-mail enviar a notificação de alarme.
7. Se você estiver satisfeito com o alarme, selecione Criar alarme.

> **Note**
>
> Forneça um nome para o alarme do CloudWatch.

Métricas do CloudWatch para trabalhos de sincronização de índices

A tabela a seguir descreve as métricas do Amazon Kendra para trabalhos de sincronização de fontes de dados.

Se você usa a API ou a CLI, deverá especificar o `Namespace` como “AWS/Kendra” além do `MetricName` da sua escolha ao usar a API GetMetricStatistics.

Métrica	Descrição
`DocumentsCrawled`	O número de documentos que o trabalho de sincronização verificou ou descobriu durante a execução. Dimensões: • IndexId • DataSourceId Unidade: contagem
`DocumentsSubmittedForIndexing`	O número de documentos que o trabalho de sincronização enviou ao índice. Dimensões: • IndexId • DataSourceId Unidade: contagem
`DocumentsSubmittedForIndexingFailed`	O número de documentos que falharam na indexação. Verifique o conteúdo do log do

Métrica	Descrição
	CloudWatch da tarefa de sincronização para obter detalhes. Dimensões: • IndexId • DataSourceId Unidade: contagem
DocumentsSubmittedForDeletion	O número de documentos que o trabalho de sincronização solicitou que fossem removidos do índice. Dimensões: • IndexId • DataSourceId Unidade: contagem
DocumentsSubmittedForDeletionFailed	O número de documentos que não foram excluídos. Verifique o conteúdo do log do CloudWatch da tarefa de sincronização para obter detalhes. Dimensões: • IndexId • DataSourceId Unidade: contagem

Métricas para fontes de dados do Amazon Kendra

A tabela a seguir descreve as métricas do Amazon Kendra para trabalhos de sincronização de fontes de dados. Métricas marcadas com um asterisco (*) são usadas apenas para fontes de dados do Amazon S3.

Se você usa a API ou a CLI, deverá especificar o `Namespace` como “AWS/Kendra” além do `MetricName` da sua escolha ao usar a API GetMetricStatistics.

Métrica	Descrição
`DocumentsSkippedNoChange *`	O número de documentos examinados e considerados inalterados e, portanto, não foram enviados para indexação. Dimensões: • IndexId • DataSourceId Unidade: contagem
`DocumentsSkippedInvalidMetadata *`	O número de documentos ignorados porque havia um problema com o arquivo de metadados associado. Verifique o conteúdo do log do CloudWatch para a execução da sincronização para obter detalhes. Dimensões: • IndexId • DataSourceId Unidade: contagem
`DocumentsCrawled`	O número de arquivos de documentos examinados.

Métrica	Descrição
	Dimensões: • IndexId • DataSourceId Unidade: contagem
DocumentsSubmittedForDeletion	O número de documentos examinados que foram excluídos da fonte de dados e enviados para exclusão. Dimensões: • IndexId • DataSourceId Unidade: contagem
DocumentsSubmittedForDeletionFailed	O número de documentos que falharam na exclusão de uma fonte de dados. Dimensões: • IndexId • DataSourceId Unidade: contagem

Métrica	Descrição
`DocumentsSubmittedForIndexing`	O número de documentos examinados e enviados para indexação. Dimensões: • IndexId • DataSourceId Unidade: contagem
`DocumentsSubmittedForIndexingFailed`	O número de documentos enviados para indexação que não puderam ser indexados. Dimensões: • IndexId • DataSourceId Unidade: contagem

Métricas para documentos indexados

A tabela a seguir descreve as métricas do Amazon Kendra para documentos indexados. Para documentos indexados usando a operação BatchputDocument, tem suporte apenas a dimensão `IndexId`.

Se você usa a API ou a CLI, deverá especificar o `Namespace` como “AWS/Kendra” além do `MetricName` da sua escolha ao usar a API GetMetricStatistics.

Métrica	Descrição
`DocumentsIndexed`	O número de documentos indexados. Dimensões: • IndexId

Métrica	Descrição
	• DataSourceId Unidade: contagem
DocumentsFailedToIndex	O número de documentos que não puderam ser indexados. Verifique o conteúdo do log do CloudWatch para obter detalhes. Dimensões: • IndexId • DataSourceId Unidade: contagem
IndexQueryCount	O número de consultas de índice por minuto. Dimensões: • IndexId Unidade: contagem

Monitoramento do Amazon Kendra com o Amazon CloudWatch Logs

O Amazon Kendra usa o Amazon CloudWatch Logs para fornecer uma visão sobre a operação de suas fontes de dados. O Amazon Kendra registra os detalhes do processo dos documentos à medida que são indexados. Ele registra erros da sua fonte de dados que ocorrem enquanto seus documentos estão sendo indexados. Você usa o CloudWatch Logs para monitorar, armazenar e acessar os arquivos de log.

O CloudWatch Logs armazena eventos de log em um fluxo de logs que faz parte de um grupo de logs. O Amazon Kendra usa esses recursos da seguinte forma:

- Grupo de logs – O Amazon Kendra armazena todos os fluxo de logs em um único grupo de logs para cada índice. O Amazon Kendra cria o grupo de logs quando o índice é criado. O identificador do grupo de logs sempre começa com “aws/kendra/”.
- Fluxo de logs – O Amazon Kendra cria um novo fluxo de logs da fonte de dados no grupo de logs para cada trabalho de sincronização de índice que você executa. Ele também cria um novo fluxo de logs de documentos quando um fluxo atinge aproximadamente 500 entradas.
- Entradas de log – O Amazon Kendra cria uma entrada de log no fluxo de logs à medida que indexa documentos. Cada entrada fornece informações sobre o processamento do documento ou sobre quaisquer erros encontrados.

Para obter mais informações sobre como usar o CloudWatch Logs, consulte O que é o Amazon Cloud Watch Logs no Guia do usuário do Amazon Cloud Watch Logs.

O Amazon Kendra cria dois tipos de fluxos de log:

- Fluxos de log da fonte de dados
- Fluxos de log de documentos

Fluxos de log da fonte de dados

Os fluxos de log da fonte de dados publicam entradas sobre seus trabalhos de sincronização de índices. Cada tarefa de sincronização cria um novo fluxo de logs que é usado para publicar entradas. O nome do fluxo de logs é:

```
data source id/YYYY-MM-DD-HH/data source sync job ID
```

Um novo fluxo de logs é criado para cada execução de trabalho de sincronização.

Há três tipos de mensagens de log publicadas em um fluxo de logs da fonte de dados:

- Uma mensagem de log para um documento que não foi enviado para indexação. Veja a seguir um exemplo dessa mensagem para um documento em uma fonte de dados do 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."
}
```

- Uma mensagem de log para um documento que não foi enviado para exclusão. A seguir está um exemplo desta mensagem:

```
{
    "DocumentId": "document ID",
    "Message": "Failed to delete document via BatchDeleteDocument.",
    "ErrorCode": "InvalidRequest",
    "ErrorMessage": "Document can't be deleted because it doesn't exist."
}
```

- Uma mensagem de log quando um arquivo de metadados inválido para um documento em um bucket do Amazon S3 é encontrado. A seguir está um exemplo desta mensagem.

```
{
    "Message": "Found invalid metadata
 file bucket/prefix/filename.extension.metadata.json."
}
```

- Para o SharePoint e os conectores de banco de dados, o Amazon Kendra só grava mensagens no fluxo de logs se um documento não puder ser indexado. Veja a seguir um exemplo da mensagem de erro registrada pelo Amazon Kendra.

```
{
    "DocumentID": "document ID",
    "IndexID": "index ID",
    "SourceURI": "",
    "CrawlStatus": "FAILED",
    "ErrorCode": "403",
    "ErrorMessage": "Access Denied",
    "DataSourceErrorCode": "403"
}
```

Fluxos de log de documentos

O Amazon Kendra registra informações sobre o processamento de documentos enquanto eles estão sendo indexados. Ele registra um conjunto de mensagens para documentos armazenados em uma

fonte de dados do Amazon S3. Ele registra erros somente para documentos armazenados em um Microsoft SharePoint ou em uma fonte de dados de banco de dados.

Se os documentos foram adicionados ao índice usando a operação BatchputDocument, o fluxo de logs será nomeado da seguinte forma:

```
YYYY-MM-DD-HH/UUID
```

Se os documentos foram adicionados ao índice usando uma fonte de dados, o fluxo de logs será nomeado da seguinte forma:

```
dataSourceId/YYYY-MM-DD-HH/UUID
```

Cada fluxo de logs contém até 500 mensagens.

Se a indexação de um documento falhar, essa mensagem será enviada para o fluxo de logs:

```
{
    "DocumentId": "document ID",
    "IndexName": "index name",
    "IndexId": "index ID"
    "SourceURI": "source URI"
    "IndexingStatus": "DocumentFailedToIndex",
    "ErrorCode": "400 | 500",
    "ErrorMessage": "message"
}
```

Segurança no Amazon Kendra

A segurança na nuvem AWS é a maior prioridade. Como AWS cliente, você se beneficia de uma arquitetura de data center e rede criada para atender aos requisitos das organizações mais sensíveis à segurança.

A segurança é uma responsabilidade compartilhada entre você AWS e você. O modelo de responsabilidade compartilhada descreve isso como segurança da nuvem e segurança na nuvem:

- Segurança da nuvem —AWS é responsável por proteger a infraestrutura que executa AWS os serviços na AWS nuvem. AWS também fornece serviços que você pode usar com segurança. Auditores terceirizados testam e verificam regularmente a eficácia de nossa segurança como parte dos Programas de Conformidade Programas de AWS de . Para saber mais sobre os programas de conformidade que se aplicam ao Amazon Kendra, AWS consulte Serviços no escopo do programa de conformidade Serviços no escopo AWS de conformidade.
- Segurança na nuvem — Sua responsabilidade é determinada pelo AWS serviço que você usa. Você também é responsável por outros fatores, incluindo a confidencialidade de seus dados, os requisitos da empresa e as leis e regulamentos aplicáveis.

Esta documentação ajuda você a entender como aplicar o modelo de responsabilidade compartilhada ao usar o Amazon Kendra. Os tópicos a seguir mostram como configurar o Amazon Kendra para atender aos seus objetivos de segurança e compatibilidade. Você também aprende a usar outros AWS serviços que ajudam você a monitorar e proteger seus recursos da Amazon Kendra.

Tópicos

- Proteção de dados no Amazon Kendra
- Amazon Kendra Amazon Kendra Classificação inteligente e interface VPC endpoints ()AWS PrivateLink
- Gerenciamento de identidade e acesso para o Amazon Kendra
- Melhores práticas de segurança
- Registrar em log e monitorar no Amazon Kendra
- Validação de conformidade para o Amazon Kendra
- Resiliência no Amazon Kendra
- Segurança da infraestrutura no Amazon Kendra

- Análise de configuração e vulnerabilidade em AWS Identity and Access Management

Proteção de dados no Amazon Kendra

O modelo de responsabilidade AWS compartilhada se aplica à proteção de dados no Amazon Kendra. Conforme descrito neste modelo, AWS é responsável por proteger a infraestrutura global que executa todos os Nuvem AWS. Você é responsável por manter o controle sobre seu conteúdo hospedado nessa infraestrutura. Você também é responsável pelas tarefas de configuração e gerenciamento de segurança dos Serviços da AWS que usa. Para obter mais informações sobre a privacidade de dados, consulte as Perguntas Frequentes sobre Privacidade de Dados.. Para obter mais informações sobre a proteção de dados na Europa, consulte a postagem do blog AWS LGPD e Modelo de Responsabilidade Compartilhada no AWS Blog de Segurança.

Para fins de proteção de dados, recomendamos que você proteja Conta da AWS as credenciais e configure usuários individuais com AWS IAM Identity Center ou AWS Identity and Access Management (IAM). Dessa maneira, cada usuário receberá apenas as permissões necessárias para cumprir suas obrigações de trabalho. Recomendamos também que você proteja seus dados das seguintes formas:

- Use uma autenticação multifator (MFA) com cada conta.
- Use SSL/TLS para se comunicar com os recursos. AWS Exigimos TLS 1.2 e recomendamos TLS 1.3.
- Configure a API e o registro de atividades do usuário com AWS CloudTrail.
- Use soluções de AWS criptografia, juntamente com todos os controles de segurança padrão Serviços da AWS.
- Use serviços gerenciados de segurança avançada, como o Amazon Macie, que ajuda a localizar e proteger dados sigilosos armazenados no Amazon S3.
- Se você precisar de módulos criptográficos validados pelo FIPS 140-2 ao acessar AWS por meio de uma interface de linha de comando ou de uma API, use um endpoint FIPS. Para ter mais informações sobre endpoints do FIPS disponíveis, consulte Federal Information Processing Standard (FIPS) 140-2.

É altamente recomendável que nunca sejam colocadas informações de identificação confidenciais, como endereços de e-mail dos seus clientes, em marcações ou campos de formato livre, como um campo Nome. Isso inclui quando você trabalha com o Amazon Kendra ou Serviços da AWS outro usando o console, a API AWS CLI ou os SDKs. AWS Quaisquer dados inseridos em tags ou campos

de texto de formato livre usados para nomes podem ser usados para logs de faturamento ou de diagnóstico. Se você fornecer um URL para um servidor externo, recomendemos fortemente que não sejam incluídas informações de credenciais no URL para validar a solicitação a esse servidor.

Criptografia em repouso

O Amazon Kendra criptografa os dados em repouso com a chave de criptografia escolhida. Você pode escolher uma das seguintes opções:

- Uma chave AWS KMS AWS de propriedade própria. Se você não especificar uma chave de criptografia, os dados serão criptografados com essa chave por padrão.
- Uma chave KMS AWS gerenciada em sua conta. Essa chave é criada, gerenciada e usada em seu nome pelo Amazon Kendra. O nome da chave é `aws/kendra`.
- Uma chave gerenciada pelo cliente Você pode informar o ARN de uma chave de criptografia criada em sua conta. Ao usar uma chave KMS gerenciada pelo cliente, você deve fornecer à chave uma política de chave que permita que a Amazon Kendra use a chave. Selecione uma chave KMS de criptografia simétrica gerenciada pelo cliente. O Amazon Kendra não oferece suporte a chaves KMS assimétricas. Para ter mais informações, consulte Gerenciamento de chaves.

Criptografia em trânsito

O Amazon Kendra usa o protocolo HTTPS para se comunicar com o aplicativo do cliente. Ele usa HTTPS e AWS assinaturas para se comunicar com outros serviços em nome do seu aplicativo. Se você usa uma VPC, você pode usá-la AWS PrivateLink para estabelecer uma conexão privada entre sua VPC e a Amazon Kendra.

Gerenciamento de chaves

O Amazon Kendra criptografa o conteúdo do índice usando um dos três tipos de chaves. Você pode escolher uma das seguintes opções:

- Um AWS AWS KMS de propriedade própria. Esse é o padrão.
- Uma chave KMS AWS gerenciada. Essa chave é criada em sua conta e é gerenciada e usada em seu nome pelo Amazon Kendra.
- Uma chave do KMS gerenciada pelo cliente Crie a chave ao criar um índice ou fonte de dados do Amazon Kendra, ou crie a chave usando o console da AWS KMS . Selecione uma chave KMS de criptografia simétrica gerenciada pelo cliente. O Amazon Kendra não oferece suporte a chaves do

KMS assimétricas. Para obter mais informações, consulte Usar chaves simétricas e assimétricas no Guia do Desenvolvedor do Key Management Service da AWS .

Amazon Kendra Amazon Kendra Classificação inteligente e interface VPC endpoints ()AWS PrivateLink

É possível estabelecer uma conexão privada entre a VPC e o Amazon Kendra criando uma endpoint da VPC de interface. Os endpoints de interface são alimentados por AWS PrivateLinkuma tecnologia que permite acessar de forma privada as APIs do Amazon Kendra sem um gateway de internet, dispositivo NAT, conexão VPN ou conexão Direct Connect. AWS As instâncias na VPC não precisam de endereços IP públicos para a comunicação com APIs do Amazon Kendra. O tráfego entre sua VPC e o Amazon Kendra não sai da rede Amazon.

Cada endpoint de interface é representado por uma ou mais Interfaces de Rede Elástica nas sub-redes.

Considerações sobre os endpoints VPC Amazon Kendra e Amazon Kendra Intelligent Ranking

Antes de configurar uma interface VPC endpoint para Amazon Kendra ou Amazon Kendra Intelligent Ranking, certifique-se de revisar os pré-requisitos no Guia do usuário do Amazon VPC.

O Amazon Kendra e o Amazon Kendra Intelligent Ranking oferecem suporte para fazer chamadas para todas as suas ações de API a partir de sua VPC.

Criação de uma interface VPC endpoint para Amazon Kendra e Amazon Kendra Intelligent Ranking

Você pode criar um VPC endpoint para o serviço Amazon Kendra ou Amazon Kendra Intelligent Ranking usando o console Amazon VPC ou o (). AWS Command Line Interface AWS CLI

Crie uma endpoint da VPC para o Amazon Kendra usando o seguinte nome de serviço:

- com.amazonaws.*region*.kendra

Crie um VPC endpoint para o Amazon Kendra Intelligent Ranking usando o seguinte nome de serviço:

- aws.api.*region*.kendra-ranking

Depois de criar um VPC endpoint, você pode usar o seguinte exemplo de AWS CLI comando que usa o `endpoint-url` parâmetro para especificar um endpoint de interface para a API Amazon Kendra:

