aws

개발자 안내서

Amazon Kendra

Amazon Kendra: 개발자 안내서

Table of Contents

Amazon Kendra란 무엇인가요?

Amazon Kendra는 자연어 처리 및 고급 기계 학습 알고리즘을 사용하여 데이터에서 검색 질문에 대한 구체적인 답변을 반환하는 지능형 검색 서비스입니다.

기존의 키워드 기반 검색과 달리, Amazon Kendra는 의미론적, 상황별 이해 기능을 사용하여 문서가 검색 쿼리와 관련이 있는지 여부를 결정합니다. 질문에 대한 구체적인 답변을 반환하여 사용자에게 마치 전문가와의 상호 작용에 가까운 경험을 제공합니다.

> **Note**
>
> Amazon Kendra의 시맨틱 검색 기능을 사용하여 다른 검색 서비스의 결과 순위를 다시 매길 수도 있습니다. 자세한 내용은 Amazon Kendra Intelligent Ranking을 참조하세요.

Amazon Kendra를 사용하면 여러 데이터 리포지토리를 인덱스에 연결하고 문서를 수집 및 크롤링하여 통합된 검색 환경을 만들 수 있습니다. 문서 메타데이터를 사용하여 다양한 기능을 갖춘 사용자 지정 검색 환경을 만들면 사용자가 쿼리에 대한 올바른 답변을 효율적으로 찾도록 지원할 수 있습니다.

Amazon Kendra란 무엇입니까?

Amazon Kendra 쿼리

Amazon Kendra에 다음과 같은 유형의 쿼리를 요청할 수 있습니다.

팩토이드 질문 - 누가, 무엇을, 언제, 어디에 관해 물어보는 간단한 질문(예: 시애틀에서 가장 가까운 서비스 센터는 어디입니까?) 팩토이드 질문에는 한 단어 또는 문구로 반환될 수 있는 사실에 기반한 답변이 있습니다. FAQ 또는 인덱싱된 문서에서 답변을 검색할 수 있습니다.

서술형 질문 - 문장, 구절 또는 문서 전체가 답일 수 있는 질문. 예를 들어 Echo Plus를 네트워크에 연결하려면 어떻게 해야 하나요? 또는 저소득 가정에 대한 세금 혜택을 받으려면 어떻게 해야 하나요?

키워드 및 자연어 질문 - 의미가 명확하지 않을 수 있는 복잡한 대화 내용이 포함된 질문. 기조 연설을 예로 들 수 있습니다. 문맥상 여러 의미를 지닌 “address”와 같은 단어를 Amazon Kendra가 발견하면 검색 쿼리에 숨은 의미를 정확하게 유추하여 관련 정보를 반환합니다.

Amazon Kendra의 이점

Amazon Kendra는 확장성이 뛰어나고 성능 요구 사항을 충족할 수 있으며 Amazon S3, Amazon Lex 등의 다른 AWS 서비스와 긴밀하게 통합되며 엔터프라이즈급 보안을 제공합니다. Amazon Kendra 사용 시 다음과 같은 이점이 있습니다.

단순성 - Amazon Kendra는 검색하려는 문서를 관리할 수 있는 콘솔과 API를 제공합니다. 간단한 검색 API를 사용하여 웹 사이트 또는 모바일 애플리케이션과 같은 클라이언트 애플리케이션에 Amazon Kendra를 통합할 수 있습니다.

연결성 - Amazon Kendra는 타사 데이터 리포지토리 또는 Microsoft SharePoint와 같은 데이터 소스에 연결할 수 있습니다. 데이터 소스를 사용하여 문서를 쉽게 인덱싱하고 검색할 수 있습니다.

정확성 - 키워드 검색을 사용하는 기존 검색 서비스와 달리, Amazon Kendra는 질문의 맥락을 이해하려고 시도하고 쿼리와 가장 관련성이 높은 단어, 스니펫 또는 문서를 반환합니다. Amazon Kendra는 기계 학습을 사용하여 시간이 지남에 따라 검색 결과를 개선합니다.

보안 - Amazon Kendra는 매우 안전한 엔터프라이즈 검색 환경을 제공합니다. 검색 결과는 조직의 보안 모델을 반영하며 문서에 대한 사용자 또는 그룹 액세스를 기준으로 필터링할 수 있습니다. 고객은 사용자 액세스를 인증하고 권한을 부여할 책임이 있습니다.

Amazon Kendra Editions

Amazon Kendra에는 Developer Edition과 Enterprise Edition, 두 가지 버전이 있습니다. 다음 표에 두 옵션의 특징과 차이점이 요약되어 있습니다.

Amazon Kendra Developer Edition	Amazon Kendra Enterprise Edition
Amazon Kendra Developer Edition은 저렴한 비용으로 Amazon Kendra의 모든 기능을 제공합니다.	Amazon Kendra Enterprise Edition은 Amazon Kendra의 모든 기능을 제공하며 프로덕션 상황에 맞게 설계되었습니다.
이상적인 사용 사례	이상적인 사용 사례
• Amazon Kendra가 문서를 인덱싱하는 방법 살펴보기 • 기능 시도 • Amazon Kendra를 사용하는 애플리케이션 개발	• 전체 엔터프라이즈 문서 라이브러리를 인덱싱합니다. • 프로덕션 환경에 애플리케이션 배포

Amazon Kendra Developer Edition	Amazon Kendra Enterprise Edition
기능 • 750시간 사용이 포함된 프리 티어 • 각각 최대 5개의 데이터 소스를 포함하는 최대 5개의 인덱스 • 문서 10,000개 또는 추출된 텍스트 3GB • 일일 약 4,000쿼리 또는 초당 0.05쿼리 • 1개의 가용 영역(AZ)에서 실행 - 가용 영역(AWS 리전 내 데이터 센터) 참조 제한 사항 • 프로덕션 애플리케이션에는 적합하지 않습니다. • 지연 시간이나 가용성에 대한 보장은 없습니다.	기능 • 각각 최대 50개의 데이터 소스를 포함하는 최대 5개의 인덱스 • 문서 100,000개 또는 추출된 텍스트 30GB • 일일 약 8,000쿼리 또는 초당 0.1쿼리 • 3개의 가용 영역(AZ)에서 실행 - 가용 영역(AWS 리전 내 데이터 센터) 참조 **Note** Service Quotas 콘솔을 사용하여 할당량을 높일 수 있습니다. 제한 사항 • 없음

Note

Amazon Kendra 지원 리전의 목록, 엔드포인트 및 서비스 할당량은 Amazon Kendra 엔드포인트 및 할당량을 참조하세요.

Amazon Kendra 요금

처음 30일 동안 최대 750시간을 사용할 수 있는 Amazon Kendra Developer Edition으로 무료로 시작할 수 있습니다.

평가판이 만료된 후에는 프로비저닝된 모든 Amazon Kendra 인덱스에 대한 요금이 부과되며, 인덱스가 비어 있고 쿼리가 실행되지 않은 경우에도 마찬가지입니다. 평가판이 만료된 후에는 Amazon Kendra 데이터 소스를 사용하여 문서를 스캔하고 동기화하는 데 추가 요금이 부과됩니다.

전체적인 요금 및 가격 목록은 Amazon Kendra 요금을 참조하세요.

Amazon Kendra를 처음 사용하십니까?

Amazon Kendra를 처음 사용한다면, 다음 단원을 순서대로 읽어보기를 권장합니다.

1	2	3	4	5	6
Amazon Kendra 작동 방식	시작하기	인덱스 생성	일괄 업로드를 사용하여 인덱스에 직접 문서 추가	데이터 소스 커넥터 생성	인덱스 검색
Amazon Kendra 구성 요소를 소개하고 구성 요소를 사용하여 검색 솔루션을 만드는 방법을 설명합니다.	계정을 설정하고 Amazon Kendra 검색 API를 테스트하는 방법을 설명합니다.	Amazon Kendra를 사용하여 검색 인덱스를 만들고 데이터 소스를 추가하여 문서를 동기화하는 방법을 설명합니다.	Amazon Kendra 인덱스에 직접 문서를 추가하는 방법을 설명합니다.	데이터 리포지토리의 문서를 Amazon Kendra 인덱스에 추가하는 방법을 설명합니다.	Amazon Kendra 검색 API를 사용하여 인덱스를 검색하는 방법을 설명합니다.

Amazon Kendra 작동 방식

Amazon Kendra 애플리케이션에 검색 기능을 제공합니다. 문서를 직접 또는 타사 문서 리포지토리에서 인덱싱하고 관련 정보를 사용자에게 지능적으로 제공합니다. 를 Amazon Kendra 사용하여 다양한 유형의 문서에 대해 업데이트 가능한 색인을 만들 수 있습니다. 에서 지원되는 문서 유형 목록은 문서 유형을 Amazon Kendra 참조하십시오.

Amazon Kendra 다른 서비스와 통합됩니다. 예를 들어 Amazon Lex 채팅 봇에 Amazon Kendra 검색 기능을 강화하여 사용자의 질문에 유용한 답변을 제공할 수 있습니다. Amazon Simple Storage Service 버킷을 데이터 소스로 사용하여 문서에 연결하고 문서를 Amazon Kendra 인덱싱할 수 있습니다. 또한 AWS Identity and Access Management를 사용하여 리소스에 대한 액세스 정책 또는 권한을 설정할 수 있습니다.

Amazon Kendra 에는 다음과 같은 구성 요소가 있습니다.

- 문서를 보관하고 검색할 수 있게 해주는 인덱스입니다.
- 문서를 저장하고 Amazon Kendra 가 연결하는 데이터 소스입니다. 데이터 원본을 색인과 자동으로 동기화하여 Amazon Kendra 색인이 원본 리포지토리와 함께 업데이트된 상태로 유지되도록 할 수 있습니다.
- 문서를 인덱스에 직접 추가하는 문서 추가 API입니다.

콘솔 또는 API를 Amazon Kendra 통해 사용할 수 있습니다. 인덱스를 생성, 업데이트 및 삭제할 수 있습니다. 색인을 삭제하면 해당 데이터 소스 커넥터가 모두 삭제되고 모든 문서 정보가 영구적으로 삭제됩니다. Amazon Kendra

주제

- 인덱스
- 문서
- 데이터 소스
- 쿼리
- Tags

인덱스

인덱스는 문서의 내용을 담고 있으며 문서를 검색할 수 있는 방식으로 구조화되어 있습니다. 인덱스에 문서를 추가하는 방법은 문서를 저장하는 방식에 따라 달라집니다.

- 문서를 Amazon S3 버킷이나 Microsoft SharePoint 사이트와 같은 일종의 리포지토리에 저장하는 경우 데이터 소스 커넥터를 사용하여 리포지토리에서 문서를 인덱싱합니다.
- 문서를 리포지토리에 저장하지 않는 경우 BatchPutDocumentAPI를 사용하여 문서를 직접 인덱싱합니다.
- Amazon Kendra (Amazon S3) 버킷에 저장해야 하는 FAQ 질문 및 답변의 경우 버킷에서 업로드합니다.

Amazon Kendra 콘솔 AWS CLI, 또는 AWS SDK를 사용하여 색인을 생성할 수 있습니다. 인덱싱할 수 있는 문서 유형에 대한 자세한 내용은 문서 유형을 참조하세요.

Amazon Kendra 예약된 문서 필드 또는 일반 문서 필드 사용

UpdateIndex API를 사용하면 예약된 인덱스 필드 이름을 `DocumentMetadataConfigurationUpdates` 사용하고 지정하여 해당 문서 속성/필드 이름에 매핑되는 Amazon Kendra 예약 또는 공통 필드를 만들 수 있습니다. 사용자 지정 필드도 생성할 수 있습니다. 데이터 소스 커넥터를 사용하는 경우 대부분의 커넥터에는 데이터 소스 문서 필드를 인덱스 필드에 매핑하는 필드 매핑이 포함됩니다. Amazon Kendra 콘솔을 사용하는 경우 데이터 소스를 선택하고 편집 작업을 선택한 다음 데이터 소스 구성을 위한 필드 매핑 섹션 옆으로 이동하여 필드를 업데이트합니다.

필드를 displayable, facetable, searchable, sortable로 설정하도록 `Search` 객체를 구성할 수 있습니다. 필드의 순위 순서, 부스트 기간 또는 부스팅에 적용할 기간, 최신성, 중요도 값 및 특정 필드 값에 매핑된 중요도 값을 설정하도록 `Relevance` 객체를 구성할 수 있습니다. 콘솔을 사용하는 경우 탐색 메뉴에서 패싯 옵션을 선택하여 필드에 대한 검색 설정을 지정할 수 있습니다. 관련성 조정을 설정하려면 탐색 메뉴에서 인덱스를 검색하는 옵션을 선택하고 쿼리를 입력한 다음 사이드 패널 옵션을 사용하여 검색 관련성을 조정합니다. 필드를 생성한 후에는 필드 유형을 변경할 수 없습니다.

Amazon Kendra 사용할 수 있는 다음과 같은 예약된 문서 필드 또는 공통 문서 필드가 있습니다.

- `_authors` - 문서 내용을 책임지는 한 명 이상의 작성자 목록.
- `_category` - 문서를 특정 그룹에 배치하는 범주.

- `_created_at` - 문서가 생성된 ISO 8601 형식의 날짜 및 시간. 예를 들어, 2012-03-25T12:30:10+01:00은 중부 유럽 시간으로 2012년 3월 25일 오후 12시 30분 10초에 대한 ISO 8601 날짜/시간 형식입니다.
- `_data_source_id` - 문서가 포함된 데이터 소스의 식별자.
- `_document_body` - 작업 문서의 내용.
- `_document_id` - 문서의 고유 식별자.
- `_document_title` - 문서의 제목.
- `_excerpt_page_number` - 문서 발췌문이 나타나는 PDF 파일의 페이지 번호. 2020년 9월 8일 이전에 인덱스를 만든 경우 이 속성을 사용하려면 먼저 문서를 다시 인덱싱해야 합니다.
- `_faq_id` - 질문-답변 유형 문서(FAQ)인 경우 FAQ의 고유 식별자.
- `_file_type` - 문서의 파일 형식(예: pdf 또는 doc).
- `_last_updated_at` - 문서가 마지막으로 업데이트된 ISO 8601 형식의 날짜 및 시간. 예를 들어, 2012-03-25T12:30:10+01:00은 중부 유럽 시간으로 2012년 3월 25일 오후 12시 30분 10초에 대한 ISO 8601 날짜/시간 형식입니다.
- `_source_uri` - 문서가 제공된 URI. 예를 들면, 회사 웹 사이트에 있는 문서의 URI.
- `_version` - 문서의 특정 버전을 나타내는 식별자.
- `_view_count` - 문서가 조회된 횟수.
- `_language_code`(문자열) - 문서에 적용되는 언어의 코드. 언어를 지정하지 않으면 영어가 기본값으로 사용됩니다. 코드를 포함하여 지원되는 언어에 대한 자세한 내용은 영어 이외의 언어로 문서 추가를 참조하세요.

사용자 지정 필드의 경우 예약된 필드 또는 공통 필드를 만들 때와 마찬가지로 `UpdateIndex` API와 `DocumentMetadataConfigurationUpdates`를 사용하여 이러한 필드를 만듭니다. 사용자 지정 필드에 적절한 데이터 유형을 설정해야 합니다. 콘솔을 사용하는 경우 데이터 소스를 선택하고 편집 작업을 선택한 다음 데이터 소스 구성을 위한 필드 매핑 섹션 옆으로 이동하여 필드를 업데이트합니다. 일부 데이터 소스는 새 필드 또는 사용자 지정 필드 추가를 지원하지 않습니다. 필드를 생성한 후에는 필드 유형을 변경할 수 없습니다.

사용자 지정 필드에 설정할 수 있는 유형은 다음과 같습니다.

- 날짜
- 숫자
- String

- 문자열 목록

BatchPutDocumentAPI를 사용하여 색인에 문서를 추가한 경우 문서의 필드/속성을 `Attributes` 나열하고 객체를 사용하여 필드를 생성합니다. `DocumentAttribute`

Amazon S3 데이터 소스에서 색인된 문서의 경우 필드 정보가 포함된 JSON 메타데이터 파일을 사용하여 필드를 만듭니다.

지원되는 데이터베이스를 데이터 소스로 사용하는 경우 필드 매핑 옵션을 사용하여 필드를 구성할 수 있습니다.

인덱스 검색

인덱스를 생성한 후 문서 검색을 시작할 수 있습니다. 자세한 내용은 인덱스 선택을 참조하세요.

문서

이 섹션에서는 지원하는 다양한 문서 형식과 문서의 다양한 필드/속성을 Amazon Kendra 인덱싱하는 방법을 설명합니다.

주제

- 문서 유형 또는 형식
- 문서 속성 또는 필드

문서 유형 또는 형식

Amazon Kendra PDF, HTML, PowerPoint Word 등과 같이 널리 사용되는 문서 유형 또는 형식을 지원합니다. 인덱스는 여러 문서 형식을 포함할 수 있습니다.

Amazon Kendra 문서를 검색할 수 있도록 문서 내의 내용을 추출합니다. 추출된 텍스트와 문서 내의 표 형식 콘텐츠(HTML 테이블)에 대한 검색을 최적화하는 방식으로 문서가 파싱됩니다. 즉, 검색에 사용되는 필드 또는 속성으로 문서를 구조화합니다. 마지막 수정 날짜와 같은 문서 메타데이터는 검색에 유용한 필드가 될 수 있습니다.

문서를 행과 열로 구성할 수 있습니다. 예를 들어, 각 문서는 행이고 각 문서 필드/속성(예: 제목 및 본문 내용)은 열입니다. 예를 들어 데이터베이스를 데이터 소스로 사용하는 경우 데이터는 행과 열로 구조화되거나 구성되어야 합니다.

다음과 같은 방법으로 인덱스에 문서를 추가할 수 있습니다.

- BatchPutDocument API
- 데이터 소스 커넥터

FAQ 파일을 추가하려면 CreateFaqAPI를 사용하여 버킷에 저장된 파일을 추가합니다. Amazon S3 기본 CSV 형식, 헤더에 사용자 지정 필드/속성을 포함하는 CSV 형식, 사용자 지정 필드가 포함된 JSON 형식 중에서 선택할 수 있습니다. 기본 형식은 기본 CSV입니다.

다음은 지원되는 각 문서 형식에 대한 정보와 Amazon Kendra 가 문서를 인덱싱할 때 각 형식을 처리하는 방법에 대한 정보를 제공합니다.

문서 형식	취급 방식	문서 처리 방법	원래 구조
휴대용 문서형식(PDF)	HTML	HTML로 변환된 후 내용이 추출됩니다.	비정형
HyperText 마크업 언어 (HTML)	HTML	HTML 태그는 콘텐츠를 추출하기 위해 필터링됩니다. 콘텐츠는 기본 HTML 시작 태그와 종료 태그 (<HTML>content</HTML>) 사이에 있어야 합니다.	반구조화
확장형 마크업 언어 (XML)	XML	XML 태그는 콘텐츠를 추출하기 위해 필터링됩니다.	반구조화
확장형 스타일시트 언어 변환(XSLT)	XSLT	태그는 콘텐츠를 추출하기 위해 필터링됩니다.	반구조화
MarkDown (MD)	일반 텍스트	내용은 MarkDown 구문이 포함된 상태로 추출됩니다.	반구조화
쉼표로 구분된 값 (CSV)	CSV	각 셀에서 추출된 콘텐츠로, 단일 파일은 단	FAQ 파일은 구조화되고 그 외에는 반정형

문서 형식	취급 방식	문서 처리 방법	원래 구조
		일 문서 결과로 처리됩니다.	
Microsoft Excel(XLS 및 XLSX)	XLS 및 XLSX	각 셀에서 추출된 콘텐츠로, 단일 파일은 단일 문서 결과로 처리됩니다.	반구조화
JavaScript 객체 표기법 (JSON)	일반 텍스트	콘텐츠는 JSON 구문이 포함된 상태로 추출됩니다.	반구조화
서식 있는 텍스트 (RTF)	RTF	RTF 구문은 필터링되어 콘텐츠를 추출합니다.	반구조화
마이크로소프트 PowerPoint (PPT)	PPT	PowerPoint 슬라이드에서는 텍스트 내용만 추출하여 검색합니다. 이미지 및 기타 콘텐츠는 추출되지 않습니다.	비정형
Microsoft Word(DOCX)	DOCX	Word 페이지에서는 검색 대상 텍스트 내용만 추출됩니다. 이미지 및 기타 콘텐츠는 추출되지 않습니다.	비정형
일반 텍스트(TXT)	TXT	텍스트 문서의 모든 텍스트가 추출됩니다.	비정형

문서 속성 또는 필드

문서에는 관련 속성 또는 필드가 있습니다. 문서의 필드는 문서의 속성 또는 문서 구조 내에 포함된 내용입니다. 예를 들어, 각 문서에는 제목, 본문 텍스트, 작성자가 포함될 수 있습니다. 특정 문서에 대한

사용자 지정 필드를 추가할 수도 있습니다. 예를 들어 인덱스가 세금 문서를 검색하는 경우 세금 문서 유형에 대한 사용자 지정 필드를 지정할 수 있습니다(예: W-2, 1099 등).

쿼리에서 문서 필드를 사용하려면 먼저 인덱스 필드에 매핑되어야 합니다. 예를 들어 제목 필드를 `_document_title` 필드에 매핑될 수 있습니다. 자세한 내용을 알아보려면 필드 매핑을 참조하세요. 새 필드를 추가하려면 필드를 매핑할 인덱스 필드를 만들어야 합니다. 콘솔이나 UpdateIndexAPI를 사용하여 인덱스 필드를 생성합니다.

문서 필드를 사용하여 응답을 필터링하고 패싯된 검색 결과를 만들 수 있습니다. 예를 들어 특정 버전의 문서만 반환하도록 응답을 필터링하거나 검색어와 일치하는 1099 유형의 세금 문서만 반환하도록 검색을 필터링할 수 있습니다. 자세한 내용은 필터링 및 패싯 검색을 참조하세요.

문서 필드를 사용하여 쿼리 응답을 수동으로 조정할 수도 있습니다. 예를 들어, 응답 시 반환할 문서를 결정할 때 제목 필드의 중요도를 높여 필드에 Amazon Kendra 할당되는 가중치를 높일 수 있습니다. 자세한 내용은 검색 관련성 조정을 참조하세요.

색인에 직접 문서를 추가하는 경우 문서 입력 매개변수의 필드를 BatchPutDocumentAPI에 지정합니다. DocumentAttribute개체 배열에 사용자 지정 필드 값을 지정합니다. 데이터 소스를 사용하는 경우 문서 필드를 추가하는 데 사용하는 방법은 데이터 소스에 따라 달라집니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

Amazon Kendra 예약된 문서 필드 또는 공용 문서 필드 사용

UpdateIndex API를 사용하면 예약된 인덱스 필드 이름을 `DocumentMetadataConfigurationUpdates` 사용하고 지정하여 해당 문서 속성/필드 이름에 매핑되는 Amazon Kendra 예약 또는 공통 필드를 만들 수 있습니다. 사용자 지정 필드도 생성할 수 있습니다. 데이터 소스 커넥터를 사용하는 경우 대부분의 커넥터에는 데이터 소스 문서 필드를 인덱스 필드에 매핑하는 필드 매핑이 포함됩니다. Amazon Kendra 콘솔을 사용하는 경우 데이터 소스를 선택하고 편집 작업을 선택한 다음 데이터 소스 구성을 위한 필드 매핑 섹션 옆으로 이동하여 필드를 업데이트합니다.

필드를 displayable, facetable, searchable, sortable로 설정하도록 `Search` 객체를 구성할 수 있습니다. 필드의 순위 순서, 부스트 기간 또는 부스팅에 적용할 기간, 최신성, 중요도 값 및 특정 필드 값에 매핑된 중요도 값을 설정하도록 `Relevance` 객체를 구성할 수 있습니다. 콘솔을 사용하는 경우 탐색 메뉴에서 패싯 옵션을 선택하여 필드에 대한 검색 설정을 지정할 수 있습니다. 관련성 조정을 설정하려면 탐색 메뉴에서 인덱스를 검색하는 옵션을 선택하고 쿼리를 입력한 다음 사이드 패널 옵션을 사용하여 검색 관련성을 조정합니다. 필드를 생성한 후에는 필드 유형을 변경할 수 없습니다.

Amazon Kendra 사용할 수 있는 다음과 같은 예약된 문서 필드 또는 공통 문서 필드가 있습니다.

- _authors - 문서 내용을 책임지는 한 명 이상의 작성자 목록.
- _category - 문서를 특정 그룹에 배치하는 범주.
- _created_at - 문서가 생성된 ISO 8601 형식의 날짜 및 시간. 예를 들어, 2012-03-25T12:30:10+01:00은 중부 유럽 시간으로 2012년 3월 25일 오후 12시 30분 10초에 대한 ISO 8601 날짜/시간 형식입니다.
- _data_source_id - 문서가 포함된 데이터 소스의 식별자.
- _document_body - 작업 문서의 내용.
- _document_id - 문서의 고유 식별자.
- _document_title - 문서의 제목.
- _excerpt_page_number - 문서 발췌문이 나타나는 PDF 파일의 페이지 번호. 2020년 9월 8일 이전에 인덱스를 만든 경우 이 속성을 사용하려면 먼저 문서를 다시 인덱싱해야 합니다.
- _faq_id - 질문-답변 유형 문서(FAQ)인 경우 FAQ의 고유 식별자.
- _file_type - 문서의 파일 형식(예: pdf 또는 doc).
- _last_updated_at - 문서가 마지막으로 업데이트된 ISO 8601 형식의 날짜 및 시간. 예를 들어, 2012-03-25T12:30:10+01:00은 중부 유럽 시간으로 2012년 3월 25일 오후 12시 30분 10초에 대한 ISO 8601 날짜/시간 형식입니다.
- _source_uri - 문서가 제공된 URI. 예를 들면, 회사 웹 사이트에 있는 문서의 URI.
- _version - 문서의 특정 버전을 나타내는 식별자.
- _view_count - 문서가 조회된 횟수.
- _language_code(문자열) - 문서에 적용되는 언어의 코드. 언어를 지정하지 않으면 영어가 기본값으로 사용됩니다. 코드를 포함하여 지원되는 언어에 대한 자세한 내용은 영어 이외의 언어로 문서 추가를 참조하세요.

사용자 지정 필드의 경우 예약된 필드 또는 공통 필드를 만들 때와 마찬가지로 UpdateIndex API와 DocumentMetadataConfigurationUpdates를 사용하여 이러한 필드를 만듭니다. 사용자 지정 필드에 적절한 데이터 유형을 설정해야 합니다. 콘솔을 사용하는 경우 데이터 소스를 선택하고 편집 작업을 선택한 다음 데이터 소스 구성을 위한 필드 매핑 섹션 옆으로 이동하여 필드를 업데이트합니다. 일부 데이터 소스는 새 필드 또는 사용자 지정 필드 추가를 지원하지 않습니다. 필드를 생성한 후에는 필드 유형을 변경할 수 없습니다.

사용자 지정 필드에 설정할 수 있는 유형은 다음과 같습니다.

- 날짜

- 숫자
- String
- 문자열 목록

BatchPutDocumentAPI를 사용하여 색인에 문서를 추가한 경우 문서의 필드/속성을 `Attributes` 나열하고 객체를 사용하여 필드를 생성합니다. `DocumentAttribute`

Amazon S3 데이터 소스에서 색인된 문서의 경우 필드 정보가 포함된 JSON 메타데이터 파일을 사용하여 필드를 만듭니다.

지원되는 데이터베이스를 데이터 소스로 사용하는 경우 필드 매핑 옵션을 사용하여 필드를 구성할 수 있습니다.

데이터 소스

데이터 소스는 문서 또는 콘텐츠에 Amazon Kendra 연결하고 색인을 생성하는 데이터 리포지토리 또는 위치입니다. 예를 들어 SharePoint Microsoft에 연결하여 이 원본에 저장된 문서를 크롤링하고 Amazon Kendra 인덱싱하도록 구성할 수 있습니다. 크롤링할 URL을 제공하여 웹 페이지를 Amazon Kendra 인덱싱할 수도 있습니다. 데이터 원본을 Amazon Kendra 색인과 자동으로 동기화하여 데이터 원본에서 추가, 업데이트 또는 삭제된 문서도 색인에서 추가, 업데이트 또는 삭제되도록 할 수 있습니다.

지원되는 데이터 소스:

- Adobe Experience Manager
- Alfresco
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (Windows)
- Amazon FSx (NetApp ONTAP)
- 데이터베이스 데이터 소스
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (오라클)
- Amazon RDS (PostgreSQL)

- Amazon S3 버킷
- Amazon Kendra 웹 크롤러
- Amazon WorkDocs
- Box
- Confluence
- 사용자 지정 데이터 소스
- Dropbox
- Drupal
- GitHub
- Gmail
- Google Workspace Drives
- IBM DB2
- Jira
- Microsoft Exchange
- 마이크로소프트 OneDrive
- 마이크로소프트 SharePoint
- Microsoft Teams
- Microsoft SQL Server
- Microsoft Yammer
- MySQL
- Oracle Database
- PostgreSQL
- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

에서 지원하는 문서 유형 또는 형식 목록은 문서 유형을 Amazon Kendra 참조하십시오. 데이터 소스에서 문서를 인덱싱할 데이터 소스 커넥터를 만들기 전에 먼저 인덱스를 만들어야 합니다.

> **Note**
>
> 문서의 인덱스를 만들려면 데이터 소스를 사용할 필요가 없습니다. 배치 업로드를 사용하여 인덱스에 직접 문서를 추가할 수 있습니다. 자세한 내용은 인덱스에 직접 문서 추가를 참조하세요.

Amazon Kendra 콘솔, AWS CLI 또는 SDK 사용에 대한 자세한 내용은 시작하기를 참조하십시오.

쿼리

답을 얻기 위해 사용자는 인덱스를 쿼리합니다. 사용자는 쿼리에 자연어를 사용할 수 있습니다. 응답에는 제목, 텍스트 발췌문, 인덱스에서 가장 좋은 답변을 제공하는 문서의 위치 등의 정보가 포함됩니다.

Amazon Kendra 문서의 내용뿐만 아니라 문서에 대해 제공하는 모든 정보를 사용하여 문서가 쿼리와 관련이 있는지 여부를 판단합니다. 예를 들어 색인에 문서가 마지막으로 업데이트된 시기에 대한 정보가 포함되어 있는 경우 최근에 업데이트된 문서에 더 높은 관련성을 Amazon Kendra 할당하도록 지시할 수 있습니다.

쿼리에는 필터 기준을 충족하는 문서만 Amazon Kendra 반환하도록 응답을 필터링하는 방법에 대한 기준도 포함될 수 있습니다. 예를 들어 부서라는 인덱스 필드를 만든 경우 부서 필드가 합법으로 설정된 문서만 반환되도록 응답을 필터링할 수 있습니다. 자세한 내용은 검색 필터링을 참조하세요.

인덱스에 있는 개별 필드의 관련성을 조정하여 쿼리 결과에 영향을 줄 수 있습니다. 조정은 결과에서 필드의 중요도를 변경합니다. 예를 들어, 새로운 범주 문서의 중요도를 높이면 이 범주의 문서가 응답에 포함될 가능성이 높아집니다. 자세한 내용은 검색 관련성 조정을 참조하세요.

쿼리 사용에 대한 자세한 내용은 인덱스 검색을 참조하세요.

Tags

태그 또는 레이블을 할당하여 인덱스, 데이터 소스 및 FAQ를 관리합니다. 태그를 사용하여 Amazon Kendra 리소스를 다양한 방식으로 분류할 수 있습니다. 예를 들어 용도별, 소유자별 또는 애플리케이션별 또는 몇 가지의 조합. 각 태그는 사용자가 정의하는 키와 값으로 구성됩니다.

태그를 통해 다음 작업을 수행할 수 있습니다.