```
aws kendra list-indices --endpoint-url https://VPC endpoint
```

0 VPC endpoint é o nome DNS gerado quando o endpoint da interface é criado. Esse nome inclui o ID do VPC endpoint e o nome do serviço Amazon Kendra, que inclui a região. Por exemplo, `vpce-1234-abcdef.kendra.us-west-2.vpce.amazonaws.com`.

Se você ativar o DNS privado para o endpoint, poderá fazer solicitações de API para a Amazon Kendra usando seu nome DNS padrão para a região. Por exemplo, `kendra.us-east-1.amazonaws.com`.

Para mais informações, consulte Criar um endpoint de interface no Guia do usuário da Amazon VPC.

Criação de uma política de VPC endpoint para Amazon Kendra e Amazon Kendra Intelligent Ranking

Você pode anexar uma política de endpoint ao seu VPC endpoint que controla o acesso ao Amazon Kendra ou ao Amazon Kendra Intelligent Ranking.

A política do Amazon Kendra ou Amazon Kendra Intelligent Ranking especifica as seguintes informações:

- O usuário principal/autorizado que pode realizar ações.
- As ações que podem ser executadas.
- Os recursos sobre os quais as ações podem ser realizadas.

Exemplo: política de endpoint da VPC para ações do Amazon Kendra

Veja a seguir um exemplo de política de endpoint para o Amazon Kendra. Quando anexada a um endpoint, essa política concede acesso a todas as ações disponíveis do Amazon Kendra para todos os usuários/diretores autorizados em todos os recursos.

```
{
```

```
   "Statement":[
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "kendra:*"
         ],
         "Resource":"*"
      }
   ]
}
```

Exemplo: política de VPC endpoint para ações do Amazon Kendra Intelligent Ranking

Veja a seguir um exemplo de uma política de endpoint para o Amazon Kendra Intelligent Ranking. Quando anexada a um endpoint, essa política concede acesso a todas as ações disponíveis do Amazon Kendra Intelligent Ranking para todos os usuários principais/autorizados em todos os recursos.

```
{
   "Statement":[
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "kendra-ranking:*"
         ],
         "Resource":"*"
      }
   ]
}
```

Para obter mais informações, consulte Controle do acesso aos endpoints da VPC usando políticas de endpoint no Guia do usuário da Amazon VPC.

Gerenciamento de identidade e acesso para o Amazon Kendra

AWS Identity and Access Management (IAM) é uma ferramenta AWS service (Serviço da AWS) que ajuda o administrador a controlar com segurança o acesso aos AWS recursos. Os administradores do IAM controlam quem pode ser autenticado (conectado) e autorizado (ter permissões) para utilizar

os recursos do Amazon Kendra. O IAM é um AWS service (Serviço da AWS) que você pode usar sem custo adicional.

Tópicos

- Público
- Autenticando com identidades
- Gerenciando acesso usando políticas
- Como o Amazon Kendra funciona com o IAM
- Exemplos de políticas baseadas em identidade do Amazon Kendra
- AWS políticas gerenciadas para o Amazon Kendra
- Solução de problemas de identidade e acesso do Amazon Kendra

Público

A forma como você usa AWS Identity and Access Management (IAM) difere, dependendo do trabalho que você faz na Amazon Kendra.

Usuário do serviço: se você usar o serviço do Amazon Kendra para fazer seu trabalho, o administrador fornecerá as credenciais e as permissões necessárias. À medida que mais recursos do Amazon Kendra forem usados para realizar o trabalho, talvez sejam necessárias permissões adicionais. Entender como o acesso é gerenciado pode ajudar você a solicitar as permissões corretas ao seu administrador. Se você não conseguir acessar um atributo no Amazon Kendra, consulte Solução de problemas de identidade e acesso do Amazon Kendra.

Administrador do serviço: se você for o responsável pelos recursos do Amazon Kendra em sua empresa, você provavelmente terá acesso total ao Amazon Kendra. Cabe a você determinar quais funcionalidades e recursos do Amazon Kendra os usuários do seu serviço devem acessar. Assim, você deve enviar solicitações ao administrador do IAM para alterar as permissões dos usuários de seu serviço. Revise as informações nesta página para entender os Introdução ao IAM. Para saber mais sobre como sua empresa pode usar o IAM com o Amazon Kendra, consulte Como o Amazon Kendra funciona com o IAM.

Administrador do IAM: se você for um administrador do IAM, talvez deseje saber detalhes sobre como escrever políticas para gerenciar o acesso ao Amazon Kendra. Para visualizar exemplos de políticas baseadas em identidade do Amazon Kendra que podem ser usadas no IAM, consulte Exemplos de políticas baseadas em identidade do Amazon Kendra.

Autenticando com identidades

A autenticação é a forma como você faz login AWS usando suas credenciais de identidade. Você deve estar autenticado (conectado AWS) como o Usuário raiz da conta da AWS, como usuário do IAM ou assumindo uma função do IAM.

Você pode entrar AWS como uma identidade federada usando credenciais fornecidas por meio de uma fonte de identidade. AWS IAM Identity Center Usuários (IAM Identity Center), a autenticação de login único da sua empresa e suas credenciais do Google ou do Facebook são exemplos de identidades federadas. Quando você faz login como identidade federada, o administrador já configurou anteriormente a federação de identidades usando perfis do IAM. Ao acessar AWS usando a federação, você está assumindo indiretamente uma função.

Dependendo do tipo de usuário que você é, você pode entrar no AWS Management Console ou no portal de AWS acesso. Para obter mais informações sobre como fazer login AWS, consulte Como fazer login Conta da AWS no Guia do Início de Sessão da AWS usuário.

Se você acessar AWS programaticamente, AWS fornece um kit de desenvolvimento de software (SDK) e uma interface de linha de comando (CLI) para assinar criptograficamente suas solicitações usando suas credenciais. Se você não usa AWS ferramentas, você mesmo deve assinar as solicitações. Para obter mais informações sobre como usar o método recomendado para assinar solicitações por conta própria, consulte Assinatura de solicitações de AWS API no Guia do usuário do IAM.

Independente do método de autenticação usado, também pode ser exigido que você forneça informações adicionais de segurança. Por exemplo, AWS recomenda que você use a autenticação multifator (MFA) para aumentar a segurança da sua conta. Para saber mais, consulte Autenticação multifator no Guia AWS IAM Identity Center do usuário e Utilizar a autenticação multifator (MFA) na AWS no Guia do usuário do IAM.

Conta da AWS usuário root

Ao criar uma Conta da AWS, você começa com uma identidade de login que tem acesso completo a todos Serviços da AWS os recursos da conta. Essa identidade é chamada de usuário Conta da AWS raiz e é acessada fazendo login com o endereço de e-mail e a senha que você usou para criar a conta. É altamente recomendável não usar o usuário raiz para tarefas diárias. Proteja as credenciais do usuário raiz e use-as para executar as tarefas que somente ele puder executar. Para obter a lista completa das tarefas que exigem login como usuário raiz, consulte Tarefas que exigem credenciais de usuário raiz no Guia do Usuário do IAM.

Grupos e usuários do IAM

Um usuário do IAM é uma identidade dentro da sua Conta da AWS que tem permissões específicas para uma única pessoa ou aplicativo. Sempre que possível, recomendamos contar com credenciais temporárias em vez de criar usuários do IAM com credenciais de longo prazo, como senhas e chaves de acesso. No entanto, se você tiver casos de uso específicos que exijam credenciais de longo prazo com usuários do IAM, recomendamos alternar as chaves de acesso. Para obter mais informações, consulte Alterne as chaves de acesso regularmente para casos de uso que exijam credenciais de longo prazo no Guia do Usuário do IAM.

Um grupo do IAM é uma identidade que especifica uma coleção de usuários do IAM. Não é possível fazer login como um grupo. É possível usar grupos para especificar permissões para vários usuários de uma vez. Os grupos facilitam o gerenciamento de permissões para grandes conjuntos de usuários. Por exemplo, você pode ter um grupo chamado IAMAdmins e conceder a esse grupo permissões para administrar recursos do IAM.

Usuários são diferentes de perfis. Um usuário é exclusivamente associado a uma pessoa ou a uma aplicação, mas um perfil pode ser assumido por qualquer pessoa que precisar dele. Os usuários têm credenciais permanentes de longo prazo, mas os perfis fornecem credenciais temporárias. Para saber mais, consulte Quando criar um usuário do IAM (em vez de um perfil) no Guia do usuário do IAM.

Perfis do IAM

Uma função do IAM é uma identidade dentro da sua Conta da AWS que tem permissões específicas. Ele é semelhante a um usuário do IAM, mas não está associado a uma pessoa específica. Você pode assumir temporariamente uma função do IAM no AWS Management Console trocando de funções. Você pode assumir uma função chamando uma operação de AWS API AWS CLI ou usando uma URL personalizada. Para obter mais informações sobre métodos para o uso de perfis, consulte Utilizar perfis do IAM no Guia do usuário do IAM.

Funções do IAM com credenciais temporárias são úteis nas seguintes situações:

- Acesso de usuário federado: para atribuir permissões a identidades federadas, você pode criar um perfil e definir permissões para ele. Quando uma identidade federada é autenticada, essa identidade é associada ao perfil e recebe as permissões definidas pelo mesmo. Para obter mais informações sobre perfis para federação, consulte Criar um perfil para um provedor de identidades de terceiros no Guia do Usuário do IAM. Se você usar o Centro de identidade do IAM, configure um conjunto de permissões. Para controlar o que suas identidades podem acessar após a

autenticação, o Centro de identidade do IAM correlaciona o conjunto de permissões a um perfil no IAM. Para obter informações sobre conjuntos de permissões, consulte Conjuntos de permissões no Guia do usuário AWS IAM Identity Center .

- Permissões temporárias para usuários do IAM –- um usuário ou um perfil do IAM pode presumir um perfil do IAM para obter temporariamente permissões diferentes para uma tarefa específica.
- Acesso entre contas –- é possível usar um perfil do IAM para permitir que alguém (uma entidade principal confiável) em outra conta acesse recursos em sua conta. Os perfis são a principal forma de conceder acesso entre contas. No entanto, com alguns Serviços da AWS, você pode anexar uma política diretamente a um recurso (em vez de usar uma função como proxy). Para saber a diferença entre funções e políticas baseadas em recursos para acesso entre contas, consulte Acesso a recursos entre contas no IAM no Guia do usuário do IAM.
- Acesso entre serviços — Alguns Serviços da AWS usam recursos em outros Serviços da AWS. Por exemplo, quando você faz uma chamada em um serviço, é comum que esse serviço execute aplicativos no Amazon EC2 ou armazene objetos no Amazon S3. Um serviço pode fazer isso usando as permissões do principal de chamada, usando um perfil de serviço ou um perfil vinculado a serviço.
  - Sessões de acesso direto (FAS) — Quando você usa um usuário ou uma função do IAM para realizar ações AWS, você é considerado principal. Ao usar alguns serviços, você pode executar uma ação que inicia outra ação em um serviço diferente. O FAS usa as permissões do diretor chamando um AWS service (Serviço da AWS), combinadas com a solicitação AWS service (Serviço da AWS) para fazer solicitações aos serviços posteriores. As solicitações do FAS são feitas somente quando um serviço recebe uma solicitação que requer interações com outros Serviços da AWS ou com recursos para ser concluída. Nesse caso, você precisa ter permissões para executar ambas as ações. Para obter detalhes da política ao fazer solicitações de FAS, consulte Encaminhar sessões de acesso.
  - Função de serviço: um perfil de serviço é um perfil do IAM que um serviço assume para realizar ações em seu nome. Um administrador do IAM pode criar, modificar e excluir um perfil de serviço do IAM. Para obter mais informações, consulte Criar um perfil para delegar permissões a um AWS service (Serviço da AWS) no Guia do Usuário do IAM.
  - Função vinculada ao serviço — Uma função vinculada ao serviço é um tipo de função de serviço vinculada a um. AWS service (Serviço da AWS) O serviço pode presumir a função de executar uma ação em seu nome. As funções vinculadas ao serviço aparecem em você Conta da AWS e são de propriedade do serviço. Um administrador do IAM pode visualizar, mas não editar as permissões para funções vinculadas ao serviço.

- Aplicativos em execução no Amazon EC2 — Você pode usar uma função do IAM para gerenciar credenciais temporárias para aplicativos que estão sendo executados em uma instância do EC2 e fazendo AWS CLI solicitações de API. AWS É preferível fazer isso a armazenar chaves de acesso na instância do EC2. Para atribuir uma AWS função a uma instância do EC2 e disponibilizá-la para todos os seus aplicativos, você cria um perfil de instância anexado à instância. Um perfil de instância contém o perfil e permite que os programas em execução na instância do EC2 obtenham credenciais temporárias. Para mais informações, consulte Utilizar um perfil do IAM para conceder permissões a aplicações em execução nas instâncias do Amazon EC2 no Guia do usuário do IAM.

Para saber se deseja usar perfis do IAM, consulte Quando criar um perfil do IAM (em vez de um usuário) no Guia do usuário do IAM.

Gerenciando acesso usando políticas

Você controla o acesso AWS criando políticas e anexando-as a AWS identidades ou recursos. Uma política é um objeto AWS que, quando associada a uma identidade ou recurso, define suas permissões. AWS avalia essas políticas quando um principal (usuário, usuário raiz ou sessão de função) faz uma solicitação. As permissões nas políticas determinam se a solicitação será permitida ou negada. A maioria das políticas é armazenada AWS como documentos JSON. Para obter mais informações sobre a estrutura e o conteúdo de documentos de políticas JSON, consulte Visão geral das políticas JSON no Guia do Usuário do IAM.

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos e em que condições.

Por padrão, usuários e funções não têm permissões. Para conceder aos usuários permissões para executar ações nos recursos que eles precisam, um administrador do IAM pode criar políticas do IAM. O administrador pode então adicionar as políticas do IAM aos perfis e os usuários podem presumir os perfis.

As políticas do IAM definem permissões para uma ação independente do método usado para executar a operação. Por exemplo, suponha que você tenha uma política que permite a ação `iam:GetRole`. Um usuário com essa política pode obter informações de função da AWS Management Console AWS CLI, da ou da AWS API.

Políticas baseadas em identidade

As políticas baseadas em identidade são documentos de políticas de permissões JSON que você pode anexar a uma identidade, como usuário do IAM, grupo de usuários ou perfil. Essas políticas

controlam quais ações os usuários e perfis podem realizar, em quais recursos e em que condições. Para saber como criar uma política baseada em identidade, consulte Criando políticas do IAM no Guia do Usuário do IAM.

As políticas baseadas em identidade podem ser categorizadas ainda adicionalmente como políticas em linha ou políticas gerenciadas. As políticas em linha são anexadas diretamente a um único usuário, grupo ou perfil. As políticas gerenciadas são políticas autônomas que você pode associar a vários usuários, grupos e funções em seu Conta da AWS. As políticas AWS gerenciadas incluem políticas gerenciadas e políticas gerenciadas pelo cliente. Para saber como escolher entre uma política gerenciada ou uma política em linha, consulte Escolher entre políticas gerenciadas e políticas em linha no Guia do Usuário do IAM.

Políticas baseadas em recursos

Políticas baseadas em recursos são documentos de políticas JSON que você anexa a um recurso. São exemplos de políticas baseadas em recursos as políticas de confiança de perfil do IAM e as políticas de bucket do Amazon S3. Em serviços que suportem políticas baseadas em recursos, os administradores de serviço podem usá-las para controlar o acesso a um recurso específico. Para o recurso ao qual a política está anexada, a política define quais ações um principal especificado pode executar nesse recurso e em que condições. Você deve especificar uma entidade principal em uma política baseada em recursos. Os diretores podem incluir contas, usuários, funções, usuários federados ou. Serviços da AWS

Políticas baseadas em recursos são políticas em linha localizadas nesse serviço. Você não pode usar políticas AWS gerenciadas do IAM em uma política baseada em recursos.

Listas de controle de acesso (ACLs)

As listas de controle de acesso (ACLs) controlam quais entidades principais (membros, usuários ou perfis da conta) têm permissões para acessar um recurso. As ACLs são semelhantes as políticas baseadas em recursos, embora não usem o formato de documento de política JSON.

O Amazon S3 e o Amazon VPC são exemplos de serviços que oferecem suporte a ACLs. AWS WAF Para saber mais sobre ACLs, consulte Visão geral da lista de controle de acesso (ACL) no Guia do Desenvolvedor do Amazon Simple Storage Service.

Outros tipos de política

AWS oferece suporte a tipos de políticas adicionais menos comuns. Esses tipos de política podem definir o máximo de permissões concedidas a você pelos tipos de política mais comuns.

- Limites de permissões: um limite de permissões é um atributo avançado no qual você define o máximo de permissões que uma política baseada em identidade pode conceder a uma entidade do IAM (usuário ou perfil do IAM). É possível definir um limite de permissões para uma entidade. As permissões resultantes são a interseção das políticas baseadas em identidade de uma entidade com seus limites de permissões. As políticas baseadas em recurso que especificam o usuário ou o perfil no campo `Principal` não são limitadas pelo limite de permissões. Uma negação explícita em qualquer uma dessas políticas substitui a permissão. Para obter mais informações sobre limites de permissões, consulte Limites de permissões para identidades do IAM no Guia do Usuário do IAM.
- Políticas de controle de serviço (SCPs) — SCPs são políticas JSON que especificam as permissões máximas para uma organização ou unidade organizacional (OU) em. AWS Organizations AWS Organizations é um serviço para agrupar e gerenciar centralmente vários Contas da AWS que sua empresa possui. Se você habilitar todos os atributos em uma organização, poderá aplicar políticas de controle de serviço (SCPs) a qualquer uma ou a todas as contas. O SCP limita as permissões para entidades nas contas dos membros, incluindo cada uma Usuário raiz da conta da AWS. Para obter mais informações sobre o Organizações e SCPs, consulte How SCPs work (Como os SCPs funcionam) no Guia do usuário do AWS Organizations .
- Políticas de sessão: são políticas avançadas que você transmite como um parâmetro quando cria de forma programática uma sessão temporária para um perfil ou um usuário federado. As permissões da sessão resultante são a interseção das políticas baseadas em identidade do usuário ou do perfil e das políticas de sessão. As permissões também podem ser provenientes de uma política baseada em atributo. Uma negação explícita em qualquer uma dessas políticas substitui a permissão. Para obter mais informações, consulte Políticas de sessão no Guia do Usuário do IAM.

Vários tipos de política

Quando vários tipos de política são aplicáveis a uma solicitação, é mais complicado compreender as permissões resultantes. Para saber como AWS determinar se uma solicitação deve ser permitida quando vários tipos de políticas estão envolvidos, consulte Lógica de avaliação de políticas no Guia do usuário do IAM.

Como o Amazon Kendra funciona com o IAM

Antes de usar o IAM para gerenciar o acesso ao Amazon Kendra, você deve entender quais recursos do IAM estão disponíveis para uso com o Amazon Kendra. Para ter uma visão de alto nível de

como o Amazon Kendra AWS e outros serviços funcionam com o IAM AWS , consulte Serviços que funcionam com o IAM no Guia do usuário do IAM.

Tópicos

- Políticas baseadas em identidade do Amazon Kendra
- Políticas baseadas em recursos do Amazon Kendra
- Listas de controle de acesso (ACLs)
- Autorização baseada em tags do Amazon Kendra
- Funções do IAM no Amazon Kendra

Políticas baseadas em identidade do Amazon Kendra

Com as políticas baseadas em identidade do IAM, é possível especificar ações ou recursos permitidos ou negados, assim como as condições sob as quais as ações são permitidas ou negadas. O Amazon Kendra oferece suporte a ações, chaves de condição e recursos específicos. Para conhecer todos os elementos usados em uma política JSON, consulte Referência de elementos de política JSON do IAM no Guia do usuário do IAM.

Ações

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos, e em que condições.

O elemento `Action` de uma política JSON descreve as ações que você pode usar para permitir ou negar acesso em uma política. As ações de política geralmente têm o mesmo nome da operação de AWS API associada. Existem algumas exceções, como ações somente de permissão, que não têm uma operação de API correspondente. Algumas operações também exigem várias ações em uma política. Essas ações adicionais são chamadas de ações dependentes.

Incluem ações em uma política para conceder permissões para executar a operação associada.

As ações de política no Amazon Kendra usam o seguinte prefixo antes da ação: `kendra:`. Por exemplo, para conceder permissão a alguém para listar índices do Amazon Kendra com a operação ListIndicesda API, você inclui a ação `kendra:ListIndices` na política dessa pessoa. As instruções de política devem incluir um elemento `Action` ou `NotAction`. O Amazon Kendra define seu próprio conjunto de ações que descrevem as tarefas que você pode executar com esse serviço.

Para especificar várias ações em uma única instrução, separe-as com vírgulas, como segue:

```
"Action": [
      "kendra:action1",
      "kendra:action2"
```

Você também pode especificar várias ações usando caracteres curinga (*). Por exemplo, para especificar todas as ações que começam com a palavra `Describe`, inclua a seguinte ação:

```
"Action": "kendra:Describe*"
```

Para ver uma lista das ações do Amazon EMR, consulte Ações definidas pelo Amazon Kendra no Guia do usuário do IAM.

Recursos

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos, e em que condições.

O elemento de política JSON `Resource` especifica o objeto ou os objetos aos quais a ação se aplica. As instruções devem incluir um elemento `Resource` ou `NotResource`. Como prática recomendada, especifique um recurso usando seu nome do recurso da Amazon (ARN). Isso pode ser feito para ações que oferecem compatibilidade com um tipo de recurso específico, conhecido como permissões em nível de recurso.

Para ações que não oferecem compatibilidade com permissões em nível de recurso, como operações de listagem, use um curinga (*) para indicar que a instrução se aplica a todos os recursos.

```
"Resource": "*"
```

O recurso do índice do Amazon Kendra tem o seguinte ARN:

```
arn:${Partition}:kendra:${Region}:${Account}:index/${IndexId}
```

Para obter mais informações sobre o formato dos ARNs, consulte Amazon Resource Names (ARNs) e AWS Service Namespaces.

Por exemplo, para especificar o índice em sua instrução, use o GUID do índice no seguinte ARN:

```
"Resource": "arn:aws:kendra:${Region}:${Account}:index/${GUID}"
```

Para especificar todas os índices que pertencem a uma conta específica, use o caractere curinga (*):

```
"Resource": "arn:aws:${Region}:${Account}:index/*"
```

Algumas ações do Amazon Kendra, como as de criação de recursos, não podem ser executadas em um recurso específico. Nesses casos, você deve utilizar o caractere curinga (*).

```
"Resource": "*"
```

Para ver uma lista dos tipos de recursos do Amazon Kendra e seus ARNs, consulte Recursos definidos pelo Amazon Kendra no Manual do usuário do IAM. Para saber com quais ações você pode especificar o ARN de cada recurso, consulte Ações definidas pelo Amazon Kendra.

Chaves de condição

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos, e em que condições.

O elemento `Condition` (ou bloco `Condition`) permite que você especifique condições nas quais uma instrução estiver em vigor. O elemento `Condition` é opcional. É possível criar expressões condicionais que usem agentes de condição, como “igual a” ou “menor que”, para fazer a condição da política corresponder aos valores na solicitação.

Se você especificar vários elementos `Condition` em uma instrução ou várias chaves em um único `Condition` elemento, a AWS os avaliará usando uma operação lógica `AND`. Se você especificar vários valores para uma única chave de condição, AWS avalia a condição usando uma `OR` operação lógica. Todas as condições devem ser atendidas antes que as permissões da instrução sejam concedidas.

Você também pode usar variáveis de espaço reservado ao especificar condições. Por exemplo, é possível conceder a um usuário do IAM permissão para acessar um atributo somente se ele estiver marcado com seu nome de usuário do IAM. Para obter mais informações, consulte Elementos da política do IAM: variáveis e tags no Guia do usuário do IAM.

AWS suporta chaves de condição globais e chaves de condição específicas do serviço. Para ver todas as chaves de condição AWS globais, consulte as chaves de contexto de condição AWS global no Guia do usuário do IAM.

O Amazon Kendra não fornece nenhuma chave de condição específica ao serviço, mas oferece suporte ao uso de algumas chaves de condição globais. Para ver todas as chaves de condição AWS globais, consulte Chaves de contexto de condição AWS global no Guia do usuário do IAM.

Exemplos

Para visualizar exemplos de políticas baseadas em identidade do Amazon Kendra, consulte Exemplos de políticas baseadas em identidade do Amazon Kendra.

Políticas baseadas em recursos do Amazon Kendra

O Amazon Kendra não oferece suporte a políticas baseadas em recursos.

Listas de controle de acesso (ACLs)

O Amazon Kendra não oferece suporte às listas de controle de acesso (ACLs) para acesso aos serviços e recursos da AWS .

Autorização baseada em tags do Amazon Kendra

Você pode associar tags a determinados tipos de recursos do Amazon Kendra para autorizar o acesso a esses recursos. Para controlar o acesso baseado em tags, forneça informações sobre as tags no elemento de condição de uma política usando as chaves de condição `aws:RequestTag/key-name` ou `aws:TagKeys`.

A tabela a seguir lista as ações e os tipos de recursos correspondentes para o controle de acesso baseado em tags. Cada ação é autorizada com base nas tags associadas ao tipo de recurso correspondente.

Ação	Tipo de recurso	Chaves de condição
CreateDataFonte		`aws:RequestTag` , `aws:TagKeys`
CreateFaq		`aws:RequestTag` , `aws:TagKeys`
CreateIndex		`aws:RequestTag` , `aws:TagKeys`

Ação	Tipo de recurso	Chaves de condição
API_ ListTags ForResource	fonte de dados, perguntas frequentes, índice	
TagResource	fonte de dados, perguntas frequentes, índice	`aws:RequestTag` , `aws:TagKeys`
UntagResource	fonte de dados, perguntas frequentes, índice	`aws:TagKeys`

Para ter mais informações sobre recursos de marcação do Amazon Kendra, consulte Tags. Para obter um exemplo de política baseada em identidade que limita o acesso a um recurso com base em tags de recurso, consulte Exemplo de política baseada em tags. Para obter informações sobre como usar tags para limitar o acesso aos seus recursos, consulte Controlar o acesso usando tags no Guia do usuário do IAM.

Funções do IAM no Amazon Kendra

Uma função do IAM é uma entidade dentro da sua AWS conta que tem permissões específicas.

Usar credenciais temporárias com o Amazon Kendra

É possível usar credenciais temporárias para fazer login com federação, assumir um perfil do IAM ou assumir um perfil entre contas. Você obtém credenciais de segurança temporárias chamando operações de AWS STS API, como AssumeRoleou GetFederationToken.

O Amazon Kendra oferece suporte ao uso de credenciais temporárias.

Perfis de serviço

Esse atributo permite que um serviço assuma um perfil de serviço em seu nome. O perfil permite que o serviço acesse recursos em outros serviços para concluir uma ação em seu nome. Os perfis de serviço aparecem em sua conta do IAM e são de propriedade da conta. Isso significa que um administrador do IAM pode alterar as permissões para esse perfil. Porém, fazer isso pode alterar a funcionalidade do serviço.

O Amazon Kendra oferece suporte às funções de serviço.

Escolher uma perfil do IAM no Amazon Kendra

Ao criar um índice, chamar a operação da `BatchPutDocument`, criar uma fonte de dados ou uma pergunta frequente, forneça uma função de acesso do nome do recurso da Amazon (ARN) que o Amazon Kendra usa para acessar os recursos necessários em seu nome. Se você já tiver criado uma função, o Amazon Kendra fornecerá uma lista de funções da qual escolher. É importante escolher uma função que permita o acesso aos recursos necessários. Para ter mais informações, consulte IAM funções de acesso para Amazon Kendra.

Exemplos de políticas baseadas em identidade do Amazon Kendra

Por padrão, usuários e perfis não têm permissão para criar ou modificar recursos do Amazon Kendra. Eles também não podem realizar tarefas usando a AWS API AWS Management Console AWS CLI, ou. Um administrador do IAM deve criar políticas do IAM que concedam aos usuários e perfis permissão para executarem operações de API específicas nos recursos especificados de que precisam. O administrador deve anexar essas políticas aos usuários ou grupos que exigem essas permissões.

Para saber como criar uma política baseada em identidade do IAM usando esses exemplos de documentos de política JSON, consulte Criar políticas na guia JSON no Guia do usuário do IAM.

Tópicos

- Melhores práticas de política
- Políticas gerenciadas pela AWS (predefinidas) para o Amazon Kendra
- Permitir que usuários visualizem suas próprias permissões
- Acessando um índice do Amazon Kendra
- Exemplo de política baseada em tags

Melhores práticas de política

As políticas baseadas em identidade determinam se alguém pode criar, acessar ou excluir recursos do Amazon Kendra em sua conta. Essas ações podem incorrer em custos para seus Conta da AWS. Ao criar ou editar políticas baseadas em identidade, siga estas diretrizes e recomendações:

- Comece com as políticas AWS gerenciadas e avance para as permissões de privilégios mínimos — Para começar a conceder permissões aos seus usuários e cargas de trabalho, use as políticas AWS gerenciadas que concedem permissões para muitos casos de uso comuns. Eles estão disponíveis no seu Conta da AWS. Recomendamos que você reduza ainda mais as permissões

definindo políticas gerenciadas pelo AWS cliente que sejam específicas para seus casos de uso. Para obter mais informações, consulte Políticas gerenciadas pela AWS ou Políticas gerenciadas pela AWS para funções de trabalho no Guia do Usuário do IAM.

- Aplique permissões de privilégio mínimo –- ao definir permissões com as políticas do IAM, conceda apenas as permissões necessárias para executar uma tarefa. Você faz isso definindo as ações que podem ser executadas em atributos específicos sob condições específicas, também conhecidas como permissões de privilégio mínimo. Para obter mais informações sobre como usar o IAM para aplicar permissões, consulte Políticas e permissões no IAM no Guia do Usuário do IAM.
- Use condições nas políticas do IAM para restringir ainda mais o acesso –- você pode adicionar uma condição às políticas para limitar o acesso a ações e recursos. Por exemplo, você pode gravar uma condição de política para especificar que todas as solicitações devem ser enviadas usando SSL. Você também pode usar condições para conceder acesso às ações de serviço se elas forem usadas por meio de uma ação específica AWS service (Serviço da AWS), como AWS CloudFormation. Para obter mais informações, consulte Elementos da política JSON do IAM: Condição no Guia do usuário do IAM.
- Use o IAM Access Analyzer para validar suas políticas do IAM a fim de garantir permissões seguras e funcionais –- o IAM Access Analyzer valida as políticas novas e existentes para que elas sigam a linguagem de política do IAM (JSON) e as práticas recomendadas do IAM. O IAM Access Analyzer oferece mais de 100 verificações de política e recomendações acionáveis para ajudá-lo a criar políticas seguras e funcionais. Para obter mais informações, consulte Validação de políticas do IAM Access Analyzer no Guia do Usuário do IAM.
- Exigir autenticação multifator (MFA) — Se você tiver um cenário que exija usuários do IAM ou um usuário root, ative Conta da AWS a MFA para obter segurança adicional. Para exigir MFA quando as operações de API forem chamadas, adicione condições de MFA às suas políticas. Para obter mais informações, consulte Configuração de acesso à API protegido por MFA no Guia do Usuário do IAM.

Para obter mais informações sobre as práticas recomendadas do IAM, consulte Práticas Recomendadas de Segurança no IAM no Guia do Usuário do IAM.

Políticas gerenciadas pela AWS (predefinidas) para o Amazon Kendra

AWS aborda muitos casos de uso comuns fornecendo políticas autônomas do IAM que são criadas e administradas pela AWS. Essas políticas são chamadas de políticas AWS gerenciadas. AWS as políticas gerenciadas facilitam a atribuição de permissões a usuários, grupos e funções do que se

você mesmo tivesse que escrever as políticas. Para obter mais informações, consulte Adicionar permissões a um usuário no Guia do usuário do IAM.

As seguintes políticas AWS gerenciadas, que você pode anexar a grupos e funções em sua conta, são específicas da Amazon Kendra:

- AmazonKendraReadOnly— Concede acesso somente para leitura aos recursos do Amazon Kendra.
- AmazonKendraFullAccess— Concede acesso total para criar, ler, atualizar, excluir, marcar e executar todos os recursos da Amazon Kendra.

Para o console, sua função também deve ter as permissões `iam:CreateRole`, `iam:CreatePolicy`, `iam:AttachRolePolicy`, e `s3:ListBucket`.

> **Note**
>
> Você pode analisar essas políticas de permissões fazendo login no console do IAM e procurando políticas específicas.

Além disso, você pode criar políticas personalizadas do para conceder permissões a ações da API do Amazon Kendra. É possível anexar essas políticas personalizadas a grupos ou perfis do IAM que exijam essas permissões. Para obter exemplos de políticas do IAM do Amazon Kendra, consulte Exemplos de políticas baseadas em identidade do Amazon Kendra.

Permitir que usuários visualizem suas próprias permissões

Este exemplo mostra como criar uma política que permita que os usuários do IAM visualizem as políticas gerenciadas e em linha anexadas a sua identidade de usuário. Essa política inclui permissões para concluir essa ação no console ou programaticamente usando a API AWS CLI ou 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": "*"
        }
    ]
}
```