- AWS 리소스를 식별하고 구성하세요. 많은 AWS 서비스가 태그 지정을 지원하므로 서로 다른 서비스의 리소스에 동일한 태그를 할당하여 리소스가 관련되어 있음을 나타낼 수 있습니다. 예를 들어 인덱스와 해당 인덱스를 사용하는 Amazon Lex 봇에 동일한 태그를 지정할 수 있습니다.

- 비용을 할당합니다. AWS Billing and Cost Management 대시보드에서 태그를 활성화합니다. AWS 태그를 사용하여 비용을 분류하고 월별 비용 할당 보고서를 제공합니다. 자세한 내용은 AWS Billing and Cost Management에 대한 비용 할당 및 태그 지정을 참조하십시오.
- 리소스에 대한 액세스를 제어합니다. AWS Identity and Access Management (IAM) 정책의 태그를 사용하여 Amazon Kendra 리소스에 대한 액세스를 제어할 수 있습니다. 이러한 정책을 IAM 역할 또는 사용자에게 연결하여 태그 기반 액세스 제어를 활성화할 수 있습니다. 자세한 내용은 태그 기반 권한 부여를 참조하세요.

AWS Management Console, AWS Command Line Interface (AWS CLI) 또는 Amazon Kendra API를 사용하여 태그를 만들고 관리할 수 있습니다.

리소스에 태그 지정

Amazon Kendra 콘솔을 사용하는 경우 리소스를 생성할 때 리소스에 태그를 지정하거나 나중에 추가할 수 있습니다. 이 콘솔을 사용하여 태그를 업데이트하거나 제거할 수도 있습니다.

AWS Command Line Interface (AWS CLI) 또는 Amazon Kendra API를 사용하는 경우 다음 작업을 사용하여 리소스의 태그를 관리하세요.

- CreateDataSource—데이터 소스를 만들 때 태그를 적용합니다.
- CreateFaq—FAQ를 생성할 때 태그를 적용합니다.
- CreateIndex—색인을 생성할 때 태그를 적용합니다.
- ListTagsForResource—리소스와 관련된 태그를 볼 수 있습니다.
- TagResource—리소스의 태그를 추가하고 수정합니다.
- UntagResource—리소스에서 태그를 제거합니다.

태그 제한

Amazon Kendra 리소스의 태그에는 다음과 같은 제한이 적용됩니다.

- 최대 태그 수 - 50
- 최대 키 길이 - 128자
- 최대 값 길이 - 256자
- 키 및 값에 사용할 수 있는 문자: a-z, A-Z, 공백 및 _ . : / = + - @ 문자
- 키와 값은 대/소문자를 구분합니다

- 키 접두사로 `aws:`를 사용하지 마세요. AWS 전용입니다.

Amazon Kendra 설정

Amazon Kendra를 사용하려면 먼저 Amazon Web Services(AWS) 계정이 있어야 합니다. AWS 계정이 있으면 Amazon Kendra 콘솔, AWS Command Line Interface () 또는 SDK를 통해 Amazon Kendra에 액세스할 수 있습니다. AWS CLI AWS

이 안내서에는 AWS CLI, Java, Python에 대한 예제가 포함되어 있습니다.

주제

가입하기 AWS

Amazon Web Services (AWS) 에 가입하면 Amazon Kendra를 AWS 포함한 모든 서비스에 계정이 자동으로 가입됩니다. 사용자에게는 사용한 서비스에 대해서만 요금이 청구됩니다.

이미 AWS 계정이 있다면 다음 작업으로 건너뛰십시오. AWS 계정이 없는 경우 다음 절차에 따라 계정을 생성합니다.

가입하려면 AWS

1. https://aws.amazon.com 를 열고 AWS 계정 만들기를 선택합니다.
2. 화면의 지침을 따라 계정 생성을 완료합니다. 12자리 AWS 계정 번호를 적어둡니다. 등록 절차 중 전화를 받고 전화 키패드를 사용하여 PIN을 입력하는 과정이 있습니다.
3. AWS Identity and Access Management (IAM) 관리자 사용자를 생성합니다. 지침은 AWS Identity and Access Management 사용 설명서의 첫 번째 IAM 사용자 및 그룹 생성을 참조하세요.

지역 및 엔드포인트

엔드포인트는 웹 서비스의 진입점인 URL입니다. 각 엔드포인트는 특정 AWS 지역과 연결됩니다. Amazon Kendra 콘솔, AWS CLI 및 Amazon Kendra SDK를 함께 사용하는 경우 해당 캠페인의 모든

Amazon Kendra 구성 요소 (인덱스, 쿼리 등) 를 동일한 지역에서 생성해야 하므로 기본 지역에 주의를 기울이십시오. 모든 Amazon Kendra 지원 리전 및 엔드포인트 정보는 리전 및 엔드포인트를 참조하세요.

설정: AWS CLI

AWS 명령줄 인터페이스 (AWS CLI) 는 Amazon Kendra를 비롯한 AWS 서비스를 관리하기 위한 통합 개발자 도구입니다. 이 도구를 설치하는 것이 좋습니다.

1. 를 설치하려면 명령줄 인터페이스 사용 설명서의 AWS 명령줄 인터페이스 설치의 AWS 지침을 따르십시오. AWS CLI
2. 를 호출하도록 프로필을 AWS CLI 구성하고 설정하려면 AWS 명령줄 인터페이스 사용 설명서의 구성에 나와 AWS CLI 있는 지침을 따르십시오. AWS CLI
3. AWS CLI 프로필이 제대로 구성되었는지 확인하려면 다음 명령을 실행합니다.

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

   프로파일이 올바르게 구성된 경우 다음과 유사한 출력이 표시됩니다.

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

4. Amazon Kendra와 함께 사용하도록 AWS CLI 구성되었는지 확인하려면 다음 명령을 실행합니다.

   ```
   aws kendra help
   ```

   AWS CLI 이 올바르게 구성된 경우 Amazon Kendra, Amazon Kendra 런타임 및 Amazon Kendra 이벤트에 지원되는 AWS CLI 명령 목록이 표시됩니다.

SDK AWS 설정

사용하려는 AWS SDK를 다운로드하고 설치합니다. 이 가이드에서는 Python용 예제를 제공합니다. 다른 AWS SDK에 대한 자세한 내용은 Amazon Web Services용 도구를 참조하십시오.

Python SDK용 패키지의 이름은 Boto3입니다.

아래 Python 명령을 실행하기 전에 먼저 운영 체제에 맞는 Python 3.6 이상을 다운로드하여 설치해야 합니다. Python 3.5 및 이전 버전에 대한 지원은 더 이상 사용되지 않습니다. Python 스크립트 디렉터리에 pip가 포함되어 있지 않은 경우, get-pip.py를 다운로드하여 스크립트 디렉터리에 저장할 수 있습니다. 터미널 프로그램을 사용하여 Python 디렉터리를 Path 또는 환경 변수로 설정할 수도 있습니다.

```
# 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
```

Boto3를 사용하려면 IAM 콘솔을 사용하여 AWS 계정의 인증 자격 증명을 설정해야 합니다.

IAM 액세스 역할 Amazon Kendra

색인, 데이터 원본 또는 FAQ를 만들 때 리소스를 만드는 데 필요한 AWS 리소스에 대한 액세스 권한이 Amazon Kendra 필요합니다. Amazon Kendra Amazon Kendra 리소스를 생성하기 전에 AWS Identity and Access Management (IAM) 정책을 만들어야 합니다. 작업을 호출할 때 연결된 정책과 함께 역할의 Amazon 리소스 이름(ARN)을 제공합니다. 예를 들어 BatchPutDocumentAPI를 호출하여 버킷의 문서를 추가하는 경우 Amazon S3 버킷에 액세스할 수 있는 정책이 포함된 역할을 제공합니다 Amazon Kendra .

Amazon Kendra 콘솔에서 새 IAM 역할을 만들거나 사용할 IAM 기존 역할을 선택할 수 있습니다. 콘솔에는 역할 이름에 “kendra” 또는 “Kendra”라는 문자열이 있는 역할이 표시됩니다.

다음 주제에서는 필수 정책에 대한 세부 정보를 제공합니다. Amazon Kendra 콘솔을 사용하여 IAM 역할을 생성하는 경우 이러한 정책이 자동으로 생성됩니다.

주제

- IAM 인덱스의 역할
- IAM BatchPutDocumentAPI의 역할
- IAM 데이터 소스의 역할
- 가상 사설 클라우드 (VPC) 역할 IAM
- IAM 자주 묻는 질문 (FAQ) 의 역할
- IAM 쿼리 제안의 역할
- IAM 사용자 및 그룹의 주요 매핑을 위한 역할
- IAM 역할: AWS IAM Identity Center
- IAMAmazon Kendra 경험에 대한 역할
- IAM 사용자 지정 문서 보강을 위한 역할

IAM 인덱스의 역할

색인을 만들 때는 에 쓸 수 있는 권한이 있는 IAM 역할을 제공해야 합니다. Amazon CloudWatch 또한 역할을 Amazon Kendra 수임할 수 있는 신뢰 정책을 제공해야 합니다. 제공해야 하는 정책은 다음과 같습니다.

IAM 인덱스의 역할

CloudWatch 로그 Amazon Kendra 액세스를 허용하는 역할 정책.

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

Amazon Kendra 액세스를 허용하는 역할 정책 AWS Secrets Manager. 사용자 컨텍스트를 주요 위치로 사용하는 경우 다음 정책을 사용할 수 있습니다. Secrets Manager

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

IAM BatchPutDocumentAPI의 역할

> **Warning**
>
> Amazon Kendra Amazon Kendra 보안 주체에 S3 버킷과 상호 작용할 수 있는 권한을 부여하는 버킷 정책을 사용하지 않습니다. 대신 IAM 역할을 사용합니다. 실수로 임의의 주체에게 권한을 부여하여 데이터 보안 문제가 발생하지 않도록 해당 Amazon Kendra 구성원이 버킷 정책에 신뢰할 수 있는 구성원으로 포함되지 않도록 하십시오. 하지만 여러 계정에서 버킷을 사용

하도록 버킷 정책을 추가할 수 있습니다 Amazon S3 . 자세한 내용은 계정 간 Amazon S3 사용 정책을 참조하세요. S3 데이터 소스의 IAM 역할에 대한 자세한 내용은 IAM 역할을 참조하세요.

BatchPutDocumentAPI를 사용하여 버킷의 문서를 인덱싱하는 경우 Amazon S3 버킷에 대한 액세스 권한이 Amazon Kendra 있는 IAM 역할을 제공해야 합니다. 또한 역할을 Amazon Kendra 수임할 수 있는 신뢰 정책을 제공해야 합니다. 버킷의 문서가 암호화된 경우 CMK (AWS KMS 고객 마스터 키) 를 사용하여 문서를 해독할 수 있는 권한을 제공해야 합니다.

IAM API의 역할 BatchPutDocument

Amazon S3 버킷 액세스를 허용하는 Amazon Kendra 데 필요한 역할 정책.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

```
    ]
}
```

신뢰 정책에 aws:sourceAccount 및 aws:sourceArn을 포함하는 것이 좋습니다. 이렇게 하면 권한이 aws:sourceAccount aws:sourceArn 제한되고 해당 sts:AssumeRole 작업에 대한 IAM 역할 정책에 제공된 것과 동일한지 안전하게 확인할 수 있습니다. 이렇게 하면 권한이 없는 주체가 사용자의 IAM 역할과 권한에 액세스하는 것을 방지할 수 있습니다. 자세한 내용은 혼동되는 대리인 문제에 대한 AWS Identity and Access Management 안내서를 참조하십시오.

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

AWS KMS 고객 마스터 키 (CMK) Amazon Kendra 를 사용하여 버킷의 문서를 해독할 수 있도록 허용하는 선택적 역할 정책입니다. 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 데이터 소스의 역할

CreateDataSourceAPI를 사용할 때는 리소스에 액세스할 권한이 Amazon Kendra 있는 IAM 역할을 부여해야 합니다. 필요한 구체적인 권한은 데이터 소스에 따라 다릅니다.

IAM 어도비 익스피리언스 매니저 데이터 소스의 역할

Adobe Experience Manager를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- Adobe Experience Manager를 인증하기 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한.
- Adobe Experience Manager 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> 를 Amazon Kendra 통해 Amazon VPC Adobe Experience Manager 데이터 소스를 연결할 수 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM 알프레스코 데이터 원본의 역할

Alfresco를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- Alfresco를 인증하기 위해 AWS Secrets Manager 비밀에 액세스할 수 있는 권한.
- Alfresco 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Alfresco 데이터 소스를 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 사용하는 경우 추가 Amazon VPC 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

```
{
```

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

IAMAurora (MySQL) 데이터 소스의 역할

Aurora (MySQL) 을 사용할 때는 다음 정책이 포함된 역할을 제공합니다.

- Aurora (MySQL) 인증을 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한.
- Aurora (MySQL) 커넥터에 필요한 공개 API를 호출할 수 있는 권한입니다.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Aurora (MySQL) 데이터 원본을 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

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

IAMAurora (PostgreSQL) 데이터 소스에 대한 역할

Aurora (PostgreSQL) 을 사용하는 경우 다음과 같은 정책이 포함된 역할을 제공합니다.

- Aurora (PostgreSQL) 인증을 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한입니다.
- Aurora (PostgreSQL) 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Aurora (PostgreSQL) 데이터 원본을 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 사용하는 경우 추가 Amazon VPC 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

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

IAMAmazon FSx 데이터 원본의 역할

Amazon FSx를 사용할 때 다음 정책이 포함된 역할을 제공합니다.

- Amazon FSx 파일 시스템을 인증하기 위해 AWS Secrets Manager 암호에 액세스할 수 있는 권한.
- Amazon FSx 파일 시스템이 있는 위치 Amazon Virtual Private Cloud (VPC) 에 대한 액세스 권한.
- Amazon FSx 파일 시스템에 사용할 Active Directory의 도메인 이름을 가져올 수 있는 권한
- Amazon FSx 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- 인덱스를 업데이트하기 위해 `BatchPutDocument` 및 `BatchDeleteDocument` API를 호출할 수 있는 권한.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

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

IAM 데이터베이스 데이터 원본의 역할

데이터베이스를 데이터 원본으로 사용하는 경우 연결에 필요한 권한이 있는 역할을 제공합니다 Amazon Kendra . 다음이 포함됩니다.

- 사이트의 사용자 이름과 AWS Secrets Manager 비밀번호가 포함된 비밀번호에 액세스할 수 있는 권한. 보안 암호의 내용에 대해 자세히 알아보려면 데이터 소스를 참조하세요.
- 에서 저장한 사용자 이름 및 암호 암호를 해독하기 위해 CMK (AWS KMS 고객 마스터 키) 를 사용할 수 있는 권한. Secrets Manager
- `BatchPutDocument` 및 `BatchDeleteDocument` 작업을 사용하여 인덱스를 업데이트할 수 있는 권한
- 사이트와 통신하는 데 사용되는 SSL 인증서가 들어 있는 Amazon S3 버킷에 액세스할 수 있는 권한.

> **Note**
>
> 를 Amazon Kendra 통해 데이터베이스 데이터 원본을 연결할 수 Amazon VPC 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

데이터 소스에 사용할 수 있는 두 가지 선택적 정책이 있습니다.

와 통신하는 데 사용되는 SSL 인증서가 들어 있는 Amazon S3 버킷을 암호화한 경우 키에 Amazon Kendra 대한 액세스 권한을 부여하는 정책을 제공하십시오.

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

VPC를 사용하는 경우 필요한 Amazon Kendra 리소스에 대한 액세스를 제공하는 정책을 제공하십시오. 필요한 정책은 데이터 소스, VPC에 대한 IAM 역할을 참조하세요.

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAMAmazon RDS (Microsoft SQL Server) 데이터 원본의 역할

Amazon RDS (Microsoft SQL Server) 데이터 원본 커넥터를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- Amazon RDS (Microsoft SQL Server) 데이터 원본 인스턴스를 인증하기 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한입니다.