Acessando um índice do Amazon Kendra

Neste exemplo, você deseja conceder a um usuário da sua AWS conta acesso para consultar um índice.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "QueryIndex",
            "Effect": "Allow",
            "Action": [
                "kendra:Query"
            ],
            "Resource": "arn:aws:kendra:${Region}:${Account}:index/${Index ID}"
        }
    ]
```

```
}
```

Exemplo de política baseada em tags

As políticas baseadas em tags são documentos de políticas JSON que especificam as ações que uma entidade principal pode executar em recursos com tags.

Exemplo: uso de uma tag para acessar um recurso

Este exemplo de política concede a um usuário ou função em sua AWS conta permissão para usar a `Query` operação com qualquer recurso marcado com a chave **`department`** e o valor**`finance`**.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kendra:Query"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/department": "finance"
                }
            }
        }
   ]
}
```

Exemplo: use uma tag para ativar as operações do Amazon Kendra

Este exemplo de política concede a um usuário ou função em sua AWS conta permissão para usar qualquer operação do Amazon Kendra, `TagResource` exceto a operação com qualquer recurso marcado com a **`department`** chave e o valor. **`finance`**

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

```
        },
        {
            "Effect": "Deny",
            "Action": [
                "kendra:TagResource"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/department": "finance"
                }
            }
        }
    ]
}
```

Exemplo: use uma tag para restringir o acesso a uma operação

Este exemplo de política restringe o acesso de um usuário ou função em sua AWS conta para usar a `CreateIndex` operação, a menos que o usuário forneça a **`department`** tag e ela tenha os valores permitidos **`finance`** e. **`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 gerenciadas para o Amazon Kendra

Para adicionar permissões a usuários, grupos e funções, é mais fácil usar políticas AWS gerenciadas do que escrever políticas você mesmo. É necessário tempo e experiência para criar políticas gerenciadas pelo cliente do IAM que fornecem à sua equipe apenas as permissões de que precisam. Para começar rapidamente, você pode usar nossas políticas AWS gerenciadas. Essas políticas abrangem casos de uso comuns e estão disponíveis em sua AWS conta. Para obter mais informações sobre políticas AWS gerenciadas, consulte políticas AWS gerenciadas no Guia do usuário do IAM.

AWS os serviços mantêm e atualizam as políticas AWS gerenciadas. Você não pode alterar as permissões nas políticas AWS gerenciadas. Os serviços ocasionalmente acrescentam permissões adicionais a uma política gerenciada pela AWS para oferecer suporte a novos recursos. Esse tipo de atualização afeta todas as identidades (usuários, grupos e funções) em que a política está anexada. É mais provável que os serviços atualizem uma política gerenciada pela AWS quando um novo recurso for iniciado ou novas operações se tornarem disponíveis. Os serviços não removem as permissões de uma política AWS gerenciada, portanto, as atualizações de políticas não violarão suas permissões existentes.

Além disso, AWS oferece suporte a políticas gerenciadas para funções de trabalho que abrangem vários serviços. Por exemplo, a política de ReadOnlyacesso AWS gerenciado fornece acesso somente de leitura a todos os AWS serviços e recursos. Quando um serviço lança um novo recurso, AWS adiciona permissões somente de leitura para novas operações e recursos. Para obter uma lista e descrições das políticas de funções de trabalho, consulte Políticas gerenciadas pela AWS para funções de trabalho no Guia do usuário do IAM.

AWS política gerenciada: AmazonKendraReadOnly

Concede acesso somente leitura a recursos do Amazon Kendra. Esta política inclui as seguintes permissões:

- `kendra`: permite que os usuários realizem ações que retornem uma lista de itens ou detalhes sobre um item. Isso inclui operações de API que começam com `Describe`, `List`, `Query`, `BatchGetDocumentStatus`, `GetQuerySuggestions` ou `GetSnapshots`.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "kendra:Describe*",
                "kendra:List*",
                "kendra:Query",
                "kendra:BatchGetDocumentStatus",
                "kendra:GetQuerySuggestions",
                "kendra:GetSnapshots"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

AWS política gerenciada: AmazonKendraFullAccess

Concede acesso total para criar, ler, atualizar, excluir e executar todos os recursos do Amazon Kendra. Esta política inclui as seguintes permissões:

- `kendra`: permite que os principais tenham acesso de leitura e gravação a todas as ações no Amazon Kendra.
- `s3`: permite que os principais obtenham a localização dos buckets do Amazon S3 e listem os buckets.
- `iam`: permite que os principais passem e listem as funções.

- kms—Permite que os diretores descrevam e listem AWS KMS chaves e aliases.
- `secretsmanager`: Permite que os principais criem, descrevam e listem as senhas.
- `ec2`: Permite que os principais descrevam grupos de segurança, VCPs (Virtual Private Cloud) e sub-redes.
- `cloudwatch`: permite que os diretores visualizem as métricas do 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": "*"
        }
    ]
}
```

Atualizações do Amazon Kendra para políticas gerenciadas AWS

Veja detalhes sobre as atualizações das políticas AWS gerenciadas do Amazon Kendra desde que esse serviço começou a monitorar essas mudanças. Para receber alertas automáticos sobre mudanças nesta página, assine o feed RSS na página Histórico de documentos do Amazon Kendra.

Alteração	Descrição	Data
AmazonKendraReadOnly—Adicione permissão ao suporte GetSnapshots, BatchGetDocumentStatus APIs	O Amazon Kendra adicionou novas APIs `GetSnapshots` e `BatchGetDocumentStatus` . `GetSnapshots` fornece dados que mostram como os usuários interagem com seu aplicativo de pesquisa. `BatchGetDocumentStatus` monitora o progresso da indexação dos documentos.	3 de janeiro de 2022
AmazonKendraReadOnly—Adicionar permissão para apoiar a operação GetQuerySuggestions	O Amazon Kendra adicionou uma nova API `GetQuerySuggestions` que permite acessar sugestões de consultas para consultas de pesquisa populares , ajudando a orientar a pesquisa dps usuários. Quando os usuários digitam a consulta da pesquisa, a consulta sugerida ajuda a preencher automaticamente a pesquisa.	27 de maio de 2021
O Amazon Kendra começa a monitorar as alterações	A Amazon Kendra começou a monitorar as mudanças em	27 de maio de 2021

Alteração	Descrição	Data
	suas políticas gerenciadas AWS .	

Solução de problemas de identidade e acesso do Amazon Kendra

Use as informações a seguir para ajudar a diagnosticar e corrigir problemas comuns que você pode encontrar ao trabalhar com o Amazon Kendra e o IAM.

Tópicos

- Não tenho autorização para executar uma ação no Amazon Kendra
- Não estou autorizado a realizar iam: PassRole
- Sou administrador e quero permitir que outros usuários tenham acesso ao Amazon Kendra.
- Quero permitir que pessoas fora da minha AWS conta acessem meus recursos do Amazon Kendra

Não tenho autorização para executar uma ação no Amazon Kendra

Se isso AWS Management Console indicar que você não está autorizado a realizar uma ação, entre em contato com o administrador para obter ajuda. Caso seu administrador seja a pessoa que forneceu suas credenciais de início de sessão.

O erro do exemplo a seguir ocorre quando o usuário `mateojackson` tenta usar o console para visualizar detalhes sobre um índice, mas não tem as permissões do `kendra:DescribeIndex`.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 kendra:DescribeIndex on resource: index ARN
```

Neste caso, Mateo pede ao administrador para atualizar suas políticas para permitir a ele o acesso ao recurso `index` usando a ação `kendra:DescribeIndex`.

Não estou autorizado a realizar iam: PassRole

Caso receba uma mensagem de erro informando que você não tem autorização para executar a ação `iam:PassRole`, as políticas deverão ser atualizadas para permitir a transmissão de um perfil ao Amazon Kendra.

Alguns Serviços da AWS permitem que você passe uma função existente para esse serviço em vez de criar uma nova função de serviço ou uma função vinculada ao serviço. Para fazer isso, é preciso ter permissões para passar o perfil para o serviço.

O erro exemplificado a seguir ocorre quando um usuário do IAM chamada `marymajor` tenta usar o console para executar uma ação no Amazon Kendra. No entanto, a ação exige que o serviço tenha permissões concedidas por um perfil de serviço. Mary não tem permissões para passar o perfil para o serviço.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

Nesse caso, as políticas de Mary devem ser atualizadas para permitir que ela realize a ação `iam:PassRole`.

Se precisar de ajuda, entre em contato com seu AWS administrador. Seu administrador é a pessoa que forneceu suas credenciais de login.

Sou administrador e quero permitir que outros usuários tenham acesso ao Amazon Kendra.

Para permitir que outras pessoas acessem o Amazon Kendra, é necessário criar uma entidade do IAM (usuário ou função) para a pessoa ou a aplicação que precisa do acesso. Elas usarão as credenciais dessa entidade para acessar a AWS. Você deve anexar uma política à entidade que concede a eles as permissões corretas no Amazon Kendra.

Para começar a usar imediatamente, consulte Criar os primeiros usuário e grupo delegados pelo IAM no Guia do usuário do IAM.

Quero permitir que pessoas fora da minha AWS conta acessem meus recursos do Amazon Kendra

Você pode criar um perfil que os usuários de outras contas ou pessoas fora da sua organização podem usar para acessar seus recursos. Você pode especificar quem é confiável para assumir o perfil. Para serviços que oferecem compatibilidade com políticas baseadas em recursos ou listas de controle de acesso (ACLs), você pode usar essas políticas para conceder às pessoas acesso aos seus recursos.

Para saber mais, consulte:

- Para saber se o Amazon Kendra é compatível com esses recursos, consulte Como o Amazon Kendra funciona com o IAM.
- Para saber como fornecer acesso aos seus recursos em todos os Contas da AWS que você possui, consulte Como fornecer acesso a um usuário do IAM em outro Conta da AWS que você possui no Guia do usuário do IAM.
- Para saber como fornecer acesso aos seus recursos a terceiros Contas da AWS, consulte Como fornecer acesso Contas da AWS a terceiros no Guia do usuário do IAM.
- Para saber como conceder acesso por meio da federação de identidades, consulte Conceder acesso a usuários autenticados externamente (federação de identidades) no Guia do usuário do IAM.
- Para saber a diferença entre usar funções e políticas baseadas em recursos para acesso entre contas, consulte Acesso a recursos entre contas no IAM no Guia do usuário do IAM.

Melhores práticas de segurança

O Amazon Kendra fornece uma série de recursos de segurança a serem considerados no desenvolvimento e na implementação das suas próprias políticas de segurança. As melhores práticas a seguir são diretrizes gerais e não representam uma solução completa de segurança. Como essas melhores práticas podem não ser adequadas ou suficientes para o seu ambiente, trate-as como considerações úteis em vez de prescrições.

Aplicação do princípio de privilégio mínimo

O Amazon Kendra fornece uma política de acesso granular para aplicativos que usam funções. IAM Recomendamos que os perfis recebam somente o conjunto de privilégios mínimos obrigatórios para o trabalho, como a cobertura da aplicação e o acesso ao destino do log. Também recomendamos auditar regularmente as permissões dos trabalhos e após qualquer alteração da aplicação.

Permissões de controle de acesso por perfil (RBAC)

Os administradores devem controlar rigorosamente as permissões de controle de acesso por perfil (RBAC) para os aplicativos do Amazon Kendra.

Registrar em log e monitorar no Amazon Kendra

O monitoramento é uma parte importante para manter a confiabilidade, a disponibilidade e o desempenho dos aplicativos do Amazon Kendra. Para monitorar as chamadas de API do Amazon

Kendra, você pode usar. AWS CloudTrail Para monitorar o status de seus trabalhos, use o Amazon CloudWatch Logs.

- Amazon CloudWatch Alarms — Usando CloudWatch alarmes, você observa uma única métrica durante um período de tempo especificado por você. Se a métrica exceder uma política. CloudWatch os alarmes não invocam ações quando uma métrica está em um estado específico. O estado deve ter sido alterado e mantido por uma quantidade especificada de períodos. Para ter mais informações, consulte Monitoramento do Amazon Kendra com o Amazon CloudWatch.
- AWS CloudTrail Registros — CloudTrail fornece um registro das ações realizadas por um usuário, função ou AWS serviço no Amazon Kendra ou no Amazon Kendra Intelligent Ranking. Usando as informações coletadas por CloudTrail, você pode determinar a solicitação que foi feita à Amazon Kendra, o endereço IP a partir do qual a solicitação foi feita, quem fez a solicitação, quando ela foi feita e detalhes adicionais. Para obter mais informações, consulte Registro de chamadas de API do Amazon Kendra com os logs do AWS CloudTrail e Registro de chamadas da API do Amazon Kendra Intelligent Ranking com logs do AWS CloudTrail.

Validação de conformidade para o Amazon Kendra

Auditores terceirizados avaliam a segurança e a conformidade do Amazon Kendra como parte de diversos programas de conformidade do Amazon Kendra. O Amazon Kendra está em conformidade com o seguinte:

- Health Insurance Portability and Accountability Act (HIPAA)
- Controles do Sistema e da Organização (CSO) 2
- Programa de avaliadores registrados de segurança da informação (IRAP)
- Programa federal de gerenciamento de autorização e risco (FedRAMP) Moderado nas regiões Leste/Oeste dos EUA
- Programa Federal de Gerenciamento de Riscos e Autorizações (FedRAMP) em alta na região da GovCloud AWS (Oeste dos EUA)

Para obter uma lista de AWS serviços no escopo de programas de conformidade específicos, consulte AWS Serviços no escopo do programa de conformidade AWS . Para obter informações gerais, consulte Programas de AWS conformidade Programas AWS de .

Você pode baixar relatórios de auditoria de terceiros usando AWS Artifact. Para obter mais informações, consulte Baixando relatórios no AWS Artifact .

Sua responsabilidade com relação à compatibilidade ao usar o Amazon Kendra é determinada pela confidencialidade dos seus dados, pelos objetivos de compatibilidade da sua empresa e pelos regulamentos e leis aplicáveis. A AWS fornece os seguintes recursos para ajudar com a compatibilidade:

- Guias de início rápido sobre sobre segurança e conformidade — Esses guias de implantação discutem considerações arquitetônicas e fornecem etapas para a implantação de ambientes básicos com foco em segurança e conformidade em. AWS
- Documento técnico sobre arquitetura para segurança e conformidade com a HIPAA — Este whitepaper descreve como as empresas podem usar para criar aplicativos compatíveis com a HIPAA. AWS
- AWS Recursos de https://aws.amazon.com/compliance/resources/ de conformidade — Essa coleção de pastas de trabalho e guias pode ser aplicada ao seu setor e local.
- Avaliação de recursos com regras no Guia do AWS Config Desenvolvedor — O AWS Config serviço avalia se suas configurações de recursos estão em conformidade com as práticas internas, as diretrizes e os regulamentos do setor.
- AWS Security Hub—Esse AWS serviço fornece uma visão abrangente do seu estado de segurança interno, AWS que ajuda você a verificar sua conformidade com os padrões e as melhores práticas do setor de segurança.

Resiliência no Amazon Kendra

A infraestrutura AWS global é construída em torno de AWS regiões e zonas de disponibilidade. AWS As regiões fornecem várias zonas de disponibilidade fisicamente separadas e isoladas, conectadas a redes de baixa latência, alta taxa de transferência e alta redundância. Com as zonas de disponibilidade, é possível projetar e operar aplicativos e bancos de dados que automaticamente executam o failover entre as zonas sem interrupção. As zonas de disponibilidade são altamente disponíveis, tolerantes a falhas e escaláveis que uma ou várias infraestruturas de datacenters tradicionais.

Para obter mais informações sobre AWS regiões e zonas de disponibilidade, consulte Infraestrutura AWS global.

Com uma infraestrutura AWS global, o Amazon Kendra Enterprise Edition é tolerante a falhas, escalável e altamente disponível. Atualmente, não há suporte para o versionamento das versões anteriores de um índice, mas é possível atualizar ou recriar partes do índice excluindo e adicionando fontes de dados existentes novamente no índice.

Segurança da infraestrutura no Amazon Kendra

Como um serviço gerenciado, o Amazon Kendra é protegido AWS pela segurança de rede global. Para obter informações sobre serviços AWS de segurança e como AWS proteger a infraestrutura, consulte AWS Cloud Security. Para projetar seu AWS ambiente usando as melhores práticas de segurança de infraestrutura, consulte Proteção de infraestrutura no Security Pillar AWS Well-Architected Framework.

Você usa chamadas de API AWS publicadas para acessar o Amazon Kendra pela rede. Os clientes devem oferecer suporte para:

- Transport Layer Security (TLS). Exigimos TLS 1.2 e recomendamos TLS 1.3.
- Conjuntos de criptografia com perfect forward secrecy (PFS) como DHE (Ephemeral Diffie-Hellman) ou ECDHE (Ephemeral Elliptic Curve Diffie-Hellman). A maioria dos sistemas modernos, como Java 7 e versões posteriores, comporta esses modos.

Além disso, as solicitações devem ser assinadas usando um ID da chave de acesso e uma chave de acesso secreta associada a uma entidade principal do IAM. Ou você pode usar o AWS Security Token Service (AWS STS) para gerar credenciais de segurança temporárias para assinar solicitações.

Análise de configuração e vulnerabilidade em AWS Identity and Access Management

AWS lida com tarefas básicas de segurança, como sistema operacional (SO) convidado e aplicação de patches em bancos de dados, configuração de firewall e recuperação de desastres. Esses procedimentos foram revisados e certificados por terceiros certificados. Para obter mais detalhes, consulte os seguintes recursos da :

- Modelo de responsabilidade compartilhada
- AWS: visão geral dos processos de segurança (whitepaper)

Os recursos a seguir também abordam a configuração e a análise de vulnerabilidades em AWS Identity and Access Management (IAM):

- Validação de conformidade para AWS Identity and Access Management

- Melhores práticas de segurança e casos de uso em AWS Identity and Access Management.

Cotas para Amazon Kendra

Regiões compatíveis da

Para obter uma lista das AWS regiões onde Amazon Kendra está disponível, consulte Amazon Kendra regiões e endpoints na Referência geral da Amazon Web Services.

Cotas

As cotas de serviço, também chamadas de limites, são o número máximo de recursos de serviço para sua AWS conta. Para obter mais informações, consulte Service Quotas da Amazon Kendra, na Referência geral da AWS .

Cotas de índice

Descrição	Padrão	Edição	Ajustável
Número máximo de índices por conta	10	Desenvolvedor, Empresa	Sim
Quantidade de texto extraída para um índice em uma única unidade (Desenvolvedor). Não é possível adicionar unidades extras para extrair texto para o Developer Edition.	3 GB	Desenvolvedor	Não
Quantidade de texto extraída para um índice em uma única unidade (Enterprise). Adicione até	30 GB	Enterprise	Sim

Descrição	Padrão	Edição	Ajustável
100 unidades extras para extrair texto para o Enterpris e Edition ou basta entrar em contato com o Suporte.			

Cotas de conectores de fonte de dados

Descrição	Padrão	Edição	Ajustável
Número máximo de conectores de fonte de dados por índice (desenvolvedor)	5	Desenvolvedor	Não
Número máximo de conectores de fonte de dados por índice (Enterprise)	50	Enterprise	Sim
Tamanho máximo de um único documento ou arquivo bruto ao usar um conector de fonte de dados	50 MB	Desenvolvedor, Empresa	Sim
Número máximo de prefixos S3 no arquivo de configuração da lista de controle de acesso incluído no conector da fonte Amazon S3 de dados	100	Desenvolvedor, Empresa	Não

Descrição	Padrão	Edição	Ajustável
Tamanho máximo do arquivo de configuração da lista de controle de acesso incluído no conector da fonte de Amazon S3 dados	50 MB	Desenvolvedor, Empresa	Sim

Perguntas frequentes sobre cotas

Descrição	Padrão	Edição	Ajustável
Número máximo de perguntas frequentes por índice	30	Desenvolvedor, Empresa	Sim
Tamanho máximo de 1 Perguntas frequentes	5 MB	Desenvolvedor, Empresa	Sim
Número máximo de resultados retornados para perguntas frequentes	4	Desenvolvedor, Empresa	Sim
Número máximo de caracteres permitido para uma pergunta de FAQ	300	Desenvolvedor, Empresa	Não
Número máximo de caracteres em uma resposta de perguntas frequentes	2000	Desenvolvedor, Empresa	Não

Cotas do dicionário de sinônimos

Descrição	Padrão	Edição	Ajustável
Número máximo de dicionários de sinônimos por índice	1	Desenvolvedor, Empresa	Não
Tamanho máximo de um arquivo de dicionário de sinônimos	5 MB	Desenvolvedor, Empresa	Sim
Número máximo de regras de sinônimos por dicionário de sinônimos	10.000	Desenvolvedor, Empresa	Sim
Número máximo de sinônimos por termo em todos os dicionários de sinônimos em um índice	10	Desenvolvedor, Empresa	Não

Amazon Kendra cotas de experiência

Descrição	Padrão	Edição	Ajustável
Número máximo de Amazon Kendra experiências por índice	50	Desenvolvedor, Empresa	Sim

Cotas de resultados de consulta e pesquisa

Descrição	Padrão	Edição	Ajustável
Quantidade de consultas por segundo para um índice em uma única unidade (Desenvolvedor). Não é possível adicionar unidades extras para consultas no Developer Edition.	0,05	Desenvolvedor	Não
Quantidade de consultas por segundo para um índice em uma única unidade (Enterprise). Adicione até 100 unidades extras para consultas ao Enterprise Edition ou basta entrar em contato com o Suporte.	0.1	Enterprise	Sim
Número máximo de caracteres por texto de consulta	1000	Desenvolvedor, Empresa	Sim
Número máximo de resultados de pesquisa por consulta. O padrão é de 100. Para obter mais de 100	100	Desenvolvedor, Empresa	Sim

Descrição	Padrão	Edição	Ajustável
resultados, basta entrar em contato com o Suporte.			
Número máximo de resultados de pesquisa por página	100	Desenvolvedor, Empresa	Sim
Número máximo de palavras simbólicas por texto de consulta antes do truncamento. O padrão é de 30. Para permitir mais de 30 palavras, basta entrar em contato com o Suporte.	30	Desenvolvedor, Empresa	Sim
Tamanho máximo da lista de grupos de usuários por atributo de consulta	1000	Desenvolvedor, Empresa	Sim
Tamanho máximo da lista de sequência de caracteres por atributo de consulta	10	Desenvolvedor, Empresa	Sim

Cotas de sugestões de consulta

Descrição	Padrão	Edição	Ajustável
Número máximo de sugestões de consulta retornada	10	Desenvolvedor, Empresa	Sim

Descrição	Padrão	Edição	Ajustável
s por chamada de GetQuerysugestões			
Número máximo de campos/atributos para sugestões de consulta por chamada de sugestões GetQuery	10	Desenvolvedor, Empresa	Sim
Número máximo de campos/atributos adicionais para sugestões de consulta por chamada de sugestões GetQuery	5	Desenvolvedor, Empresa	Sim
Número máximo de listas de bloqueio por índice	1	Desenvolvedor, Empresa	Não
Tamanho máximo de um arquivo de texto de lista de bloqueio	2 MB	Desenvolvedor, Empresa	Sim
Número máximo de itens (palavras ou frases) em uma lista de bloqueio	20.000	Desenvolvedor, Empresa	Sim
Número máximo de sugestões de consulta com correção ortográfica a serem retornadas em uma chamada de API de`Query`.	1	Desenvolvedor, Empresa	Sim

Cotas de documentos

Descrição	Padrão	Edição	Ajustável
Quantidade de texto extraída para um índice em uma única unidade (Desenvolvedor). Não é possível adicionar unidades extras para extrair texto para o Developer Edition.	3 GB	Desenvolvedor	Não
Quantidade de texto extraída para um índice em uma única unidade (Enterprise). Adicione até 100 unidades extras para extrair texto para o Enterprise Edition ou basta entrar em contato com o Suporte.	30 GB	Enterprise	Sim
Tamanho máximo de um único documento ou arquivo bruto ao usar um conector de fonte de dados	50 MB	Desenvolvedor, Empresa	Sim
Tamanho máximo de um único documento ou arquivo bruto ao usar a `BatchPutDocument` API	5 MB	Desenvolvedor, Empresa	Sim

Descrição	Padrão	Edição	Ajustável
Quantidade máxima de texto extraída de um único documento	5 MB	Desenvolvedor, Empresa	Não
Número máximo de campos/atributos personalizados por índice	500	Desenvolvedor, Empresa	Não

Cotas de resultados de pesquisa em destaque

Descrição	Padrão	Edição	Ajustável
Número máximo de documentos em destaque por conjunto de resultados em destaque	4	Enterprise	Sim
Número máximo de textos de consulta por conjunto de resultados em destaque	49	Enterprise	Não
Número máximo de caracteres por texto de consulta em um conjunto de resultados em destaque	1000	Enterprise	Sim
Número máximo de conjuntos de resultados em destaque por índice	50	Enterprise	Sim

Rescore/reclassifique as cotas dos resultados da pesquisa

Descrição	Padrão	Edição	Ajustável
Número máximo de solicitações de `Rescore` por segundo para um plano de execução de reclassificação ou uma única unidade de capacidade. Você poderá adicionar até 100 unidades de capacidade adicional.	0,01	Enterprise	Não
Número máximo de planos de execução de reclassificação por conta.	50	Enterprise	Sim
Número máximo de tokens no `Title` para um documento em uma solicitação de `Rescore`.	100	Enterprise	Não
Número máximo de tokens no `Body` para um documento em uma solicitação de `Rescore`.	200	Enterprise	Não
Número máximo de documentos em uma solicitação de `Rescore`.	25	Enterprise	Não

Descrição	Padrão	Edição	Ajustável
Número máximo de documento s por grupo em uma solicitação de `Rescore`.	3	Enterprise	Não

Para obter mais informações sobre cotas Amazon Kendra de serviço e para solicitar um aumento de cota, consulte Cotas de serviço.