- Amazon RDS (Microsoft SQL Server) 데이터 원본 커넥터에 필요한 공개 API를 호출할 수 있는 권한입니다.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Amazon RDS (Microsoft SQL Server) 데이터 원본을 Amazon Kendra 통해 연결할 수 Amazon VPC 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAMAmazon RDS (MySQL) 데이터 소스의 역할

Amazon RDS (MySQL) 데이터 소스 커넥터를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- AWS Secrets Manager 시크릿에 액세스하여 Amazon RDS (MySQL) 데이터 소스 인스턴스를 인증할 수 있는 권한입니다.

- Amazon RDS (MySQL) 데이터 소스 커넥터에 필요한 공개 API를 호출할 수 있는 권한입니다.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Amazon RDS (MySQL) 데이터 원본을 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAMAmazon RDS (Oracle) 데이터 소스의 역할

Amazon RDS Oracle 데이터 소스 커넥터를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- AWS Secrets Manager 비밀번호에 액세스하여 Amazon RDS (Oracle) 데이터 소스 인스턴스를 인증할 수 있는 권한.
- Amazon RDS (Oracle) 데이터 소스 커넥터에 필요한 퍼블릭 API를 호출할 수 있는 권한.

- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> 를 Amazon Kendra 통해 Amazon RDS Amazon VPC Oracle 데이터 소스를 연결할 수 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAMAmazon RDS (PostgreSQL) 데이터 소스에 대한 역할

Amazon RDS (PostgreSQL) 데이터 소스 커넥터를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- Amazon RDS (PostgreSQL) AWS Secrets Manager 데이터 소스 인스턴스를 인증하기 위해 시크릿에 액세스할 수 있는 권한입니다.
- Amazon RDS (PostgreSQL) 데이터 소스 커넥터에 필요한 공개 API를 호출할 수 있는 권한.

- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Amazon RDS (PostgreSQL) 데이터 원본을 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 사용하는 경우 추가 Amazon VPC 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAMAmazon S3 데이터 원본의 역할

> **Warning**
>
> Amazon Kendra Amazon Kendra 보안 주체에게 S3 버킷과 상호 작용할 수 있는 권한을 부여하는 버킷 정책을 사용하지 않습니다. 대신 IAM 역할을 사용합니다. 실수로 임의의 주체에게 권한을 부여하여 데이터 보안 문제가 발생하지 않도록 해당 Amazon Kendra 구성원을 신뢰할 수 있는 구성원으로 버킷 정책에 포함하지 않도록 하세요. 하지만 여러 계정에서 Amazon S3

버킷을 사용하도록 버킷 정책을 추가할 수 있습니다. 자세한 내용은 여러 계정에서 Amazon S3를 사용할 수 있는 정책(아래로 스크롤) 섹션을 참조하세요.

버킷을 데이터 원본으로 사용하는 경우 Amazon S3 버킷에 액세스하고 및 작업을 사용할 권한이 있는 역할을 제공합니다. `BatchPutDocument` `BatchDeleteDocument` Amazon S3 버킷의 문서가 암호화된 경우 CMK (AWS KMS 고객 마스터 키) 를 사용하여 문서를 해독할 수 있는 권한을 제공해야 합니다.

다음 역할 정책은 역할 수임을 Amazon Kendra 허용해야 합니다. 아래로 스크롤하여 역할 수임을 위한 신뢰 정책을 확인하세요.

Amazon S3 버킷을 데이터 원본으로 사용할 수 있도록 허용하는 Amazon Kendra 데 필요한 역할 정책입니다.

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

AWS KMS 고객 마스터 키 (CMK) Amazon Kendra 를 사용하여 버킷의 문서를 해독할 수 있도록 허용하는 선택적 역할 정책입니다. Amazon S3

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

권한을 사용하거나 권한을 AWS KMS 활성화하거나 공유하지 않고도 Amazon Kendra Amazon S3 버킷에 액세스할 수 있도록 허용하는 선택적 역할 정책입니다 Amazon VPC. AWS KMS

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

를 사용하는 동안 Amazon S3 버킷에 액세스할 수 Amazon Kendra 있도록 허용하고 AWS KMS 권한이 활성화된 상태에서 버킷에 Amazon VPC 액세스할 수 있도록 허용하는 선택적 역할 정책입니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

여러 계정에서 Amazon S3 를 사용할 수 있는 정책

Amazon Kendra 인덱스에 사용하는 계정과 다른 계정에 Amazon S3 버킷이 있는 경우 여러 계정에서 사용할 수 있는 정책을 만들 수 있습니다.

Amazon S3 버킷이 Amazon Kendra 인덱스와 다른 계정에 있는 경우 버킷을 데이터 원본으로 사용하기 위한 역할 정책입니다. 참고로 `s3:PutObject` 및 `s3:PutObjectAcl`은 선택 사항이며 액세스 제어 목록용 구성 파일을 포함하려는 경우 이 옵션을 사용하세요.

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

Amazon S3 데이터 원본 역할이 여러 계정에서 버킷에 액세스할 수 있도록 허용하는 Amazon S3 버킷 정책. 참고로 `s3:PutObject` 및 `s3:PutObjectAcl`은 선택 사항이며 액세스 제어 목록용 구성 파일을 포함하려는 경우 이 옵션을 사용하세요.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

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

IAMAmazon Kendra 웹 크롤러 데이터 원본의 역할

Amazon Kendra Web Crawler를 사용할 때는 다음 정책이 포함된 역할을 제공합니다.

- 기본 인증으로 뒷받침되는 웹 사이트 또는 웹 프록시 서버에 연결하기 위한 자격 증명이 포함된 AWS Secrets Manager 암호에 액세스할 수 있는 권한입니다. 보안 암호의 내용에 대해 자세히 알아보려면 웹 크롤러 데이터 소스 사용을 참조하세요.
- 에서 저장한 사용자 이름 및 암호 암호를 해독하기 위해 AWS KMS 고객 마스터 키 (CMK) 를 사용할 수 있는 권한 Secrets Manager
- `BatchPutDocument` 및 `BatchDeleteDocument` 작업을 사용하여 인덱스를 업데이트할 수 있는 권한
- 버킷을 사용하여 시드 URL 또는 사이트맵 목록을 저장하는 경우 Amazon S3 버킷에 액세스할 수 있는 권한을 포함하세요. Amazon S3

> **Note**
>
> 를 통해 Amazon Kendra 웹 크롤러 데이터 소스를 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 사용하는 경우 추가 Amazon VPC 권한을 추가해야 합니다.

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

시드 URL 또는 사이트맵을 Amazon S3 버킷에 저장하는 경우 이 권한을 역할에 추가해야 합니다.

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

역할 수임을 허용하는 Amazon Kendra 신뢰 정책입니다.

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

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

IAMAmazon WorkDocs 데이터 원본의 역할

를 사용할 Amazon WorkDocs 때 다음 정책이 포함된 역할을 제공합니다.

- Amazon WorkDocs 사이트 저장소에 해당하는 디렉터리 ID(조직 ID)를 확인할 수 있는 권한.
- Amazon WorkDocs 사이트 디렉터리가 포함된 Active Directory의 도메인 이름을 가져올 수 있는 권한.
- Amazon WorkDocs 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- 인덱스를 업데이트하기 위해 `BatchPutDocument` 및 `BatchDeleteDocument` API를 호출할 수 있는 권한.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

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

IAM Box 데이터 소스의 역할

Box를 사용할 때는 다음 정책이 적용된 역할을 제공합니다.

- Slack을 인증하기 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한.
- Box 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Box 데이터 소스를 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC Amazon VPCa를 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

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

IAM Confluence 데이터 원본의 역할

IAM 컨플루언스 커넥터 v1.0의 역할

Confluence 서버를 데이터 소스로 사용하는 경우 다음 정책이 적용된 역할을 제공합니다.

- Confluence에 연결하는 데 필요한 자격 증명이 포함된 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한. 보안 암호의 내용에 대해 자세히 알아보려면 Confluence 데이터 소스를 참조하세요.
- 에서 저장한 사용자 이름 및 암호 암호를 해독하기 위해 AWS KMS 고객 마스터 키 (CMK) 를 사용할 수 있는 권한. Secrets Manager
- `BatchPutDocument` 및 `BatchDeleteDocument` 작업을 사용하여 인덱스를 업데이트할 수 있는 권한

> **Note**
>
> Confluence 데이터 소스를 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 사용하는 경우 추가 Amazon VPC 권한을 추가해야 합니다.

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

VPC를 사용하는 경우 필요한 Amazon Kendra 리소스에 대한 액세스를 제공하는 정책을 제공하십시오. 필요한 정책은 데이터 소스, VPC에 대한 IAM 역할을 참조하세요.

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM 컨플루언스 커넥터 v2.0의 역할

Confluence 커넥터 v2.0 데이터 소스의 경우 다음 정책이 포함된 역할을 제공합니다.

- Confluence의 인증 자격 증명이 포함된 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한. 보안 암호의 내용에 대해 자세히 알아보려면 Confluence 데이터 소스를 참조하세요.
- 에서 저장한 사용자 이름 및 암호 암호를 해독하기 위해 AWS KMS 고객 마스터 키 (CMK) 를 사용할 수 있는 권한. AWS Secrets Manager
- `BatchPutDocument` 및 `BatchDeleteDocument` 작업을 사용하여 인덱스를 업데이트할 수 있는 권한

또한 역할을 수임할 수 있는 신뢰 정책을 첨부해야 Amazon Kendra 합니다.

> **Note**
>
> Confluence 데이터 원본을 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 사용하는 경우 추가 Amazon VPC 권한을 추가해야 합니다.

Confluence에 연결할 수 Amazon Kendra 있도록 허용하는 역할 정책.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

```
}
```

IAM Dropbox 데이터 소스의 역할

Dropbox를 사용할 때는 다음 정책이 적용된 역할을 제공합니다.

- AWS Secrets Manager 비밀번호에 액세스하여 Dropbox를 인증할 수 있는 권한
- Dropbox 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Dropbox 데이터 소스를 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM Drupal 데이터 소스의 역할

Drupal을 사용할 때는 다음 정책이 적용된 역할을 제공합니다.

- Drupal을 인증하기 위해 AWS Secrets Manager 비밀에 액세스할 수 있는 권한.

- Drupal 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Drupal 데이터 소스를 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM GitHub 데이터 원본의 역할

GitHub를 사용할 때 다음 정책이 포함된 역할을 제공합니다.

- 인증을 위해 AWS Secrets Manager 비밀에 액세스할 수 있는 권한. GitHub
- 커넥터에 필요한 공개 API를 호출할 수 있는 권한. GitHub

- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> 를 Amazon Kendra 통해 GitHub Amazon VPC 데이터 소스를 연결할 수 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM Gmail 데이터 소스의 역할

Gmail을 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- AWS Secrets Manager 비밀번호에 액세스하여 Gmail을 인증할 수 있는 권한.
- Gmail 커넥터에 필요한 공개 API를 호출할 수 있는 권한.

- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> 를 통해 Gmail 데이터 소스를 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM Google 드라이브 데이터 소스의 역할

Google Workspace 드라이브 데이터 소스를 사용하는 경우 사이트에 연결하는 데 필요한 권한이 있는 역할을 제공합니다 Amazon Kendra . 다음이 포함됩니다.

- Google 드라이브 사이트에 연결하는 데 필요한 클라이언트 계정 이메일, 관리자 계정 이메일, 비공개 키가 포함된 AWS Secrets Manager 비밀번호를 가져오고 해독할 수 있는 권한입니다. 보안 암호의 내용에 대해 자세히 알아보려면 Google Drive 데이터 소스를 참조하세요.
- BatchPutDocument및 BatchDeleteDocumentAPI 사용 권한.

> **Note**
>
> 를 Amazon Kendra 통해 Amazon VPC Google 드라이브 데이터 소스를 연결할 수 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

다음 IAM 정책은 필요한 권한을 제공합니다.

```
{
  "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"
```

```
  }]
}
```

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM IBM DB2 데이터 소스의 역할

IBM DB2 데이터 소스 커넥터를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- IBM DB2 데이터 소스 인스턴스를 인증하기 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한.
- IBM DB2 데이터 소스 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> 를 통해 IBM DB2 데이터 소스를 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

```
}
```

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM Jira 데이터 소스의 역할

Jira를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- AWS Secrets Manager 시크릿에 액세스하여 Jira를 인증할 수 있는 권한.
- Jira 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Jira 데이터 소스를 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM Microsoft Exchange 데이터 원본의 역할

Microsoft Exchange 데이터 원본을 사용하는 경우 사이트에 연결하는 데 필요한 권한이 있는 역할을 제공합니다 Amazon Kendra . 다음이 포함됩니다.

- Microsoft Exchange 사이트에 연결하는 데 필요한 응용 프로그램 ID 및 AWS Secrets Manager 암호 키가 포함된 암호를 가져오고 해독할 수 있는 권한입니다. 보안 암호의 내용에 대해 자세히 알아보려면 Microsoft Exchange 데이터 소스를 참조하세요.
- BatchPutDocument및 BatchDeleteDocumentAPI를 사용할 수 있는 권한.

> **Note**
>
> 를 Amazon Kendra 통해 Microsoft Exchange 데이터 원본을 연결할 수 Amazon VPC 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

다음 IAM 정책은 필요한 권한을 제공합니다.

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

인덱싱할 사용자 목록을 Amazon S3 버킷에 저장하는 경우 S3 `GetObject` 작업을 사용할 권한도 제공해야 합니다. 다음 IAM 정책은 필수 권한을 제공합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

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

IAM Microsoft OneDrive 데이터 원본의 역할

Microsoft OneDrive 데이터 원본을 사용하는 경우 사이트에 연결하는 데 필요한 권한이 있는 역할을 제공합니다 Amazon Kendra . 다음이 포함됩니다.

- 사이트에 연결하는 데 필요한 응용 프로그램 ID 및 AWS Secrets Manager 암호 키가 들어 있는 암호를 가져오고 해독할 수 있는 권한입니다. OneDrive 암호의 내용에 대한 자세한 내용은 Microsoft OneDrive 데이터 원본을 참조하십시오.
- BatchPutDocument및 BatchDeleteDocumentAPI를 사용할 수 있는 권한.

> **Note**
>
> 를 Amazon Kendra 통해 Microsoft OneDrive 데이터 원본을 연결할 수 Amazon VPC 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

다음 IAM 정책은 필요한 권한을 제공합니다.

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

인덱싱할 사용자 목록을 Amazon S3 버킷에 저장하는 경우 S3 `GetObject` 작업을 사용할 권한도 제공해야 합니다. 다음 IAM 정책은 필수 권한을 제공합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

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

IAM Microsoft SharePoint 데이터 원본의 역할

IAM SharePoint커넥터 v1.0의 역할

Microsoft SharePoint 커넥터 v1.0 데이터 원본의 경우 다음 정책이 포함된 역할을 제공합니다.

- SharePoint 사이트의 사용자 이름과 AWS Secrets Manager 암호가 포함된 암호에 액세스할 수 있는 권한. 암호의 내용에 대한 자세한 내용은 Microsoft SharePoint 데이터 원본을 참조하십시오.
- 에서 저장한 사용자 이름 및 암호 암호를 해독하기 위해 CMK (AWS KMS 고객 마스터 키) 를 사용할 수 있는 권한. AWS Secrets Manager
- `BatchPutDocument` 및 `BatchDeleteDocument` 작업을 사용하여 인덱스를 업데이트할 수 있는 권한
- 사이트와 통신하는 데 사용되는 SSL 인증서가 들어 있는 Amazon S3 버킷에 액세스할 수 있는 권한. SharePoint

또한 역할을 수임할 수 Amazon Kendra 있는 신뢰 정책을 첨부해야 합니다.

> **Note**
>
> 를 Amazon Kendra 통해 Microsoft SharePoint 데이터 원본을 연결할 수 Amazon VPC 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

SharePoint 사이트와 통신하는 데 사용되는 SSL 인증서가 들어 있는 Amazon S3 버킷을 암호화한 경우 키에 Amazon Kendra 대한 액세스 권한을 부여하는 정책을 제공하십시오.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

IAM SharePoint커넥터 v2.0의 역할

Microsoft SharePoint 커넥터 v2.0 데이터 원본의 경우 다음 정책이 포함된 역할을 제공합니다.