Solução de problemas

Esta seção pode ajudá-lo a resolver problemas comuns que você pode encontrar ao trabalhar com Amazon Kendra.

Tópicos

- Solucionar problemas de origens de dados
- Solucionar problemas de resultados da pesquisa de documentos
- Solução de problemas gerais

Solucionar problemas de origens de dados

Esta seção pode ajudá-lo a resolver problemas comuns ao configurar e usar conectores de fonte Amazon Kendra de dados.

Meus documentos não foram indexados

Ao sincronizar seu Amazon Kendra índice com uma fonte de dados, você pode ter problemas que impedem que os documentos sejam indexados. A indexação é um processo em duas etapas. Primeiro, a fonte de dados é verificada em busca de documentos novos e atualizados para indexar e para encontrar documentos a serem removidos do índice. Segundo, no nível do documento, cada documento é acessado e indexado.

Um erro pode ocorrer em qualquer uma dessas etapas. Os erros no nível da fonte de dados são relatados no console na seção Histórico de execução da sincronização da página de detalhes da fonte de dados. O status da tarefa de sincronização pode ser Bem-sucedido, Incompleto ou Falha. Visualize também o número de documentos indexados e excluídos durante o trabalho. Se o status for Falha, uma mensagem será exibida na coluna Detalhes.

Os erros no nível do documento são relatados em Amazon CloudWatch Logs. Você pode ver os erros usando o CloudWatch console.

Para gerar um relatório de status de sincronização de documentos, consulte Quero gerar um relatório de status de sincronização para meus documentos.

Meu trabalho de sincronização falhou

Normalmente, um trabalho de sincronização falha quando há um erro de configuração no índice ou na fonte de dados. No console, encontre a mensagem de erro na seção Histórico de execução da sincronização da página de detalhes da fonte de dados, na coluna Detalhes. Os erros no nível do documento são relatados no Amazon CloudWatch Logs. A mensagem de erro fornece informações sobre o que deu errado. O problema geralmente é que o índice ou a fonte de dados não tem as IAM permissões adequadas. A mensagem de erro descreve as permissões ausentes. Aqui estão algumas das mensagens de erro que você poderá receber:

`Failed to create log group for job. Please make sure that the IAM role provided has sufficient permissions.`

Se sua função de índice não tiver permissão de uso CloudWatch, a fonte de dados não poderá criar um CloudWatch registro. Se você receber esse erro, deverá adicionar CloudWatch permissões à função 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.`

Ao usar uma fonte de Amazon S3 dados, você Amazon Kendra deve ter permissão para acessar o bucket que contém os documentos. Você precisa adicionar permissão Amazon Kendra para ler o bucket à IAM função de fonte de dados.

`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 precisa de permissão para assumir as IAM funções de índice e fonte de dados. Adicione uma política de confiança às funções com permissão para a ação de `sts:AssumeRole`.

Para as IAM políticas que Amazon Kendra precisam indexar uma fonte de dados, consulte IAM funções.

Para gerar um relatório de status de sincronização de documentos, consulte Quero gerar um relatório de status de sincronização para meus documentos.

Meu trabalho de sincronização está incompleto

Os trabalhos geralmente ficam incompletos quando concluem o processo no nível da fonte de dados, mas apresentam alguns erros durante o processo no nível do documento. Quando um trabalho está

incompleto, alguns dos documentos podem não ter sido indexados com êxito. Para uma fonte de dados do Amazon S3 , um trabalho incompleto geralmente é causado por:

- Os metadados de um ou mais documentos eram inválidos.
- Quando os documentos são enviados para indexação, mas pelo menos um documento não foi enviado.
- Quando os documentos são enviados para exclusão do índice, mas pelo menos um documento não foi enviado.

Para solucionar um trabalho de sincronização incompleto, consulte primeiro seus CloudWatch registros.

1. Na coluna de detalhes, escolha Exibir detalhes em CloudWatch.
2. Examine as mensagens de erro para visualizar o que causou a falha do documento.

Para gerar um relatório de status de sincronização de documentos, consulte Quero gerar um relatório de status de sincronização para meus documentos.

Meu trabalho de sincronização foi bem-sucedido, mas não há documentos indexados

Ocasionalmente, a execução de uma tarefa de sincronização de índice será marcada como bem-sucedida, mas não há documentos novos ou atualizados indexados conforme o esperado. Os possíveis motivos incluem:

- Verifique a CloudWatch `DocumentsSubmittedForIndexingFailed` métrica para ver se algum documento falhou na sincronização. Verifique seus CloudWatch registros para obter detalhes.
- Para uma fonte Amazon S3 de dados, você pode ter fornecido Amazon Kendra o nome ou prefixo incorreto do bucket. Verifique se o bucket que Amazon Kendra está usando é aquele que contém os documentos a serem indexados.
- Ao reindexar um documento que não foi indexado em um trabalho anterior, o Amazon Kendra não o indexará, a menos que você tenha alterado o documento ou o arquivo de metadados associado.

Para gerar um relatório de status de sincronização de documentos, consulte Quero gerar um relatório de status de sincronização para meus documentos.

Estou enfrentando problemas de formato de arquivo ao sincronizar minha fonte de dados

Se tiver problemas de formato de arquivo ao adicionar arquivos à sua fonte de dados ou sincronizar sua fonte de dados, verifique se os tipos de documentos têm suporte pelo Amazon Kendra . Para obter uma lista dos tipos de documentos suportados pelo, Amazon Kendra consulte Tipos ou formatos de documentos.

Se estiver usando a API de `BatchPutDocument` com arquivos de texto simples, especifique o `PLAIN_TEXT` como tipo de conteúdo.

Quero gerar um relatório de status de sincronização para meus documentos

Quando você sincroniza seu conector de fonte de Amazon Kendra dados, Amazon Kendra pode gerar relatórios de status de sincronização para cada documento em sua fonte de dados e copiá-los para um Amazon S3 bucket. Durante esse processo, seus dados são criptografados usando chaves do AWS KMS e só podem ser visualizados por você. O status do documento relatado pode ser um dos seguintes: Falha, Concluído ou Bem-sucedido com erros.

Antes de gerar relatórios de status de sincronização, você deverá fazer o seguinte:

- Adicione o seguinte principal Amazon Kendra de serviço à sua política de Amazon S3 acesso

```
{
    “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/*”
        }
    ]
}
```

- Crie um Amazon S3 bucket com permissões de acesso para Amazon Kendra

No console, para gerar um relatório de status de sincronização, opte por ativar a opção Geração do histórico de sincronização na página Detalhes da fonte de dados. Em seguida, insira a localização do Amazon S3 bucket e escolha entre as opções de configuração disponíveis. Os relatórios serão gerados na próxima sincronização depois que você ativar a geração de relatórios.

Se você excluir o Amazon S3 bucket, perderá seus dados de log e precisará configurar um novo bucket para armazenar novos relatórios de sincronização.

No momento, a geração do status de relatórios de sincronização é compatível somente com o conector do Amazon S3.

Quanto tempo demora a sincronização de uma fonte de dados?

Se não houver atualizações nos documentos, o tempo de sincronização de um Amazon Kendra índice aumenta em proporção linear ao número de documentos. Por exemplo, 1.000 documentos sem nenhuma atualização levariam cerca de cinco minutos para serem sincronizados e 2.000 documentos sem nenhuma atualização levariam cerca de 10 minutos. Se houver alguma atualização nos documentos, o tempo de sincronização aumentará com base no número de documentos atualizados.

Qual é a cobrança pela sincronização de uma fonte de dados?

Quando você sincroniza seu índice, leva dois minutos para aquecer e ativar Amazon EC2 para estabelecer as conexões necessárias. Você não é cobrado durante esse processo. Seu medidor de uso começa somente após o início da tarefa de sincronização. Para obter mais informações sobre Amazon Kendra preços, consulte Amazon Kendra preços.

Estou recebendo um erro Amazon EC2 de autorização

Se ocorrer um erro de operação Amazon EC2 não autorizada durante a sincronização de uma fonte de dados de nuvem privada virtual (VPC), é provável que sua função de IAM VPC não tenha as permissões necessárias. Verifique se a IAM função que você usa para sua fonte de dados tem as permissões anexadas. Para obter mais informações, consulte IAM Função de nuvem privada virtual.

Não consigo usar links de índice de pesquisa para abrir meus Amazon S3 objetos

Seu Amazon Kendra índice só pode acessar arquivos que uma fonte de Amazon S3 dados concede permissão para acessar. Por exemplo, Amazon Kendra não é possível modificar as Amazon S3

permissões que determinam se um objeto deve ser público ou criptografado. Amazon Kendra também não tem as permissões padrão para criar ou retornar um link assinado para Amazon S3 objetos. Se você quiser ativar a vinculação assinada para Amazon S3 objetos em um Amazon Kendra índice, você tem duas opções:

- Assine os resultados da consulta de índice com o objeto URI de origem antes de retornar o resultado à página de pesquisa. Para ver um step-by-step passo a passo desse processo, consulte Compartilhamento de objetos usando URLs pré-assinados.
- Você pode substituir o uri da fonte de metadados do Amazon S3 objeto e disponibilizar seu serviço por meio de uma rede de distribuição de CloudFront conteúdo (CDN) conectada a um bucket. Amazon S3 Ou você pode usar um endpoint de API Gateway proxy que retorna uma URL pré-assinada e redireciona para ela.

Estou recebendo uma mensagem de erro AccessDenied Ao usar o arquivo de certificado SSL

Se você estiver recebendo um erro de acesso negado ao usar um certificado SSL com sua fonte de dados, certifique-se de que sua IAM função tenha permissão para acessar o arquivo do certificado SSL no local especificado. Se o certificado for criptografado com uma AWS KMS chave, sua IAM função também deverá ter permissão para descriptografar usando a chave. AWS KMS Para obter mais informações, consulte Autenticação e controle de acesso para o AWS KMS.

Estou recebendo um erro de autorização ao usar uma fonte SharePoint de dados

Se você estiver recebendo um erro de autorização ao sincronizar seu índice com uma fonte de SharePoint dados, confirme se você tem uma função de administrador do site atribuída a você em SharePoint.

Meu índice não rastreia documentos da minha fonte de dados do Confluence

Se seu Amazon Kendra índice não estiver rastreando documentos da sua fonte de dados do Confluence durante o processo de sincronização, confirme se você faz parte dos grupos de administradores no Confluence.

Solucionar problemas de resultados da pesquisa de documentos

Esta seção pode ajudá-lo a corrigir problemas nos resultados Amazon Kendra da pesquisa.

Meus resultados de pesquisa não são relevantes para minha consulta de pesquisa

Se os resultados da pesquisa parecerem irrelevantes, talvez seja pelos seguintes motivos:

- Resultados com confiança `LOW` são incluídos nos resultados. Você pode filtrar os resultados com `LOW` confiança usando QueryResultItemo `ScoreAttributes` campo s para excluir qualquer resultado com um valor de`LOW`. Amazon Kendra atribui a cada resultado um valor de intervalo de confiança de `VERY_HIGHHIGH`, `MEDIUM` e. `LOW` Esses valores indicam o nível de confiança de que um resultado é relevante para uma consulta. Além disso, independentemente dos intervalos de confiança, Amazon Kendra retorna três tipos de resultados na seguinte ordem: `ANSWER` (trecho da resposta sugerida), (FAQ) e `QUESTION_ANSWER` `DOCUMENT` (trecho do documento). Portanto, é possível que um resultado da `QUESTION_ANSWER` de confiança `LOW` seja posicionado acima de um resultado do `DOCUMENT` de confiança `VERY_HIGH`. No entanto, nem sempre é necessariamente verdade que a confiança `LOW` da `QUESTION_ANSWER` é um resultado melhor do que a confiança `VERY_HIGH` do `DOCUMENT`.
- Certos campos ou atributos de metadados são aumentados para um valor muito alto, afetando a classificação dos resultados. Amazon Kendra pesquisa seu índice usando vários parâmetros, como título do documento, texto, data e campos ou atributos de texto personalizados. Você poderá experimentar diferentes valores de aumento para obter os melhores resultados em todas as consultas. Você também poderá usar o ajuste dinâmico de relevância no nível da consulta para usar valores de aumento diferentes para cada consulta.
- Seus usuários estão usando termos especializados quando consultam informações e não há sinônimos personalizados configurados em seu índice para lidar com esses termos especializados. Para obter mais detalhes sobre como e quando usar sinônimos, consulte Adicionar sinônimos personalizados a um índice.

Por que eu só vejo 100 resultados?

Amazon Kendra retorna a contagem total de documentos relevantes. Por padrão, os 100 principais são retornados por consulta. Os resultados são paginados. Use o `PageNumber` para acessar páginas diferentes.

Você pode configurar Amazon Kendra para retornar até 1.000 documentos ou resultados de pesquisa por consulta, com até 100 resultados por página. Para retornar mais de 100 resultados, você poderá solicitar isso entrando em contato com o Suporte a cotas. Aumentar o número de resultados da pesquisa pode afetar a latência.

Por que os documentos que eu espero ver estão ausentes?

Amazon Kendra suporta listas de controle de acesso (ACLs) com base em usuários e grupos. Amazon Kendra ingere políticas de ACL por meio de conectores. Se um índice não configurar uma ACL, somente documentos que correspondam ao filtro de atributos para usuário e grupo serão exibidos. Se um filtro de atributo de usuário ou grupo for fornecido, os documentos sem uma ACL não serão exibidos.

Se estiver usando controle de acesso baseado em tokens, documentos sem uma política de ACL e documentos que correspondam ao usuário e aos grupos serão exibidos.

Por que vejo documentos que têm uma política de ACL?

Se um índice não configurar uma política de controle de acesso, usuários e grupos poderão ser fornecidos pelo filtro. Se nenhum filtro de usuário e grupo for aplicado, todos os documentos relacionados serão retornados. Qualquer política de ACL será ignorada.

Solução de problemas gerais

Amazon Kendra usa CloudWatch métricas e registros para fornecer informações sobre a sincronização de suas fontes de dados. Use as métricas e os registros para determinar o que deu errado com uma execução de sincronização e como corrigi-lo.

Para solucionar problemas gerais, comece com suas CloudWatch métricas.

- Verifique a métrica do `DocumentsCrawled` para visualizar quantos documentos sua fonte de dados verificou. Para um Amazon S3 intervalo, se o número for menor do que o esperado, verifique se sua fonte de dados está apontando para o intervalo certo.
- Verifique a métrica do `DocumentsSkippedNoChange` para visualizar quantos documentos foram ignorados porque não foram alterados desde a última sincronização. Se o número não corresponder ao esperado, verifique se o repositório foi atualizado corretamente.
- Verifique a métrica do `DocumentsSkippedInvalidMetadata` para visualizar quantos documentos tinham metadados inválidos. Verifique seus CloudWatch registros para ver os erros específicos que ocorreram.

- Verifique a métrica do `DocumentsSubmittedForIndexingFailed` para visualizar quantos documentos foram enviados da fonte de dados para o índice, mas não foram indexados. Por exemplo, se usar um atributo de metadados em uma fonte de dados do Amazon S3 que não tenha sido definida como um campo de índice personalizado, o documento não será indexado. Verifique seus CloudWatch registros para ver os erros específicos que ocorreram.
- Verifique a métrica `DocumentsSubmittedForDeletionFailed` para visualizar quantos documentos que a fonte de dados tentou remover do índice não foram excluídos do índice. Verifique seus CloudWatch registros para ver os erros específicos que ocorreram.

Você pode consultar os CloudWatch registros de uma execução de sincronização específica para obter detalhes dos erros que ocorreram durante a execução. Para obter mais informações sobre CloudWatch registros com Amazon Kendra, consulte CloudWatch Logs.

Intelligent Ranking do Amazon Kendra

O Intelligent Ranking do Amazon Kendra usa recursos de pesquisa semântica do Amazon Kendra para reclassificar de forma inteligente os resultados de um serviço de pesquisa.

Tópicos

- Amazon Kendra Classificação inteligente para autogestão OpenSearch
- Classificando semanticamente os resultados de um serviço de pesquisa

Amazon Kendra Classificação inteligente para autogestão OpenSearch

Você pode aproveitar os recursos Amazon Kendra de pesquisa semântica do Apache 2.0 para melhorar os resultados de OpenSearchpesquisa do serviço de pesquisa autogerenciado de código aberto baseado na Licença Apache 2.0. O plugin Amazon Kendra Intelligent Ranking reclassifica semanticamente os resultados OpenSearch usando. Amazon Kendra Ele faz isso entendendo o significado e o contexto de uma consulta de pesquisa usando campos específicos, como o corpo ou o título do documento, a partir dos resultados de OpenSearch pesquisa padrão.

Veja, por exemplo, esta consulta: “endereço principal da palestra”. Como “endereço” tem vários significados, é Amazon Kendra possível inferir o significado por trás da consulta para retornar informações relevantes alinhadas com o significado pretendido. Nesse contexto, é um discurso de abertura da conferência. Um serviço de pesquisa mais simples pode não levar em conta a intenção e possivelmente retornar resultados para um endereço na Main Street, por exemplo.

O plug-in Intelligent Ranking para OpenSearch está disponível para a versão 2.4.0 OpenSearch (autogerenciada) e posterior. Você pode instalar o plug-in usando um script Bash de início rápido para criar uma nova imagem do Docker OpenSearch com o plug-in Intelligent Ranking incluído. Veja o Configurando o plug-in de pesquisa inteligente. Ele é um exemplo de configuração para você começar a trabalhar rapidamente.

Como funciona o plug-in de pesquisa inteligente

O processo geral do plugin Intelligent Ranking para OpenSearch (autogerenciado) é o seguinte:

1. Um OpenSearch usuário emite uma consulta e OpenSearch fornece uma resposta à consulta ou uma lista de documentos que são relevantes para a consulta.

2. O plug-in do Intelligent Ranking pega a resposta da consulta e extrai informações dos documentos.
3. O plugin Intelligent Ranking faz uma chamada para a API Rescore do Amazon Kendra Intelligent Ranking.
4. A API `Rescore` pega as informações extraídas dos documentos e reclassifica semanticamente os resultados da pesquisa.
5. A API `Rescore` envia os resultados da pesquisa reclassificados de volta ao plug-in. O plug-in reorganiza os resultados da pesquisa na resposta da OpenSearch pesquisa para refletir a nova classificação semântica.

O plugin Intelligent Ranking reclassifica os resultados usando os campos “corpo” e “título”. Esses campos de plug-in podem ser mapeados para campos em seu OpenSearch índice que melhor se ajustem à definição de corpo e título de um documento. Por exemplo, o índice contém capítulos de um livro com campos como “chapter_title” e “chapter_contents”, mepeie primeiro para “título” e depois o “corpo” para obter os melhores resultados.

Configurando o plug-in de pesquisa inteligente

A seguir, descrevemos como configurar rapidamente OpenSearch (autogerenciado) com o plug-in Intelligent Ranking.

Configuração OpenSearch (autogerenciada) com o plug-in Intelligent Ranking (configuração rápida)

Se você já estiver usando a imagem do Docker`opensearch:2.4.0`, poderá usar esse Dockerfile para criar uma nova imagem da versão OpenSearch 2.4.0 com o plug-in Intelligent Ranking. Você inclui um contêiner para a nova imagem no arquivo docker-compose.yml ou no arquivo opensearch.yml. Você também inclui o ID do plano de execução de repontuação gerado pela criação de um plano de execução de repontuação, junto com as informações da região e do endpoint. Consulte a etapa 2 para criar um plano de execução de repontuação.

Se você já baixou uma versão da imagem do Docker `opensearch` anterior à 2.4.0, você deve usar a imagem do Docker `opensearch:2.4.0` ou posterior e criar uma nova imagem com o plug-in do Intelligent Ranking incluído.

1. Baixe e instale o Docker Desktop para seu sistema operacional. O Docker Desktop inclui o Docker Compose e o Docker Engine. É recomendável verificar se o computador atende aos requisitos de sistema mencionados nos detalhes de instalação do Docker.

Você também pode aumentar os requisitos de uso de memória nas configurações do Docker Desktop. Você é responsável pelos requisitos de uso do Docker fora dos limites de uso disponíveis gratuitamente para os serviços do Docker. Consulte as assinaturas do Docker.

Verifique se o status do Docker Desktop está “em execução”.

2. Provisione o Amazon Kendra Intelligent Ranking e seus requisitos de capacidade. Depois de provisionar o Intelligent Ranking do Amazon Kendra , você é cobrado por hora com base nas unidades de capacidade definidas. Veja as informações sobre o nível gratuito e os preços.

   Você usa a CreateRescoreExecutionPlanAPI para provisionar `Rescore API` o. Se não precisar de mais unidades de capacidade do que a unidade padrão, não adicione mais unidades e forneça apenas um nome para o plano de execução de repontuação. Você também pode atualizar seus requisitos de capacidade usando a UpdateRescoreExecutionPlanAPI. Para obter mais informações, consulte Classificação semântica dos resultados de um serviço de pesquisa.

   Opcionalmente, vá para a etapa 3 para criar um plano de execução de repontuação padrão ao executar o script Bash de início rápido.

   Para a etapa 4, o ID do plano de execução de repontuação está incluído na resposta.

   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. Baixe o script Bash de início rápido GitHub para sua versão do OpenSearch selecionando a ramificação da versão no menu suspenso da ramificação principal.

   Esse script usa imagens do Docker OpenSearch e OpenSearch painéis usando a versão que você selecionou no GitHub repositório para o script. Ele baixa um arquivo zip para o plug-in Intelligent Ranking e gera um `Dockerfile` para criar uma nova imagem do Docker OpenSearch que inclua o plug-in. Ele também cria um arquivo docker-compose.yml que inclui contêineres OpenSearch com o plug-in e os painéis do Intelligent Ranking. OpenSearch O script adiciona o ID do plano de execução do rescore, as informações da região e o endpoint (usa a região) ao arquivo docker-compose.yml. Em seguida, o script é executado `docker-compose up` para iniciar os contêineres OpenSearch com o Intelligent Ranking incluído e os OpenSearch painéis. Para parar os contêineres sem removê-los, execute `docker-compose stop`. Para remover os contêineres, execute `docker-compose down`.

4. Abra o terminal e, no diretório do script Bash, execute o seguinte comando:

   ```
   bash search_processing_kendra_quickstart.sh -p <execution-plan-id> -r <region>
   ```

   Ao executar esse comando, você fornece o ID do plano de execução de rescore que você anotou na etapa 2 ao provisionar o Amazon Kendra Intelligent Ranking, junto com as informações da sua região. Opcionalmente, você pode provisionar o Intelligent Ranking do Amazon Kendra usando a opção `--create-execution-plan`. Isso cria um plano de execução de repontuação com nome e capacidade padrão.

   Para não perder seu índice quando o contêiner temporário padrão for removido, você pode fazer com que seu índice persista em todas as execuções fornecendo o nome do volume de dados usando a opção `--volume-name`. Se você criou um índice anteriormente, pode especificar o volume no arquivo docker-compose.yml ou opensearch.yml. Para deixar os volumes intactos, não execute`docker-compose down -v`.

   O script Bash de início rápido configura suas AWS credenciais no OpenSearch armazenamento de chaves para se conectar ao Intelligent Ranking. Amazon Kendra Para fornecer suas AWS credenciais ao script, use a `--profile` opção para especificar o AWS perfil. Se a `--profile` opção não for especificada, o script Bash de início rápido tentará ler as AWS credenciais (chave de acesso/secreta, token de sessão opcional) das variáveis de ambiente e, em seguida, do perfil padrão. AWS Se a `--profile` opção não for especificada e nenhuma credencial for encontrada, o script não passará as credenciais para o OpenSearch armazenamento de chaves. Se nenhuma credencial for especificada no OpenSearch repositório de chaves, o plug-in ainda

verificará as credenciais na cadeia de fornecedores de credenciais padrão, incluindo credenciais de Amazon ECS contêiner ou credenciais de perfil de instância fornecidas pelo serviço de metadados. Amazon EC2

Certifique-se de ter criado uma IAM função com as permissões necessárias para invocar o Amazon Kendra Intelligent Ranking. Veja a seguir um exemplo de uma IAM política para conceder permissão para usar a `Rescore` API para um plano de execução 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}"
        }
    ]
}
```

Exemplo de docker-compose.yml

Um exemplo de um arquivo docker-compose.yml usando OpenSearch 2.4.0 ou posterior com o plug-in Intelligent Ranking e o Dashboards 2.4.0 ou 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
```

Exemplo de um Dockerfile e construção de uma imagem

Um exemplo de um `Dockerfile` para usar a OpenSearch versão 2.4.0 ou posterior com o plug-in Intelligent Ranking.

```
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
```

Construindo uma imagem do Docker OpenSearch com o plug-in Intelligent Ranking.

```
docker build --tag=<Docker image tag name of OpenSearch with Intelligent Ranking
 plugin>
```

Interagindo com o plug-in de pesquisa inteligente

Depois de configurar OpenSearch (autogerenciado) com o plug-in Intelligent Ranking, você pode interagir com o plug-in usando comandos curl ou bibliotecas de OpenSearch clientes. As credenciais padrão para acesso OpenSearch com o plug-in Intelligent Ranking são o nome de usuário 'admin' e a senha 'admin'.

Para aplicar as configurações do plug-in Intelligent Ranking a um 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)
```

Você deve incluir o nome do campo de texto principal que deseja usar para se reclassificar, como o corpo do documento ou o campo de conteúdo do documento. Você também pode incluir outros campos de texto, como título do documento ou resumo do documento.

Agora você pode emitir qualquer consulta e os resultados são classificados usando o plug-in do 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 remover as configurações do plug-in Intelligent Ranking de um 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 testar o plug-in do Intelligent Ranking em uma determinada consulta ou para testar em determinados campos de corpo e 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)
```

Comparando OpenSearch resultados com Amazon Kendra resultados

Você pode comparar os resultados classificados side-by-side OpenSearch (autogerenciados) com Amazon Kendra os resultados reclassificados. OpenSearch A versão 2.4.0 e posterior do Dashboards oferece side-by-side resultados para que você possa comparar como OpenSearch classifica os documentos com a forma como Amazon Kendra o plug-in classifica os documentos para uma consulta de pesquisa.

Antes de comparar os resultados OpenSearch classificados com os resultados Amazon Kendra reclassificados, certifique-se de que seus OpenSearch painéis sejam apoiados por um OpenSearch servidor com o plug-in Intelligent Ranking. Você pode configurar isso usando o Docker e um script Bash de início rápido. Consulte Configurando o plug-in de pesquisa inteligente.

A seguir, descrevemos como comparar OpenSearch e Amazon Kendra pesquisar resultados em OpenSearch painéis. Para obter mais informações, consulte a OpenSearchdocumentação.

Comparando resultados de pesquisa em OpenSearch painéis

1. Abra http://localhost:5601 e faça login nos OpenSearch painéis. As credenciais padrão são o nome de usuário e senha “admin”.
2. Selecione Relevância da pesquisa nos OpenSearch plug-ins no menu de navegação.
3. Insira o texto de pesquisa na barra de pesquisa.
4. Selecione seu índice para a Consulta 1 e insira uma OpenSearch consulta na Consulta DSL. Você pode usar a variável `%SearchText%` para se referir ao texto de pesquisa inserido na barra de pesquisa. Para ver um exemplo dessa consulta, consulte a OpenSearch documentação. Os

resultados retornados para essa consulta são os OpenSearch resultados sem o uso do plug-in Intelligent Ranking.

5. Selecione o mesmo índice para a Consulta 2 e insira a mesma OpenSearch consulta na Consulta DSL. Além disso, inclua a extensão `kendra_intelligent_ranking` e especifique a obrigatoriedade na qual `body_field` para se classificar. Você também pode especificar o campo do título, mas o campo do corpo é obrigatório. Para ver um exemplo dessa consulta, consulte a OpenSearch documentação. Os resultados retornados para essa consulta são os resultados Amazon Kendra reclassificados usando o plug-in Intelligent Ranking. O plugin classifica até 25 resultados.
6. Selecione Pesquisar para retornar e comparar os resultados.

Classificando semanticamente os resultados de um serviço de pesquisa

Amazon Kendra O Intelligent Ranking usa Amazon Kendra os recursos de pesquisa semântica para reclassificar os resultados de um serviço de pesquisa. Ele faz isso levando em consideração o contexto da consulta de pesquisa, além de todas as informações disponíveis nos documentos do serviço de pesquisa. Amazon Kendra A classificação inteligente pode melhorar a correspondência simples de palavras-chave.

A CreateRescoreExecutionPlanAPI cria um recurso de classificação Amazon Kendra inteligente usado para provisionar a API Rescore. A `Rescore` API reclassifica os resultados da pesquisa de um serviço de pesquisa, como OpenSearch (autogerenciado).

Ao chamar `CreateRescoreExecutionPlan`, você define as unidades de capacidade necessárias para reclassificar os resultados de um serviço de pesquisa. Se você não precisar de mais unidades de capacidade além do padrão de unidade única, não altere o padrão. Forneça somente um nome para seu plano de execução de repontuação. É possível configurar até 1000 unidades extras. Para obter informações sobre o que está incluído em uma única unidade de capacidade, consulte Ajustando a capacidade. Depois de provisionar o Amazon Kendra Intelligent Ranking, você é cobrado por hora com base nas unidades de capacidade definidas. Veja as informações sobre o nível gratuito e os preços.

Um ID do plano de execução de repontuação é gerado e retornado na resposta quando você chama `CreateRescoreExecutionPlan`. A API `Rescore` usa o ID do plano de execução de repontuação para reclassificar os resultados de um serviço de pesquisa usando a capacidade definida. Você

inclui o ID do plano de execução de repontuação nos arquivos de configuração do serviço de pesquisa. Por exemplo, se você usa OpenSearch (autogerenciado), inclui o ID do plano de execução de rescore em seu arquivo docker-compose.yml ou opensearch.yml — consulte Resultados de classificação inteligente (autoatendimento). OpenSearch

Um nome do recurso da Amazon (ARN) também é gerado na resposta ao chamar `CreateRescoreExecutionPlan`. Você pode usar esse ARN para criar uma política de permissões em AWS Identity and Access Management (IAM) para restringir o acesso do usuário a um ARN específico para um plano de execução de rescore específico. Para ver um exemplo de IAM política para conceder permissão para usar a `Rescore` API para um plano de execução de rescore específico, consulte Classificação Amazon Kendra inteligente para OpenSearch autogerenciamento.