- SharePoint 사이트의 인증 자격 증명이 포함된 AWS Secrets Manager 암호에 액세스할 수 있는 권한. 암호의 내용에 대한 자세한 내용은 Microsoft SharePoint 데이터 원본을 참조하십시오.
- 에서 저장한 사용자 이름 및 암호 암호를 해독하기 위해 CMK (AWS KMS 고객 마스터 키) 를 사용할 수 있는 권한. AWS Secrets Manager
- `BatchPutDocument` 및 `BatchDeleteDocument` 작업을 사용하여 인덱스를 업데이트할 수 있는 권한

- 사이트와 통신하는 데 사용되는 SSL 인증서가 들어 있는 Amazon S3 버킷에 액세스할 수 있는 권한. SharePoint

또한 역할을 수임할 수 Amazon Kendra 있는 신뢰 정책을 첨부해야 합니다.

> **Note**
>
> 를 Amazon Kendra 통해 Microsoft SharePoint 데이터 원본을 연결할 수 Amazon VPC 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

```
{
  "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": "*"
    }
```

```
  ]
}
```

SharePoint 사이트와 통신하는 데 사용되는 SSL 인증서가 들어 있는 Amazon S3 버킷을 암호화한 경우 키에 Amazon Kendra 대한 액세스 권한을 부여하는 정책을 제공하십시오.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

IAM Microsoft SQL Server 데이터 원본의 역할

Microsoft SQL Server를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- AWS Secrets Manager 암호에 액세스하여 Microsoft SQL Server 인스턴스를 인증할 수 있는 권한입니다.

- Microsoft SQL Server 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Microsoft SQL Server 데이터 원본을 Amazon Kendra 통해 연결할 수 Amazon VPC 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM 마이크로소프트 팀즈 데이터 소스의 역할

Microsoft Teams 데이터 원본을 사용하는 경우 사이트에 연결하는 데 필요한 권한이 있는 역할을 제공합니다 Amazon Kendra . 다음이 포함됩니다.

- Microsoft Teams에 연결하는 데 필요한 클라이언트 ID 및 클라이언트 암호가 포함된 암호를 가져오고 해독할 수 있는 권한입니다. AWS Secrets Manager 보안 암호의 내용에 대해 자세히 알아보려면 Microsoft Teams 데이터 소스를 참조하세요.

> **Note**
>
> Microsoft Teams 데이터 원본을 Amazon Kendra 통해 연결할 수 Amazon VPC 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

다음 IAM 정책은 필요한 권한을 제공합니다.

```
{
  "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"
```

```
  }]
}
```

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM Microsoft Yammer 데이터 원본의 역할

Microsoft Yammer 데이터 원본을 사용하는 경우 사이트에 연결하는 데 필요한 권한이 있는 역할을 제공합니다 Amazon Kendra . 다음이 포함됩니다.

- Microsoft Yammer 사이트에 연결하는 데 필요한 응용 프로그램 ID 및 AWS Secrets Manager 암호 키가 포함된 암호를 가져오고 해독할 수 있는 권한입니다. 보안 암호의 내용에 대해 자세히 알아보려면 Microsoft Yammer 데이터 소스를 참조하세요.
- 및 API를 사용할 수 있는 권한. BatchPutDocumentBatchDeleteDocument

> **Note**
>
> 를 Amazon Kendra 통해 Amazon VPC Microsoft Yammer 데이터 원본을 연결할 수 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

다음 IAM 정책은 필요한 권한을 제공합니다.

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

인덱싱할 사용자 목록을 Amazon S3 버킷에 저장하는 경우 S3 `GetObject` 작업을 사용할 권한도 제공해야 합니다. 다음 IAM 정책은 필수 권한을 제공합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

IAM MySQL 데이터 소스의 역할

MySQL 데이터 소스 커넥터를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- AWS Secrets Manager 비밀번호에 액세스하여 My SQL 데이터 원본 인스턴스를 인증할 수 있는 권한
- MySQL 데이터 소스 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> 를 통해 MySQL 데이터 원본을 연결할 Amazon Kendra 수 있습니다. Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM Oracle 데이터 소스의 역할

Oracle 데이터 소스 커넥터를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- 오라클 데이터 소스 인스턴스를 인증하기 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한.
- Oracle 데이터 소스 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> 를 Amazon Kendra 통해 Amazon VPC Oracle 데이터 소스를 연결할 수 있습니다. 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

```
{
```

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

IAM PostgreSQL 데이터 소스의 역할

PostgreSQL 데이터 소스 커넥터를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- PostgreSQL 데이터 소스 인스턴스를 인증하기 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한입니다.
- PostgreSQL 데이터 소스 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> PostgreSQL 데이터 소스를 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 사용하는 경우 추가 Amazon VPC 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

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

IAM Quip 데이터 소스의 역할

Quip을 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- Quip을 인증하기 위해 AWS Secrets Manager 비밀에 액세스할 수 있는 권한.
- Quip 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Quip 데이터 소스를 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

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

IAM Salesforce 데이터 소스의 역할

Salesforce를 데이터 소스로 사용하는 경우 다음 정책이 적용된 역할을 제공합니다.

- Salesforce 사이트의 AWS Secrets Manager 사용자 이름과 암호가 포함된 암호에 액세스할 수 있는 권한. 보안 암호의 내용에 대해 자세히 알아보려면 Salesforce 데이터 소스를 참조하세요.
- 에서 저장한 사용자 이름 및 암호 암호를 해독하기 위해 AWS KMS 고객 마스터 키 (CMK) 를 사용할 수 있는 권한. Secrets Manager
- `BatchPutDocument` 및 `BatchDeleteDocument` 작업을 사용하여 인덱스를 업데이트할 수 있는 권한

> **Note**
>
> 를 통해 Salesforce 데이터 소스를 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 사용하는 경우 추가 Amazon VPC 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM ServiceNow 데이터 원본의 역할

를 데이터 ServiceNow 원본으로 사용하는 경우 다음 정책이 적용된 역할을 제공합니다.

- ServiceNow 사이트의 사용자 이름과 Secrets Manager 암호가 포함된 암호에 액세스할 수 있는 권한. 보안 암호의 내용에 대해 자세히 알아보려면 ServiceNow 데이터 소스를 참조하세요.
- 에서 저장한 사용자 이름 및 암호 암호를 해독하기 위해 CMK (AWS KMS 고객 마스터 키) 를 사용할 수 있는 권한. Secrets Manager
- `BatchPutDocument` 및 `BatchDeleteDocument` 작업을 사용하여 인덱스를 업데이트할 수 있는 권한

> **Note**
>
> 를 통해 ServiceNow 데이터 소스를 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM Slack 데이터 소스의 역할

Slack을 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- Slack을 인증하기 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한.
- Slack 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> Slack 데이터 소스를 통해 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM Zendesk 데이터 소스의 역할

Zendesk를 사용하는 경우 다음 정책이 포함된 역할을 제공합니다.

- Zendesk Suite를 인증하기 위해 AWS Secrets Manager 시크릿에 액세스할 수 있는 권한
- Zendesk 커넥터에 필요한 공개 API를 호출할 수 있는 권한.
- `BatchPutDocument`, `BatchDeleteDocument`, `PutPrincipalMapping`, `DeletePrincipalMapping`, `DescribePrincipalMapping`, `ListGroupsOlderThanOrderingId` API를 호출할 수 있는 권한.

> **Note**
>
> 를 통해 Zendesk 데이터 소스를 연결할 수 있습니다. Amazon Kendra Amazon VPC 를 Amazon VPC 사용하는 경우 추가 권한을 추가해야 합니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

가상 사설 클라우드 (VPC) 역할 IAM

가상 사설 클라우드 (VPC) 를 사용하여 데이터 소스에 연결하는 경우 다음과 같은 추가 권한을 제공해야 합니다.

VPC 역할 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": "*"
    }
}
```

역할 수임을 Amazon Kendra 허용하는 신뢰 정책.

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

IAM 자주 묻는 질문 (FAQ) 의 역할

CreateFaqAPI를 사용하여 질문과 답변을 인덱스로 로드할 때는 소스 파일이 포함된 Amazon S3 버킷에 액세스할 수 있는 IAM 역할을 제공해야 Amazon Kendra 합니다. 원본 파일이 암호화된 경우 AWS KMS 고객 마스터 키 (CMK) 를 사용하여 파일을 해독할 수 있는 권한을 제공해야 합니다.

IAM FAQ의 역할

Amazon S3 버킷 액세스를 허용하는 Amazon Kendra 데 필요한 역할 정책.

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

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

AWS KMS 고객 마스터 키 (CMK) Amazon Kendra 를 사용하여 버킷의 파일을 해독할 수 있도록 허용하는 선택적 역할 정책입니다. 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"
                    ]
                }
            }
        }
    ]
}
```

역할을 맡을 수 Amazon Kendra 있도록 허용하는 신뢰 정책입니다.

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

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

IAM 쿼리 제안의 역할

파일을 쿼리 제안 차단 목록으로 사용하는 경우 Amazon S3 Amazon S3 파일 및 Amazon S3 버킷에 액세스할 권한이 있는 역할을 제공합니다. Amazon S3 버킷의 차단 목록 텍스트 파일 (Amazon S3 파일) 이 암호화된 경우 CMK (AWS KMS 고객 마스터 키) 를 사용하여 문서를 해독할 수 있는 권한을 제공해야 합니다.

IAM 쿼리 제안의 역할

Amazon S3 파일을 쿼리 제안 차단 목록으로 사용할 수 있도록 허용하는 Amazon Kendra 데 필요한 역할 정책입니다.

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

AWS KMS 고객 마스터 키 (CMK) Amazon Kendra 를 사용하여 버킷의 문서를 해독할 수 있도록 허용하는 선택적 역할 정책입니다. Amazon S3

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

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

역할을 맡을 수 Amazon Kendra 있도록 허용하는 신뢰 정책입니다.

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

IAM 사용자 및 그룹의 주요 매핑을 위한 역할

PutPrincipalMappingAPI를 사용하여 사용자를 그룹에 매핑하여 사용자 컨텍스트별로 검색 결과를 필터링하는 경우 그룹에 속하는 사용자 또는 하위 그룹의 목록을 제공해야 합니다. 한 그룹의 사용자 또는 하위 그룹이 1000명을 초과하는 경우 목록 및 버킷의 Amazon S3 파일에 액세스할 수 있는 권한이 있는 역할을 제공해야 합니다. Amazon S3 Amazon S3 버킷에 있는 목록의 텍스트 Amazon S3 파일 (파일) 이 암호화된 경우 AWS KMS 고객 마스터 키 (CMK) 를 사용하여 문서를 해독할 수 있는 권한을 제공해야 합니다.

IAM 주체 매핑의 역할

Amazon S3 파일을 그룹에 속하는 사용자 및 하위 그룹 목록으로 사용할 수 있도록 허용하는 Amazon Kendra 데 필요한 역할 정책입니다.

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

AWS KMS 고객 마스터 키 (CMK) Amazon Kendra 를 사용하여 버킷의 문서를 해독할 수 있도록 허용하는 선택적 역할 정책입니다. Amazon S3

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

역할을 맡을 수 Amazon Kendra 있도록 허용하는 신뢰 정책입니다.

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

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

신뢰 정책에 aws:sourceAccount 및 aws:sourceArn을 포함하는 것이 좋습니다. 이렇게 하면 권한이 aws:sourceAccount aws:sourceArn 제한되고 해당 sts:AssumeRole 작업에 대한 IAM 역할 정책에 제공된 것과 동일한지 안전하게 확인할 수 있습니다. 이렇게 하면 권한이 없는 주체가 사용자의 IAM 역할과 권한에 액세스하는 것을 방지할 수 있습니다. 자세한 내용은 혼동되는 대리인 문제에 대한 AWS Identity and Access Management 안내서를 참조하십시오.

```
{
    "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 역할: AWS IAM Identity Center

UserGroupResolutionConfiguration개체를 사용하여 AWS IAM Identity Center ID 소스에서 그룹 및 사용자의 액세스 수준을 가져올 때는 액세스 IAM Identity Center 권한이 있는 역할을 제공해야 합니다.

IAM 역할: AWS IAM Identity Center

액세스를 허용하는 Amazon Kendra 데 필요한 역할 정책 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"
              ]
            }
          }
        }
     ]
}
```

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

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

IAMAmazon Kendra 경험에 대한 역할

CreateExperience또는 UpdateExperienceAPI를 사용하여 검색 애플리케이션을 만들거나 업데이트할 때는 필요한 작업과 IAM Identity Center에 액세스할 권한이 있는 역할을 제공해야 합니다.

IAM 검색 경험을 위한 Amazon Kendra 역할

`Query`운영, 운영, `QuerySuggestions` `SubmitFeedback` 운영 및 사용자 및 그룹 정보를 저장하는 IAM Identity Center에 액세스할 수 있도록 허용하는 Amazon Kendra 데 필요한 역할 정책입니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

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

신뢰 정책에 aws:sourceAccount 및 aws:sourceArn을 포함하는 것이 좋습니다. 이렇게 하면 권한이 aws:sourceAccount aws:sourceArn 제한되고 해당 sts:AssumeRole 작업에 대한 IAM 역할 정책에 제공된 것과 동일한지 안전하게 확인할 수 있습니다. 이렇게 하면 권한이 없는 주체가 사용자의 IAM 역할과 권한에 액세스하는 것을 방지할 수 있습니다. 자세한 내용은 혼동되는 대리인 문제에 대한 AWS Identity and Access Management 안내서를 참조하십시오.

```
{
    "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 사용자 지정 문서 보강을 위한 역할

CustomDocumentEnrichmentConfiguration객체를 사용하여 문서 메타데이터 및 콘텐츠의 고급 변경을 적용할 때는 실행 및/또는 에 필요한 권한이 있는 역할을 제공해야 합니다.

PreExtractionHookConfiguration PostExtractionHookConfiguration 수집 프로세스 중에 문서 메타데이터 및 콘텐츠의 고급 변경을 적용하려면 PreExtractionHookConfiguration 및/또는 PostExtractionHookConfiguration에 대한 Lambda 함수를 구성합니다. Amazon S3 버킷에 대해 서버측 암호화를 활성화하기로 선택한 경우, AWS KMS 고객 마스터 키 (CMK) 를 사용하여 버킷에 저장된 객체를 암호화하고 해독할 수 있는 권한을 제공해야 합니다. Amazon S3

IAM 사용자 지정 문서 보강을 위한 역할

실행을 PreExtractionHookConfiguration 허용하고 PostExtractionHookConfiguration 버킷을 암호화하는 Amazon Kendra 데 필요한 역할 정책. Amazon S3

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

Amazon S3 버킷을 PostExtractionHookConfiguration 암호화하지 않고 Amazon Kendra PreExtractionHookConfiguration 실행하도록 허용하는 선택적 역할 정책입니다.

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

역할 수임을 Amazon Kendra 허용하는 신뢰 정책입니다.

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

신뢰 정책에 aws:sourceAccount 및 aws:sourceArn을 포함하는 것이 좋습니다. 이렇게 하면 권한이 aws:sourceAccount aws:sourceArn 제한되고 해당 sts:AssumeRole 작업에 대한 IAM 역할 정책에 제공된 것과 동일한지 안전하게 확인할 수 있습니다. 이렇게 하면 권한이 없는 주체가 사용자의 IAM 역할과 권한에 액세스하는 것을 방지할 수 있습니다. 자세한 내용은 혼동되는 대리인 문제에 대한 AWS Identity and Access Management 안내서를 참조하십시오.

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

```
}
```

Amazon Kendra 배포

웹 사이트에 Amazon Kendra 검색을 배포할 시기가 되면 React와 함께 사용하여 애플리케이션을 빠르게 시작할 수 있는 소스 코드를 제공합니다. 수정된 MIT 라이선스에 따라 소스 코드가 무료로 제공됩니다. 그대로 사용하거나 필요에 따라 변경할 수 있습니다. 제공된 React 앱은 시작하는 데 도움이 되는 예입니다. 프로덕션 지원 앱이 아닙니다.

코드 없이 검색 애플리케이션을 배포하고 액세스 제어가 가능한 검색 페이지의 엔드포인트 URL을 생성하려면 Amazon Kendra Experience Builder를 참조하세요.

다음 예제 코드는 기존 React 웹 애플리케이션에 Amazon Kendra 검색을 추가합니다.

- https://kendrasamples.s3.amazonaws.com/kendrasamples-react-app.zip - 개발자가 기존 React 웹 애플리케이션에 기능적인 검색 환경을 구축하는 데 사용할 수 있는 샘플 파일입니다.

이 예제는 Amazon Kendra 콘솔의 검색 페이지를 모델로 삼았습니다. 검색 및 검색 결과 표시 기능은 동일합니다. 전체 예제를 사용할 수도 있고, 기능 중 하나만 선택하여 사용할 수도 있습니다.

Amazon Kendra 콘솔에서 검색 페이지의 세 가지 구성 요소를 보려면 오른쪽 메뉴에서 코드 아이콘 (</>)을 선택합니다. 각 섹션 위에 포인터를 놓으면 구성 요소에 대한 간략한 설명과 구성 요소 소스의 URL을 확인할 수 있습니다.

주제

개요

기존 React 웹 애플리케이션에 예제 코드를 추가하여 검색을 활성화합니다. 예제 코드에는 새로운 React 개발 환경을 설정하는 단계가 포함된 Readme 파일이 포함되어 있습니다. 코드 예제의 예제 데이터를 사용하여 검색을 시연할 수 있습니다. 예제 코드의 검색 파일 및 구성 요소는 다음과 같이 구성되어 있습니다.

- 기본 검색 페이지(`Search.tsx`) - 모든 구성 요소가 포함된 기본 페이지입니다. 여기에서 애플리케이션을 Amazon Kendra API와 통합할 수 있습니다.
- 검색 창 - 사용자가 검색어를 입력하고 검색 함수를 호출하는 구성 요소입니다.
- 결과 - Amazon Kendra의 결과를 표시하는 구성 요소입니다. 제안된 답변, FAQ 결과, 권장 문서라는 세 가지 구성 요소로 구성됩니다.
- 패싯 - 검색 결과에 패싯을 표시하는 구성 요소로, 패싯을 선택하여 검색 범위를 좁힐 수 있습니다.
- 페이지 매김 - Amazon Kendra의 응답에 페이지를 매기는 구성 요소입니다.

필수 조건

시작하려면 다음이 필요합니다.

- Node.js 및 npm이 설치되었습니다. Node.js 버전 19 이상이 필요합니다.
- Python 3 또는 Python 2가 다운로드 및 설치되었습니다.
- Amazon Kendra에 대한 API 호출이 가능한 SDK for Java 또는 AWS SDK for JavaScript.
- 기존 React 웹 애플리케이션. 예제 코드에는 필수 프레임워크/라이브러리 사용을 포함하여 새로운 React 개발 환경을 설정하는 방법에 대한 단계가 포함된 Readme 파일이 포함되어 있습니다. React 웹 앱 생성에 대한 React 설명서의 빠른 시작 지침을 따를 수도 있습니다.
- 개발 환경에 구성된 필수 라이브러리 및 종속성. 예제 코드에는 필수 라이브러리와 패키지 종속성을 나열하는 Readme 파일이 포함되어 있습니다. 더 이상 `node-sass`가 사용되지 않으므로 `sass`가 필수 항목입니다. 이전에 `node-sass`를 설치한 경우, 제거하고 `sass`를 설치합니다.

예제 설정

React 애플리케이션에 Amazon Kendra 검색을 추가하는 전체 절차는 코드 예제에 포함된 Readme 파일에 있습니다.

kendrasamples-react-app.zip 사용을 시작하려면

1. Node.js 및 npm 다운로드 및 설치를 포함한 필수 조건 작업을 완료했는지 확인하세요.
2. kendrasamples-react-app.zip 다운로드 후 압축을 해제합니다.
3. 터미널을 열고 `aws-kendra-example-react-app/src/services/`로 이동합니다. `local-dev-credentials.json` 열고 보안 인증을 제공합니다. 이 파일을 퍼블릭 리포지토리에 추가하면 안 됩니다.
4. `aws-kendra-example-react-app`으로 이동하여 `package.json`의 종속성을 설치합니다. `npm install`를 실행합니다.
5. 로컬 서버에서 앱의 데모 버전을 실행합니다. `npm start`를 실행합니다. 키보드로 `Cmd/Ctrl + C`를 입력하여 로컬 서버를 중지할 수 있습니다.
6. `package.json`으로 이동하여 포트 또는 호스트(예: IP 주소)를 변경하고 호스트 및 포트를 업데이트할 수 있습니다. `"start": "HOST=[host] PORT=[port] react-scripts start"`. Windows를 사용하는 경우: `"start": "set HOST=[host] && set PORT=[port] && react-scripts start"`.
7. 등록된 웹사이트 도메인이 있는 경우 앱 이름 뒤의 `package.json`에 이를 지정할 수 있습니다. 예: `"homepage": "https://mywebsite.com"`. `npm install`을 다시 실행하여 새 종속성을 업데이트한 다음 `npm start`를 실행해야 합니다.
8. 앱을 빌드하려면 `npm build`를 수행합니다. 빌드 디렉터리의 콘텐츠를 호스팅 제공업체에 업로드합니다.

> **Warning**
>
> React 앱은 아직 프로덕션 준비가 되지 않았습니다. Amazon Kendra 검색용 앱을 배포하는 예시입니다.

기본 검색 페이지

기본 검색 페이지(`Search.tsx`)에는 모든 예제 검색 구성 요소가 들어 있습니다. 여기에는 출력용 검색 창 구성 요소, Query API의 응답을 표시하는 결과 구성 요소, 응답을 통한 페이징을 위한 페이지 매김 구성 요소가 포함됩니다.

검색 구성 요소

검색 구성 요소는 쿼리 텍스트를 입력할 수 있는 텍스트 상자를 제공합니다. onSearch 함수는 Search.tsx에서 기본 함수를 호출하여 Amazon Kendra Query API 호출을 수행하는 후크입니다.

결과 구성 요소

결과 구성요소는 Query API의 응답을 보여줍니다. 결과는 세 개의 개별 영역에 표시됩니다.

- 제안된 답변 - Query API에서 가장 많이 반환된 결과입니다. 제안된 답변은 최대 3개까지 포함됩니다. 응답에는 ANSWER 결과 유형이 있습니다.
- FAQ 답변 - 응답에서 반환하는 자주 묻는 질문 결과입니다. FAQ는 인덱스에 별도로 추가됩니다. 응답에는 QUESTION_ANSWER 유형이 있습니다. 자세한 내용은 질문 및 답변을 참조하세요.
- 권장 문서 - Amazon Kendra가 응답으로 반환하는 추가 문서입니다. Query API의 응답에는 DOCUMENT 유형이 있습니다.

결과 구성 요소는 강조 표시, 제목, 링크 등의 기능에 대한 구성 요소 세트를 공유합니다. 결과 구성 요소가 작동하려면 공유 구성 요소가 있어야 합니다.

패싯 구성 요소

패싯 구성 요소는 검색 결과에서 사용할 수 있는 패싯을 나열합니다. 각 패싯은 작성자와 같은 특정 차원에 따라 응답을 분류합니다. 목록에서 하나를 선택하여 특정 패싯으로 검색을 세분화할 수 있습니다.

패싯을 선택하면 구성 요소가 속성 필터로 Query를 호출하여 검색을 패싯과 일치하는 문서로 제한합니다.

페이지 매김 구성 요소

페이지 매김 구성 요소를 사용하면 Query API의 검색 결과를 여러 페이지에 표시할 수 있습니다. PageSize 및 PageNumber 파라미터와 함께 Query API를 호출하여 특정 결과 페이지를 가져옵니다.

코드 없이 검색 환경 구축

프런트엔드 코드 없이도 Amazon Kendra 검색 애플리케이션을 구축하고 배포할 수 있습니다. Amazon Kendra Experience Builder를 사용하면 클릭 몇 번으로 모든 기능을 갖춘 검색 애플리케이션을 구축하

고 배포하여 즉시 검색을 시작할 수 있습니다. 검색 페이지를 사용자 지정으로 디자인하고 검색을 조정하여 사용자의 요구에 맞게 환경을 조정할 수 있습니다. Amazon Kendra는 검색 페이지의 완전히 호스팅된 고유한 엔드포인트 URL을 생성하여 문서 및 FAQ 검색을 시작합니다. 검색 경험에 대한 개념 증명을 신속하게 구축하여 다른 사람들과 공유할 수 있습니다.

빌더에 제공되는 검색 경험 템플릿을 사용하여 검색을 사용자 지정할 수 있습니다. 검색 경험을 구축하기 위해 협업하거나 조정 목적으로 검색 결과를 평가할 다른 사람들을 초대할 수 있습니다. 사용자가 검색을 시작할 수 있는 검색 환경이 준비되면 보안 엔드포인트 URL을 공유하기만 하면 됩니다.

검색 Experience Builder의 작동 방식

검색 경험을 구축하는 전체 프로세스는 다음과 같습니다.

1. 검색 환경에 이름과 설명을 지정하고, 사용할 데이터 소스를 선택하여 검색 환경을 만듭니다.
2. AWS IAM Identity Center에서 사용자 및 그룹 목록을 구성한 다음, 검색 환경에 대한 액세스 권한을 할당합니다. 본인을 경험의 소유자로 포함하세요. 자세한 내용은 the section called “검색 페이지에 대한 액세스 권한 제공” 섹션을 참조하세요.
3. Amazon Kendra Experience Builder를 열어 검색 페이지를 설계하고 조정할 수 있습니다. 소유-편집 액세스 권한 또는 보기-검색 액세스 권한을 할당한 다른 사람들과 검색 경험의 엔드포인트 URL을 공유할 수 있습니다.

CreateExperience API를 호출하여 검색 환경을 만들고 구성합니다. 콘솔을 사용하는 경우 인덱스를 선택한 다음 탐색 메뉴에서 경험을 선택하여 경험을 구성합니다.

검색 경험을 설계하고 조정

검색 환경을 만들고 구성한 후에는 엔드포인트 URL을 사용하여 검색 환경을 열고 편집자 액세스 권한을 가진 소유자로서 검색을 사용자 지정하기 시작합니다. 검색 창에 쿼리를 입력한 다음, 사이드 패널의 편집 옵션을 사용하여 검색을 사용자 지정하여 페이지에 어떻게 적용되는지 확인합니다. 게시할 준비가 되면 게시를 선택합니다. 실시간 보기로 전환을 토글하여 검색 페이지의 최신 게시 버전을 보거나 빌드 모드로 전환을 토글하여 검색 페이지를 편집하거나 사용자 지정할 수도 있습니다.

검색 환경을 사용자 지정할 수 있는 방법은 다음과 같습니다.

필터

패싯된 검색을 추가하거나 문서 속성별로 필터링합니다. 이 항목에는 사용자 지정 속성이 포함됩니다. 자체 구성된 메타데이터 필드를 사용하여 필터를 추가할 수 있습니다. 예를 들어, 각 도시 범주별로 패싯 검색을 하려면 모든 도시 범주가 포함된 `_category` 사용자 지정 문서 속성을 사용합니다.

제안된 답변

기계 학습으로 생성된 답변을 사용자 쿼리에 추가합니다. 예: '이 과정은 얼마나 어려운가요?'. Amazon Kendra는 과정 난이도와 관련하여 모든 문서에서 가장 관련성이 높은 텍스트를 검색하고 가장 관련성이 높은 답변을 제안할 수 있습니다.

FAQ

FAQ 문서를 추가하여 자주 묻는 질문에 대한 답변을 제공합니다. 예: '이 과정을 완료하는 데 몇 시간이 걸리나요?'. Amazon Kendra가 이 질문에 대한 답변이 들어 있는 FAQ 문서를 사용하여 답변을 제공할 수 있습니다.

정렬

검색 결과 정렬 기능을 추가하여 사용자가 관련성, 만든 시간, 마지막 업데이트 시간 및 기타 정렬 기준에 따라 결과를 정리할 수 있도록 합니다.

문서

문서 또는 검색 결과가 검색 페이지에 표시되는 방식을 구성합니다. 페이지에 표시할 결과 수를 구성하고, 페이지 번호와 같은 페이지 매김을 포함하고, 사용자 피드백 버튼을 활성화하고, 검색 결과에 문서 메타데이터 필드가 표시되는 방식을 정렬할 수 있습니다.

언어

선택한 언어로 검색 결과 또는 문서를 필터링할 언어를 선택합니다.

검색 창

검색 창의 크기 및 자리 표시자 텍스트를 구성하고, 쿼리 제안을 허용할 수 있습니다.

관련성 조정

사용자가 문서를 검색할 때 이러한 필드에 더 많은 가중치를 부여하도록 문서 메타데이터 필드에 부스팅을 추가합니다. 1부터 10까지 점진적으로 가중치를 추가할 수 있습니다. 텍스트, 날짜 및 숫자 필드 유형을 부스팅할 수 있습니다. 예를 들어 `_last_updated_at` 및 `_created_at`에 다른 필드보다 더 높은 가중치 또는 중요도를 부여하려면 해당 필드에 중요도에 따라 1~10의 가중치를 부여합니다. 각 검색 애플리케이션 또는 경험에 대해 서로 다른 관련성 조정 구성을 적용할 수 있습니다.

검색 페이지에 대한 액세스 권한 제공

검색 환경에 대한 액세스는 IAM Identity Center를 통해 이루어집니다. 검색 환경을 구성하면 Identity Center 디렉터리에 나열된 다른 사람들에게 Amazon Kendra 검색 페이지에 대한 액세스 권한을 부여합니다. 해당 사용자는 IAM Identity Center의 보안 인증을 사용하여 로그인하여 검색 페이지에 액세스하도록 지시하는 이메일을 받습니다. AWS Organizations에 조직 수준 또는 계정 소유자 수준에서 IAM Identity Center를 설정해야 합니다. IAM Identity Center 설정에 대한 자세한 내용은 IAM Identity Center 시작하기를 참조하세요.

IAM Identity Center에서 검색 환경으로 사용자 ID를 활성화하고, API 또는 콘솔을 사용하여 최종 사용자 또는 소유자에게 액세스 권한을 할당합니다.

- 최종 사용자: 쿼리를 발행하고, 검색과 관련하여 제안된 답변을 받고, Amazon Kendra에 피드백을 제공하여 검색을 지속적으로 개선할 수 있도록 허용했습니다.
- 소유자: 검색 페이지 디자인을 사용자 지정하고, 검색을 조정하고, 검색 애플리케이션을 최종 사용자로 사용하도록 허용했습니다. 콘솔에서 최종 사용자 액세스를 비활성화하는 기능은 현재 지원되지 않습니다.

다른 사람에게 검색 환경에 대한 액세스 권한을 할당하려면 먼저 ExperienceConfiguration 객체를 사용하여 Amazon Kendra 경험을 통해 IAM Identity Center에서 사용자 ID를 활성화해야 합니다. 사용자 이름 또는 이메일 주소와 같이 사용자 식별자가 포함된 필드 이름을 지정합니다. 그런 다음 AssociateEntitiesToExperience API를 사용하여 사용자 목록에 검색 환경 액세스 권한을 부여하고 AssociatePersonasToEntities API를 사용하여 해당 권한을 최종 사용자 또는 소유자로 정의합니다. EntityConfiguration 객체를 사용하여 각 사용자 또는 그룹을 지정하고 EntityPersonaConfiguraton 객체를 사용하여 해당 사용자 또는 그룹이 최종 사용자인지 소유자인지 지정합니다.

콘솔을 사용하여 다른 사용자에게 검색 환경에 대한 액세스 권한을 할당하려면 먼저 경험을 만들고 자신의 ID와 소유자임을 확인해야 합니다. 그런 다음 다른 사용자나 그룹을 최종 사용자 또는 소유자로 할당할 수 있습니다. 콘솔에서 색인을 선택한 다음 탐색 메뉴에서 경험을 선택합니다. 경험을 만든 후 목록에서 경험을 선택할 수 있습니다. 액세스 관리로 이동하여 사용자 또는 그룹을 최종 사용자 또는 소유자로 할당합니다.

검색 환경 구성

다음은 검색 환경을 구성 또는 생성하는 예입니다.

Console

Amazon Kendra 검색 경험을 만들려면

1. 왼쪽 탐색 창의 인덱스에서 경험을 선택한 다음 경험 생성을 선택합니다.
2. 경험 구성 페이지에서 경험의 이름과 설명을 입력하고, 콘텐츠 소스를 선택하고, 경험에 대한 IAM 역할을 선택합니다. IAM 역할에 대한 자세한 내용은 Amazon Kendra 경험을 위한 IAM 역할을 참조하세요.
3. Identity Center 디렉터리에서 ID 확인 페이지에서 사용자 ID(예: 이메일)를 선택합니다. Identity Center 디렉터리가 없는 경우 전체 이름과 이메일을 입력하여 Identity Center 디렉터리를 생성하기만 하면 됩니다. 여기에는 경험의 사용자도 포함되며 소유자 액세스 권한이 자동으로 할당됩니다.
4. Experience Builder 열기 검토 페이지에서 구성 세부 정보를 검토하고 경험 생성 및 Experience Builder 열기를 선택하여 검색 페이지 편집을 시작합니다.

CLI

Amazon Kendra 경험 생성