Veja a seguir um exemplo de criação de um plano de execução de repontuação com unidades de capacidade definidas como 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.");
  }
}
```

Veja a seguir um exemplo de atualização de um plano de execução de repontuação para definir unidades de capacidade como 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.");
  }
}
```

Veja a seguir um exemplo de uso da 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());
  }
}
```

Histórico do documento para Amazon Kendra

- Última atualização da documentação: 27 de fevereiro de 2024

A tabela a seguir descreve mudanças importantes em cada versão do Amazon Kendra. Para receber notificações sobre atualizações dessa documentação, é possível assinar o feed RSS.

Alteração	Descrição	Data
Novo recurso	Amazon Kendra agora oferece suporte a uma versão atualizada do conector da fonte de GitHub dados. Para obter mais informações, consulte GitHub.	27 de fevereiro de 2024
Novo recurso	Amazon Kendra agora oferece suporte a versões atualizad as do conector da fonte de Amazon FSx dados. Para obter mais informaçõ es, consulte Amazon FSx (Windows) e Amazon FSx (NetAppONTAP).	8 de fevereiro de 2024
Novo recurso	Amazon Kendra agora é compatível com uma versão atualizada do conector da fonte de dados do Slack. Para obter mais informações, consulte Slack.	11 de janeiro de 2024
Novo recurso	Amazon Kendra agora suporta o colapso e a expansão de seus resultados de pesquisa. Para obter mais informações,	19 de outubro de 2023

	consulte Recolher/expandir os resultados da pesquisa.	
Novo recurso	Amazon Kendra agora oferece suporte a um Aurora conector de fonte de dados (MySQL). Para obter mais informações, consulte Aurora (MySQL).	28 de setembro de 2023
Novo recurso	Amazon Kendra agora oferece suporte a um Aurora conector de fonte de dados (PostgreSQL). Para obter mais informações, consulte Aurora (PostgreSQL).	28 de setembro de 2023
Novo recurso	Amazon Kendra agora oferece suporte a um Amazon RDS conector de fonte de dados (MySQL). Para obter mais informações, consulte Amazon RDS (MySQL).	28 de setembro de 2023
Novo recurso	Amazon Kendra agora oferece suporte a um conector de fonte de dados Amazon RDS (Microsoft SQL Server). Para obter mais informaçõ es, consulte Amazon RDS (Microsoft SQL Server).	28 de setembro de 2023
Novo recurso	Amazon Kendra agora oferece suporte a um conector de fonte de dados Amazon RDS (Oracle). Para ter mais informações, consulte Amazon RDS (Oracle).	28 de setembro de 2023

Novo recurso	Amazon Kendra agora oferece suporte a um Amazon RDS conector de fonte de dados (PostgreSQL). Para obter mais informações, consulte Amazon RDS (PostgreSQL).	28 de setembro de 2023
Novo recurso	Amazon Kendra agora suporta um conector de fonte de dados IBM DB2. Para obter mais informações, consulte IBM DB2.	28 de setembro de 2023
Novo recurso	Amazon Kendra agora oferece suporte a um conector de fonte de dados do Microsoft SQL Server. Para obter mais informações, consulte Microsoft SQL Server.	28 de setembro de 2023
Novo recurso	Amazon Kendra agora oferece suporte a um conector de fonte de dados MySQL. Para obter mais informações, consulte (MySQL).	28 de setembro de 2023
Novo recurso	Amazon Kendra agora oferece suporte a um conector de fonte de dados do Oracle Database. Para obter mais informações, consulte Oracle Database.	28 de setembro de 2023

Novo recurso	Amazon Kendra agora oferece suporte a um conector de fonte de dados PostgreSQL. Para obter mais informações, consulte PostgreSQL.	28 de setembro de 2023
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Drupal. Para obter mais informações, consulte Drupal.	6 de setembro de 2023
Novo recurso	Recupere passagens semanticamente relevantes usando a API do Amazon Kendra Recuperar para sistemas de geração aumentada de recuperação (RAG).	22 de junho de 2023
Novo recurso	Amazon Kendra agora oferece suporte a uma versão atualizada do conector de fonte de dados do Amazon Kendra Web Crawler. Para obter mais informações, consulte Amazon Kendra Web Crawler v2.0.	21 de junho de 2023
Expansão da região	Amazon Kendra agora está disponível na Europa (Londres) (eu-west-2).	5 de junho de 2023

Novo recurso	Amazon Kendra agora oferece suporte a uma versão atualizada do conector de fonte de dados Alfresco. Para obter mais informações, consulte Alfresco.	16 de maio de 2023
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Adobe Experience Manager. Para obter mais informações, consulte Adobe Experience Manager.	11 de maio de 2023
Novo recurso	Amazon Kendra agora suporta a configuração de campos/atributos do documento quando você liga. GetQuerySuggestions Agora você pode basear as sugestões de consulta no conteúdo dos campos do documento. Para obter mais informações, consulte Sugestões de consulta.	2 de maio de 2023
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Gmail. Para obter mais informações, consulte Gmail.	13 de abril de 2023

Novo recurso	Amazon Kendra agora oferece suporte a uma versão atualizada do conector de fonte OneDrive de dados da Microsoft. Para obter mais informações, consulte Microsoft OneDrive v2.0.	3 de abril de 2023
Novo recurso	Melhore a visibilidade de novos documentos ou promova determinados documentos quando os usuários digitarem determina das consultas usando os Resultados em destaque.	30 de março de 2023
Novo recurso	Amazon Kendra agora oferece suporte a um conector de fonte de dados atualizado para a Microsoft SharePoint. Para obter mais informações, consulte Microsoft SharePoint.	2 de março de 2023
Novo recurso	Amazon Kendra agora oferece suporte a uma versão atualizada do conector de fonte de dados do Confluence. Para obter mais informações, consulte Confluence.	1 de março de 2023
Expansão da região	Amazon Kendra agora está disponível na Ásia-Pacífico (Tóquio) (ap-northeast-1).	7 de fevereiro de 2023

Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Microsoft Exchange. Para obter mais informações, consulte Microsoft Exchange.	12 de janeiro de 2023
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Microsoft Yammer. Para obter mais informações, consulte Microsoft Yammer.	12 de janeiro de 2023
Novo recurso	Amazon Kendra agora oferece suporte à indexação dos tipos de documentos RTF, XML, XSLT, MS_EXCEL, CSV, JSON e MD. Para obter mais informações, consulte Tipos de documentos.	11 de janeiro de 2023
Novo recurso	Amazon Kendra agora oferece suporte a uma versão atualizada do conector da fonte de Amazon S3 dados. Para ter mais informações, consulte Amazon S3.	10 de janeiro de 2023
Novo recurso	OpenSearchOs resultados de pesquisa (autogerenciados) podem ser classificados semanticamente usando o Amazon Kendra Intelligent Ranking.	9 de janeiro de 2023

Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Microsoft Teams. Para obter mais informações, consulte Microsoft Teams.	5 de janeiro de 2023
Novo recurso	Amazon Kendra tem um conector de fonte de dados atualizado para o Google Drive. Para obter mais informações, consulte Google Drive .	5 de janeiro de 2023
Novo recurso	Amazon Kendra tem um conector de fonte de dados atualizado para ServiceNow. Para obter mais informações, consulte ServiceNow.	21 de dezembro de 2022
Novo recurso	Amazon Kendra tem um conector de fonte de dados atualizado para o Salesforce. Para obter mais informações, consulte Salesforce.	21 de dezembro de 2022
Expansão da região	Amazon Kendra agora está disponível na Ásia-Pacífico (Mumbai) (ap-south-1).	14 de dezembro de 2022
Novo recurso	O recurso de pesquisa tabular do Amazon Kendra pode pesquisar e extrair respostas de tabelas incorporadas em documentos HTML.	27 de novembro de 2022

Novo recurso	Amazon Kendra suporta pesquisa semântica para um conjunto selecionado de idiomas.	27 de novembro de 2022
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Dropbox. Para obter mais informações, consulte Dropbox.	27 de setembro de 2022
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Zendesk. Para obter mais informações, consulte o Zendesk.	17 de agosto de 2022
Novo atributo	O controle de acesso em nível de documento agora pode ser reconfigurado após a indexação dos documento s. Para obter mais informaçõ es, consulte Configuração de controle de acesso.	14 de julho de 2022
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Alfresco. Para obter mais informações, consulte Alfresco.	30 de junho de 2022
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para GitHub. Para obter mais informações, consulte GitHub.	2 de junho de 2022

Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Jira. Para obter mais informações, consulte Jira.	12 de maio de 2022
Novo recurso	As facetas aninhadas dentro de uma faceta podem ser exibidas nos resultados da pesquisa. Para obter mais informações, consulte Facetas.	5 de maio de 2022
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Quip. Para obter mais informações, consulte a Quip.	19 de abril de 2022
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Box. Para obter mais informações, consulte Box.	6 de abril de 2022
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o Slack. Para obter mais informações, consulte Slack.	14 de março de 2022
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para Amazon FSx. Para ter mais informações, consulte Amazon FSx.	8 de fevereiro de 2022

AWS atualizações gerenciadas de políticas - Novas políticas	Amazon Kendra adicionou novas políticas AWS gerenciadas. Para obter mais informações, consulte Políticas gerenciadas pela AWS para o Amazon Kendra.	3 de janeiro de 2022
Novo recurso	Amazon Kendra o aplicativo de pesquisa pode ser implantado em alguns cliques sem a necessidade de nenhum código de front-end. Para obter mais informações, consulte Implantar uma aplicação de Pesquisa sem código.	1º de dezembro de 2021
Novo recurso	Os metadados e o conteúdo do documento podem ser enriquecidos durante o processo de ingestão de documentos. Para obter mais informações, consulte Personalização de metadados de documentos durante o processo de ingestão.	1º de dezembro de 2021
Novo recurso	Amazon Kendra oferece análises de pesquisa para obter informações úteis sobre seu aplicativo de pesquisa. Para obter mais informações, consulte Obter informações com a análise de pesquisa.	1º de dezembro de 2021

Expansão da região	Amazon Kendra agora está disponível em AWS GovCloud (Oeste dos EUA) (us-gov-west-1).	13 de outubro de 2021
Novo recurso	Amazon Kendra agora pode indexar documentos em vários idiomas e filtrar os resultados da pesquisa por idioma. Consulte Adicionar documentos em outros idiomas além do inglês e Pesquisar em idiomas.	7 de outubro de 2021
Novo recurso	Amazon Kendra agora se integra ao diretório do Identity Center para obter níveis de acesso de grupos e usuários para filtragem do contexto do usuário. Consulte Configuração do grupo de usuários para o IAM Identity Center.	6 de outubro de 2021
Novo tutorial	Amazon Kendra agora fornece um tutorial que mostra como criar uma solução de pesquisa enriquecida com metadados. Consulte Criação de uma solução de pesquisa inteligente.	13 de agosto de 2021
Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para Amazon WorkDocs. Para ter mais informações, consulte Amazon WorkDocs.	20 de julho de 2021

Novo recurso	Amazon Kendra agora fornece um rastreador da Web para rastrear e indexar páginas da Web. Para obter mais informações, consulte Web crawler.	17 de junho de 2021
Expansão da região	Amazon Kendra agora está disponível no Canadá (Central) (ca-central-1).	16 de junho de 2021
Expansão da região	Amazon Kendra agora está disponível no Leste dos EUA (Ohio) (us-east-2).	7 de junho de 2021
Novo recurso	Amazon Kendra agora oferece suporte a sugestões de consultas, nas quais os usuários recebem sugestões de consultas populares relevantes para suas pesquisas. Para obter mais informações, consulte Sugerir consultas de pesquisa populares.	27 de maio de 2021
AWS atualizações gerenciadas de políticas - Novas políticas	Amazon Kendra adicionou novas políticas AWS gerenciadas. Para obter mais informações, consulte Políticas gerenciadas pela AWS para o Amazon Kendra.	27 de maio de 2021
Expansão da região	Amazon Kendra agora está disponível na Ásia-Pacífico (Cingapura) (ap-southeast-1).	5 de maio de 2021

Novo recurso	Amazon Kendra agora suporta o ajuste da relevância da pesquisa na consulta, substituindo as configurações de ajuste definidas no nível do índice. Para obter mais informações, consulte Ajustar a relevância da pesquisa e Ajustar as respostas.	20 de abril de 2021
Novo recurso	Amazon Kendra agora oferece suporte à autenticação OAuth 2.0 e ao uso de ServiceNow consultas para selecionar documentos para indexação. Para obter mais informações, consulte ServiceNow.	1º de abril de 2021
Novo atributo	Amazon Kendra agora oferece suporte ao aprendizado incremental para documentos de perguntas frequentes. Para obter mais informações, consulte Enviar feedback para o aprendizado incremental.	17 de fevereiro de 2021
Novo recurso	Amazon Kendra agora suporta sinônimos de índice. Para obter mais informações, consulte Adicionar sinônimos a um índice..	10 de dezembro de 2020

Novo recurso	Amazon Kendra agora fornece um conector de banco de dados para o Google Workspace Drive. Para obter mais informações, consulte Usar uma fonte de dados do Google Workspace Drive.	8 de dezembro de 2020
Novo recurso	Amazon Kendra agora fornece uma JavaScript biblioteca que facilita o fornecimento de feedback sobre consultas Amazon Kendra. Para obter mais informações, consulte Enviar feedback.	8 de dezembro de 2020
Novo recurso	Amazon Kendra agora oferece suporte ao controle de acesso do usuário baseado em tokens. Para obter mais informações, consulte Configurar um documento de índice.	5 de novembro de 2020
Novo recurso	O conector da fonte de dados do Amazon Kendra Confluenc e agora funciona com a nuvem do Confluence. Para obter mais informações, consulte Using a Confluence data source (Uso de uma fonte de dados do Confluence).	5 de novembro de 2020
Expansão da região	Amazon Kendra agora está disponível na Ásia-Pacífico (Sydney) (ap-southeast-2).	2 de novembro de 2020

Novo recurso	Amazon Kendra agora fornece um conector de fonte de dados para o servidor Confluence. Para obter mais informações, consulte Using a Confluence data source (Uso de uma fonte de dados do Confluence).	26 de outubro de 2020
Novo recurso	Amazon Kendra agora fornece uma fonte de dados que você pode usar para gerar estatísticas para seus conectores personalizados. Para obter mais informações, consulte Uso de uma fonte dos dados personalizada.	21 de outubro de 2020
Novo recurso	Amazon Kendra agora oferece suporte a atributos personalizados para perguntas frequentes. Para obter mais informações, consulte Adicionar perguntas e respostas.	17 de setembro de 2020
Novo recurso	Amazon Kendra agora retorna pontuações de confiança para os resultados da consulta. Para obter mais informações, consulte QueryResultItem.	15 de setembro de 2020

Novo atributo	AWS CloudFormation agora suporta Amazon Kendra. Para obter mais informações, consulte a referência do tipo de Amazon Kendra recurso - AWS CloudFormation.	10 de setembro de 2020
Novo recurso	Amazon Kendra adiciona suporte para AWS PrivateLink. Para obter mais informações, consulte Endpoints de VPC do Amazon Kendra e interface (AWS PrivateLink).	7 de julho de 2020
Novo guia	Esta é a primeira versão do Guia do desenvolvedor do Amazon Kendra .	11 de maio de 2020

Referência de API

A documentação de referência da API agora é um guia separado.

Glossário do AWS

Para obter a terminologia mais recente da AWS, consulte o glossário da AWS na Referência do Glossário da AWS.

As traduções são geradas por tradução automática. Em caso de conflito entre o conteúdo da tradução e da versão original em inglês, a versão em inglês prevalecerá.