```
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

Amazon Kendra 경험 생성

```
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

Amazon Kendra 생성

```
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.");
    }
}
```

용량 조정

Amazon Kendra 인덱스에 대한 자원을 용량 단위로 제공합니다. 각 용량은 인덱스에 사용할 추가 리소스를 제공합니다. 문서 스토리지와 쿼리에는 별도의 용량 단위가 있습니다. Amazon Kendra Enterprise Edition 인덱스에는 용량 단위만 추가할 수 있습니다. 개발자 에디션 인덱스에는 용량을 추가할 수 없습니다.

문서 저장 용량 단위는 인덱스에 다음과 같은 추가 저장 공간을 제공합니다.

- 100,000개의 문서 또는 30GB의 저장 공간.

쿼리 용량 단위는 인덱스에 다음과 같은 추가 쿼리를 제공합니다.

- 초당 0.1쿼리 또는 일일 8,000쿼리.

각 인덱스의 기본 용량은 1용량 단위(스토리지 30GB, 초당 0.1쿼리)와 같습니다. 각 추가 용량 단위마다 추가 비용이 부과됩니다. 자세한 내용은 Amazon Kendra 요금을 참조하십시오.

최대 100 추가 용량 단위를 스토리지에 추가하고 인덱스에 대한 리소스를 쿼리할 수 있습니다. 단위가 더 필요한 경우 지원 팀에 문의하면 됩니다.

사용량 요구 사항에 맞게 하루 최대 5회까지 용량 단위를 조정할 수 있습니다. 문서 저장 용량을 인덱스에 저장된 문서 개수 이하로 줄일 수는 없습니다. 예를 들어 150,000개의 문서를 저장하는 경우 저장 용량을 추가 단위 1개 이하로 줄일 수 없습니다.

콘솔에서 인덱스 이름을 선택하여 인덱스 설정 및 기타 정보를 열어 인덱스가 사용 중인 리소스를 보거나 DescribeIndexAPI를 사용할 수 있습니다.

Amazon Kendra 인덱스 용량을 초과한 경우에도 예외를 반환합니다. 모든 문서의 추출된 총 크기가 인덱스 한도를 초과하면 `ServiceQuotaExceededException`이 발생합니다. 문서 수가 인덱스 한도를 초과할 경우 각 문서에 대해 `InvalidRequest`가 표시됩니다. 초당 쿼리 수가 한도를 초과할 경우 `ThrottlingException`이 발생합니다. 한도에 대한 자세한 내용은 Amazon Kendra할당량 섹션을 참조하십시오.

누적된 쿼리는 최대 24시간까지 지속됩니다.

용량 확인

Amazon Kendra 콘솔에서 인덱스 이름을 선택하여 세부 정보에 액세스하여 인덱스가 사용 중인 리소스를 확인하세요. 콘솔도 사용량 그래프를 제공하므로 인덱스에서 사용하는 스토리지 및 쿼리 용량을 확인할 수 있습니다. 이 정보를 사용하여 언제 추가 용량을 추가할 것인지 계획할 수 있습니다.

문서 스토리지 및 쿼리 사용을 보려면(콘솔)

1. AWS Management Console 로그인하고 https://console.aws.amazon.com/kendra/home 에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 목록에서 액세스할 인덱스를 선택합니다.
3. 설정 섹션으로 스크롤하여 현재 총 문서 스토리지와 쿼리 용량을 확인합니다.

API를 사용하여 용량을 보려면 Amazon Kendra DescribeIndexAPI의 `CapacityUnits` 매개변수를 사용하십시오.

용량 추가 및 제거

인덱스에 추가 용량이 필요한 경우 콘솔 또는 Amazon Kendra API를 사용하여 인덱스를 추가할 수 있습니다.

스토리지 또는 쿼리 용량을 추가하거나 제거하려면(콘솔)

1. AWS Management Console 로그인하고 https://console.aws.amazon.com/kendra/home 에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 목록에서 액세스할 인덱스를 선택합니다.
3. 편집을 선택하거나 작업 드롭다운에서 편집을 선택합니다.
4. 다음을 선택하여 프로비저닝 세부 정보 페이지로 이동합니다.
5. 문서 스토리지 및/또는 쿼리 용량 단위를 추가 또는 제거합니다.
6. 계속 다음을 선택하여 검토 페이지로 이동한 다음 업데이트를 선택하여 변경 내용을 저장합니다.

인덱스의 용량을 업데이트할 경우 변경 사항이 적용되기까지 몇 분 정도 걸릴 수 있습니다.

API를 사용하여 용량을 추가하거나 제거하려면 Amazon Kendra UpdateIndexAPI의 `CapacityUnits` 매개변수를 사용하십시오.

Amazon Kendra 인텔리전트 랭킹 용량

용량 단위는 재평가 실행 계획에 대해 초당 다음과 같은 추가 재평가 요청을 제공합니다. 재평가 실행 계획은 Rescore API를 프로비저닝하는 데 사용되는 리소스입니다.

- 초당 0.01 요청.

각 재평가 실행 계획에는 용량 단위 1의 기본 용량이 제공됩니다(초당 0.01 요청). 각 추가 용량 단위마다 추가 비용이 부과됩니다. 자세한 내용은 Amazon Kendra 요금을 참조하십시오.

재평가 실행 계획에 최대 1000의 추가 용량 단위를 추가할 수 있습니다. 단위가 더 필요한 경우 지원 팀에 문의하면 됩니다.

쿼리 제안 용량

쿼리 제안을 사용하는 경우 기본 쿼리 용량은 초당 2.5회 GetQuerySuggestions호출입니다. `GetQuerySuggestions` 용량은 인덱스에 대해 프로비저닝된 쿼리 용량의 5배 또는 기본 용량인 초당 2.5회 호출 중 더 큰 값입니다. 예를 들어, 인덱스의 기본 용량은 초당 쿼리 0.1개이고 `GetQuerySuggestions`의 기본 용량은 초당 호출 2.5건입니다. 인덱스에 대해 초당 쿼리 총 0.2개에 초당 쿼리 0.1개를 추가하면 `GetQuerySuggestions` 용량은 초당 호출 2.5건(초당 쿼리 0.2개의 5배 이상)입니다.

Amazon Kendra 경험 수용 능력

검색 환경 용량

Amazon Kendra 초당 요청 15개 `QueryQuerySuggestions`, 쿼리 버스팅의 `SubmitFeedback` 경우 초당 40개 요청으로 Amazon Kendra 경험에 맞게 조절하기 시작합니다. 쿼리 용량 단위가 150을 초과하는 인덱스의 경우 이러한 제한이 계속 적용됩니다.

예를 들어 인덱스의 쿼리 용량 단위는 150이므로 검색 환경 애플리케이션은 초당 15개의 요청을 처리할 수 있습니다. 하지만 200 쿼리 용량 단위로 확장해도 검색 환경 앱은 여전히 초당 15개의 요청만 처리할 수 있습니다. 인덱스를 100 쿼리 용량 단위로 제한하면 검색 환경 앱은 초당 10개의 요청만 처리할 수 있습니다.

적응형 쿼리 버스팅

Amazon Kendra 프로비저닝된 기본 용량은 쿼리 용량 단위 1개입니다. 하루 최대 8,000 쿼리를 사용할 수 있으며 최소 처리량은 초당 0.1쿼리(쿼리 용량 단위당)입니다. 누적된 쿼리는 최대 24시간 지속되며 트래픽 폭증을 수용할 수 있습니다. 허용되는 버스트 양은 특정 시점의 클러스터 부하에 따라 달라지므로 바뀔 수 있습니다. 최대 부하 수준을 처리할 수 있을 만큼 충분한 쿼리 용량 단위를 프로비저닝하십시오.

기본 제공되는 적응형 쿼리 버스팅은 프로비저닝된 처리량을 넘어서는 예상치 못한 트래픽 폭주를 처리하는 적응형 접근 Amazon Kendra방식입니다. 적응형 쿼리 버스팅은 Amazon Kendra엔터프라이즈 에디션에서 사용할 수 있습니다.

적응형 쿼리 버스팅은 사용하지 않는 쿼리 용량을 적용하여 예상치 못한 트래픽을 처리할 수 있는 기본 제공 기능입니다. Amazon Kendra 초당 프로비저닝된 쿼리의 속도로 인덱스에 프로비저닝한 최대 쿼리 수까지 미사용 쿼리를 1초마다 누적합니다. Amazon Kendra 이렇게 누적된 쿼리는 할당된 용량을 초과하는 예상치 못한 트래픽에 사용됩니다. 적응형 쿼리 버스팅의 최적 성능은 총 인덱스 크기, 쿼리 복잡성, 누적된 미사용 쿼리, 인덱스의 전체 부하와 같은 여러 요인에 따라 달라질 수 있습니다. 버스팅 용량을 정확하게 측정하려면 자체 부하 테스트를 수행하는 것이 좋습니다.

시작하기

이 섹션에서는 데이터 소스를 만들고 Amazon Kendra 색인에 문서를 추가하는 방법을 보여줍니다. AWS 콘솔, 를 사용하는 Python 프로그램 및 를 사용하는 Java 프로그램에 대한 지침이 제공됩니다 AWS SDK for Java. AWS CLI AWS SDK for Python (Boto3)

주제

필수 조건

다음 단계는 시작하기 연습의 전제 조건을 나타냅니다. 이 단계에서는 계정을 설정하고, 사용자 대신 전화를 걸 수 있는 Amazon Kendra 권한을 부여하는 IAM 역할을 만들고, Amazon S3 버킷에서 문서를 인덱싱하는 방법을 보여줍니다. S3 버킷이 예로 사용되지만 Amazon Kendra 지원하는 다른 데이터 소스를 사용할 수 있습니다. 데이터 소스를 참조하세요.

가입하십시오. AWS 계정

계정이 없는 경우 다음 단계를 완료하여 계정을 만드세요. AWS 계정

가입하려면 AWS 계정

1. https://portal.aws.amazon.com/billing/signup을 여세요.
2. 온라인 지시 사항을 따르세요.

   등록 절차 중에는 전화를 받고 키패드로 인증 코드를 입력하는 과정이 있습니다.

   에 AWS 계정가입하면 AWS 계정 루트 사용자a가 생성됩니다. 루트 사용자에게는 계정의 모든 AWS 서비스 및 리소스 액세스 권한이 있습니다. 보안 모범 사례는 사용자에게 관리 액세스 권한

을 할당하고, 루트 사용자만 사용하여 루트 사용자 액세스 권한이 필요한 작업을 수행하는 것입니다.

AWS 가입 절차가 완료된 후 확인 이메일을 보냅니다. 언제든지 https://aws.amazon.com/으로 가서 내 계정(My Account)을 선택하여 현재 계정 활동을 보고 계정을 관리할 수 있습니다.

관리자 액세스 권한이 있는 사용자 생성

등록한 AWS 계정후에는 일상적인 작업에 루트 사용자를 사용하지 않도록 관리 사용자를 보호하고 AWS IAM Identity Center활성화하고 생성하십시오 AWS 계정 루트 사용자.

보안을 유지하세요. AWS 계정 루트 사용자

1. 루트 사용자를 선택하고 AWS 계정 이메일 주소를 입력하여 계정 AWS Management Console소유자로 로그인합니다. 다음 페이지에서 비밀번호를 입력합니다.

   루트 사용자를 사용하여 로그인하는 데 도움이 필요하면AWS 로그인 사용 설명서의 루트 사용자로 로그인을 참조하세요.
2. 루트 사용자의 다중 인증(MFA)을 활성화합니다.

   지침은 IAM 사용 설명서의 AWS 계정 루트 사용자 (콘솔) 에 대한 가상 MFA 디바이스 활성화를 참조하십시오.

관리자 액세스 권한이 있는 사용자 생성

1. IAM Identity Center를 활성화합니다.

   지침은 AWS IAM Identity Center 사용 설명서의 AWS IAM Identity Center설정을 참조하세요.
2. IAM Identity Center에서 사용자에게 관리 액세스 권한을 부여합니다.

   를 ID 소스로 사용하는 방법에 대한 자습서는 사용 설명서의 기본값으로 IAM Identity Center 디렉터리사용자 액세스 구성을 참조하십시오. IAM Identity Center 디렉터리 AWS IAM Identity Center

관리 액세스 권한이 있는 사용자로 로그인

- IAM IDentity Center 사용자로 로그인하려면 IAM IDentity Center 사용자를 생성할 때 이메일 주소로 전송된 로그인 URL을 사용합니다.

IAM Identity Center 사용자를 사용하여 로그인하는 데 도움이 필요하면 사용 설명서의 AWS 액세스 포털 로그인을 참조하십시오.AWS 로그인

추가 사용자에게 액세스 권한 할당

1. IAM Identity Center에서 최소 권한 적용 모범 사례를 따르는 권한 세트를 생성합니다.

   지침은AWS IAM Identity Center 사용 설명서의 Create a permission set를 참조하세요.
2. 사용자를 그룹에 할당하고, 그룹에 Single Sign-On 액세스 권한을 할당합니다.

   지침은AWS IAM Identity Center 사용 설명서의 Add groups를 참조하세요.

- 테스트할 문서가 포함된 S3 버킷을 사용하는 경우 Amazon Kendra, 사용 중인 동일한 리전에 S3 버킷을 생성하십시오. Amazon Kendra지침은 Amazon Simple Storage Service Console 사용 설명서의 S3 버킷 생성 및 구성을 참조하세요.

  S3 버킷에 문서를 업로드합니다. 지침은 Amazon Simple Storage Service 사용 설명서의 객체 업로드, 다운로드 및 관리를 참조하세요.

  다른 데이터 소스를 사용하는 경우 데이터 소스에 연결하려면 활성 사이트와 보안 인증이 있어야 합니다.

콘솔을 사용하여 시작하는 경우, Amazon Kendra 콘솔 시작하기로 시작하세요.

Amazon Kendra 리소스: AWS CLI, SDK, 콘솔

CLI, SDK 또는 콘솔을 사용하는 경우 특정 권한이 필요합니다.

CLI, SDK 또는 Amazon Kendra 콘솔에 사용하려면 사용자를 대신하여 리소스를 생성하고 관리할 수 Amazon Kendra 있는 권한이 있어야 합니다. 사용 사례에 따라 이러한 권한에는 Amazon Kendra API 자체에 대한 액세스, 사용자 지정 CMK를 통해 데이터를 암호화하려는 AWS KMS keys 경우 Identity Center 디렉터리 (검색 경험과 AWS IAM Identity Center 통합하거나 생성하려는 경우) 에 대한 액세스가 포함됩니다. 다양한 사용 사례에 대한 전체 권한 목록은 IAM 역할을 참조하세요.

먼저, 아래 권한을 IAM 사용자에게 연결해야 합니다.

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

둘째, CLI 또는 SDK를 사용하는 경우 액세스할 IAM 역할과 정책도 생성해야 합니다. Amazon CloudWatch Logs콘솔을 사용하는 경우 이를 위한 IAM 역할과 정책을 만들 필요가 없습니다. 콘솔 절차의 일부로 이를 생성합니다.

액세스를 허용하는 Amazon Kendra 및 SDK의 IAM 역할 AWS CLI 및 정책을 만들려면 Amazon CloudWatch Logs

1. 에 AWS Management Console 로그인하고 https://console.aws.amazon.com/iam/ 에서 IAM 콘솔을 엽니다.
2. 왼쪽에서 정책을 선택한 후 정책 생성을 선택합니다.
3. JSON을 선택한 다음, 기본 정책을 다음과 같이 바꿉니다.

```
{
    "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. 정책 검토를 선택합니다.
5. "KendraPolicyForGettingStartedIndex" 정책 이름을 입력한 후 정책 생성을 선택합니다.
6. 왼쪽 메뉴에서 역할을 선택한 후 역할 생성을 선택합니다.
7. 다른 AWS 계정을 선택한 다음 계정 ID에 계정 ID를 입력합니다. 다음: 권한을 선택합니다.
8. 위에서 만든 정책을 선택한 후 다음: 태그를 선택합니다.
9. 태그를 추가하지 마세요. 다음: 검토를 선택합니다.
10. "KendraRoleForGettingStartedIndex" 역할 이름을 지정한 다음 역할 생성을 선택합니다.
11. 방금 생성한 역할을 찾습니다. 역할 이름을 선택하여 요약을 엽니다. 신뢰 관계를 선택한 후 신뢰 관계 편집을 선택합니다.
12. 기존 신뢰 관계를 다음과 같이 교체합니다.

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

13. 신뢰 정책 업데이트를 선택합니다.

셋째, 를 사용하여 문서를 저장하거나 S3를 사용하여 Amazon Kendra테스트하는 경우 버킷에 액세스하기 위한 IAM 역할 및 정책도 생성해야 합니다. Amazon S3 다른 데이터 소스를 사용하는 경우 데이터 소스의IAM 역할을 참조하세요.

Amazon S3 버킷에 액세스하고 Amazon Kendra 인덱싱할 수 있는 IAM 역할 및 정책을 만들려면

1. https://console.aws.amazon.com/iam/ 에서 AWS Management Console 로그인하고 IAM 콘솔을 엽니다.
2. 왼쪽에서 정책을 선택한 후 정책 생성을 선택합니다.
3. JSON을 선택한 다음, 기본 정책을 다음과 같이 바꿉니다.

```
{
    "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. 정책 검토를 선택합니다.
5. 정책 이름을 KendraPolicyForGettingStartedDataSource ""로 지정한 다음 Create policy (정책 생성) 를 선택합니다.
6. 왼쪽 메뉴에서 역할을 선택한 후 역할 생성을 선택합니다.
7. 다른 AWS 계정을 선택한 다음 계정 ID에 계정 ID를 입력합니다. 다음: 권한을 선택합니다.
8. 위에서 만든 정책을 선택한 후 다음: 태그를 선택합니다.
9. 태그를 추가하지 마세요. 다음: 검토를 선택합니다.
10. 역할 이름을 KendraRoleForGettingStartedDataSource ""로 지정한 다음 역할 생성을 선택합니다.
11. 방금 생성한 역할을 찾습니다. 역할 이름을 선택하여 요약을 엽니다. 신뢰 관계를 선택한 후 신뢰 관계 편집을 선택합니다.
12. 기존 신뢰 관계를 다음과 같이 교체합니다.

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

```
    ]
}
```

13. 신뢰 정책 업데이트를 선택합니다.

원하는 Amazon Kendra API 사용 방식에 따라 다음 중 하나를 수행하십시오.

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

Amazon Kendra 콘솔 시작하기

다음 절차는 AWS 콘솔을 사용하여 Amazon Kendra 색인을 만들고 테스트하는 방법을 보여줍니다. 이 절차에서 인덱스와 인덱스의 데이터 소스를 만듭니다. 마지막으로 검색 요청을 하여 인덱스를 테스트합니다.

1단계: 인덱스 생성하기(콘솔)

1. AWS 관리 콘솔에 로그인하고 https://console.aws.amazon.com/kendra/ 에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 섹션에서 인덱스 생성을 선택합니다.
3. 인덱스 세부정보 지정 페이지에서 인덱스 이름과 설명을 입력합니다.
4. IAM 역할에서 새 역할 생성을 선택한 다음, 역할에 이름을 지정합니다. IAM 역할의 접두사는 "AmazonKendra- “입니다.
5. 다른 모든 필드의 기본값은 그대로 둡니다. 다음을 선택합니다.
6. 사용자 액세스 제어 구성 페이지에서 다음을 선택합니다.
7. 프로비저닝 세부 정보 페이지에서 개발자 에디션을 선택합니다.
8. 생성을 선택하여 인덱스를 생성합니다.
9. 색인이 생성될 때까지 기다리세요. Amazon Kendra 인덱스의 하드웨어를 프로비저닝합니다. 이 작업에는 다소 시간이 걸릴 수 있습니다.

2단계: 인덱스에 데이터 소스 추가하기(콘솔)

1. 문서에 연결하고 문서를 색인화하는 데 사용할 수 Amazon Kendra 있는 데이터 소스를 확인하세요.
2. 탐색 창에서 데이터 소스를 선택한 다음 선택한 데이터 소스에 대한 데이터 소스 추가를 선택합니다.
3. 단계에 따라 데이터 소스를 구성합니다.

3단계: 인덱스 검색하기(콘솔)

1. 탐색 창에서 인덱스 검색 옵션을 선택합니다.
2. 인덱스에 적합한 검색어를 입력합니다. 가장 많이 반환된 결과 및 가장 많이 반환된 문서 결과가 표시됩니다.

시작하기(AWS CLI)

다음 절차는 를 사용하여 Amazon Kendra 색인을 만드는 방법을 보여줍니다 AWS CLI. 이 절차는 데이터 소스, 인덱스를 만들고 인덱스에서 쿼리를 실행합니다.

Amazon Kendra 인덱스 (CLI) 를 만들려면

1. 필수 조건을 수행합니다.
2. 다음 명령을 입력하여 인덱스를 생성합니다.

```
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. 인덱스가 생성될 때까지 Amazon Kendra 기다려 주세요. 다음 명령을 사용하여 진행률을 확인합니다. 상태 필드가 ACTIVE이면 다음 단계로 계속 진행합니다.

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

4. 명령 프롬프트에 다음 명령을 입력하여 데이터 소스를 생성합니다.

```
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"}}'
```

템플릿 스키마를 사용하여 데이터 소스에 연결하는 경우 템플릿 스키마를 구성합니다.

```
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. 데이터 소스를 만드는 데 시간이 Amazon Kendra 좀 걸릴 수 있습니다. 진행률을 확인하려면 다음 명령을 입력합니다. 상태가 ACTIVE이면 다음 단계로 계속 진행합니다.

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

6. 다음 명령을 입력하여 데이터 소스를 동기화합니다.

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

7. Amazon Kendra 데이터 원본을 인덱싱합니다. 소요 시간은 문서 수에 따라 달라집니다. 다음 명령을 사용하여 동기화 작업 상태를 확인할 수 있습니다. 상태가 ACTIVE이면 다음 단계로 계속 진행합니다.

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

8. 다음 명령을 입력하여 쿼리를 수행합니다.

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

검색 결과는 JSON 형식으로 표시됩니다.

시작하기(AWS SDK for Python (Boto3))

다음 프로그램은 Python Amazon Kendra 프로그램에서 사용하는 예입니다. 프로그램이 실행하는 작업은 다음과 같습니다.

1. CreateIndex작업을 사용하여 새 인덱스를 만듭니다.
2. 인덱스 생성이 완료될 때까지 기다립니다. DescribeIndex작업을 사용하여 인덱스 상태를 모니터링합니다.
3. 인덱스가 활성화되면 CreateDataSource작업을 사용하여 데이터 소스를 만듭니다.
4. 데이터 소스 생성이 완료될 때까지 기다립니다. DescribeDataSource작업을 사용하여 데이터 원본의 상태를 모니터링합니다.
5. 데이터 원본이 활성 상태이면 StartDataSourceSyncJob작업을 사용하여 색인을 데이터 원본의 내용과 동기화합니다.

```
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.")
```

시작하기(AWS SDK for Java)

다음 프로그램은 Java Amazon Kendra 프로그램에서 사용하는 예제입니다. 프로그램이 실행하는 작업은 다음과 같습니다.

1. CreateIndex작업을 사용하여 새 인덱스를 만듭니다.
2. 인덱스 생성이 완료될 때까지 기다립니다. DescribeIndex작업을 사용하여 인덱스 상태를 모니터링합니다.
3. 인덱스가 활성화되면 CreateDataSource작업을 사용하여 데이터 소스를 만듭니다.
4. 데이터 소스 생성이 완료될 때까지 기다립니다. DescribeDataSource작업을 사용하여 데이터 원본의 상태를 모니터링합니다.
5. 데이터 원본이 활성 상태이면 StartDataSourceSyncJob작업을 사용하여 색인을 데이터 원본의 내용과 동기화합니다.

```
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");
    }
}
```

Amazon S3 데이터 소스 시작하기(콘솔)

Amazon Kendra 콘솔을 사용하여 Amazon S3 버킷을 데이터 스토어로 사용하여 시작할 수 있습니다. 콘솔을 사용할 때 버킷의 내용을 인덱싱하는 데 필요한 모든 연결 정보를 지정합니다. 자세한 내용은 Amazon S3 섹션을 참조하세요.

다음 절차에 따라 기본 구성을 사용하여 기본 S3 버킷 데이터 소스를 생성합니다. 이 절차에서는 Amazon Kendra 콘솔 시작하기의 1단계에 따라 이미 인덱스를 생성한 것으로 가정합니다.

Amazon Kendra 콘솔을 사용하여 S3 버킷 데이터 소스를 생성하려면

1. AWS Management Console에 로그인하고 https://console.aws.amazon.com/kendra/home에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 목록에서 데이터 소스를 추가할 인덱스를 선택합니다.
3. 데이터 소스 추가를 선택합니다.
4. 데이터 소스 커넥터 목록에서 Amazon S3를 선택합니다.
5. 속성 정의 페이지에서 데이터 소스의 이름과 설명(선택 사항)을 지정합니다. 태그 필드는 비워 둡니다. 다음을 선택하여 계속 진행합니다.

6. 데이터 소스 위치 입력 필드에 문서가 포함된 S3 버킷의 이름을 입력합니다. 이름을 직접 입력하거나 찾아보기를 선택하여 이름을 찾아볼 수 있습니다. 버킷은 인덱스와 동일한 리전에 있어야 합니다.
7. IAM 역할에서 새 역할 생성을 선택한 다음, 역할 이름을 입력합니다. 자세한 내용은 Amazon S3 데이터 소스에 대한 IAM 역할을 참조하세요.
8. 동기화 실행 일정 설정 섹션에서 온디맨드 실행을 선택합니다.
9. 다음을 선택하여 계속 진행합니다.
10. 검토 및 생성 페이지에서 S3 데이터 소스의 세부 정보를 검토합니다. 변경하려면 변경하려는 항목 옆에 있는 편집 버튼을 선택합니다. 선택에 만족하면 생성을 선택하여 S3 데이터 소스를 생성합니다.

생성을 선택한 후 Amazon Kendra가 데이터 소스 생성을 시작합니다. 데이터 소스가 생성되는 데 몇 분 정도 걸립니다. 작업이 끝나면 데이터 소스의 상태가 생성 중에서 활성으로 변경됩니다.

데이터 소스를 만든 후에는 Amazon Kendra 인덱스를 데이터소스와 동기화해야 합니다. 지금 동기화를 선택하여 동기화 프로세스를 시작합니다. 문서 수와 크기에 따라 데이터 소스를 동기화하는 데 몇 분에서 몇 시간이 걸릴 수 있습니다.

MySQL 데이터베이스 데이터 소스 시작하기(콘솔)

Amazon Kendra 콘솔을 사용하여 MySQL 데이터베이스를 데이터 소스로 사용하여 시작할 수 있습니다. 콘솔을 사용할 때 MySQL 데이터베이스의 내용을 인덱싱하는 데 필요한 연결 정보를 지정합니다. 자세한 내용은 데이터베이스 데이터 소스 사용을 참조하세요.

먼저 MySQL 데이터베이스를 만든 다음 데이터베이스의 데이터 소스를 만들 수 있습니다.

다음 절차에 따라 기본 MySQL 데이터베이스를 생성합니다. 이 절차에서는 Amazon Kendra 콘솔 시작하기의 1단계에 따라 이미 인덱스를 생성한 것으로 가정합니다.

MySQL 데이터베이스 생성

1. AWS Management Console에 로그인한 후 https://console.aws.amazon.com/rds/에서 Amazon RDS 콘솔을 엽니다.
2. 탐색 창에서 서브넷 그룹을 선택한 다음 DB 서브넷 그룹 생성을 선택합니다.
3. 그룹 이름을 지정하고 Virtual Private Cloud(VPC)를 선택합니다. VPC 구성에 대한 자세한 내용은 VPC를 사용하기 위한 Amazon Kendra 구성을 참조하세요.

4. VPC의 프라이빗 서브넷을 추가합니다. 프라이빗 서브넷은 NAT에 연결되지 않은 서브넷입니다. 생성을 선택합니다.
5. 탐색 창에서 데이터베이스를 선택하고 데이터베이스 생성을 선택합니다.
6. 다음 파라미터를 사용하여 데이터베이스를 생성합니다. 다른 모든 파라미터는 기본값으로 유지합니다.

   - 엔진 옵션 - MySQL
   - 템플릿 - 프리 티어
   - 보안 인증 설정 - 암호를 입력하고 확인합니다.
   - 연결에서 추가 연결 구성을 선택합니다. 다음과 같이 변경합니다.
     - 서브넷 그룹 - 4단계에서 생성한 서브넷 그룹을 선택합니다.
     - VPC 보안 그룹 - VPC에서 만든 인바운드 규칙과 아웃바운드 규칙을 모두 포함하는 그룹을 선택합니다. 예: **`DataSourceSecurityGroup`**. VPC 구성에 대한 자세한 내용은 VPC를 사용하기 위한 Amazon Kendra 구성을 참조하세요.
   - 추가 구성 아래 초기 데이터베이스 이름에 **`content`**를 입력합니다.
7. 데이터베이스 생성을 선택합니다.
8. 데이터베이스 목록에서 새 데이터베이스를 선택합니다. 데이터베이스 엔드포인트를 기록해 둡니다.
9. 데이터베이스를 생성한 후 문서를 보관할 테이블을 생성해야 합니다. 테이블을 만드는 것은 이 지침의 범위를 벗어납니다. 테이블을 생성할 때 다음 사항에 유의하세요.

   - 데이터베이스 이름 - **`content`**
   - 테이블 이름 - **`documents`**
   - 열 - **`ID`**, **`Title`**, **`Body`** 및 **`LastUpdate`**. 원하는 경우 열을 더 포함할 수 있습니다.

MySQL 데이터베이스를 만들었으니 데이터베이스의 데이터 소스를 만들 수 있습니다.

MySQL 데이터 소스 생성

1. AWS Management Console에 로그인한 후 https://console.aws.amazon.com/kendra/home에서 Amazon Kendra 콘솔을 엽니다.
2. 탐색 창에서 인덱스를 선택한 후 해당 인덱스를 선택합니다.
3. 데이터 소스 추가를 선택한 다음 Amazon RDS를 선택합니다.
4. 새 역할의 이름과 설명을 입력하고, 다음을 선택합니다.

5. MySQL을 선택합니다.
6. 연결 액세스에서 다음 정보를 입력하세요.
   - 엔드포인트 - 이전에 만든 데이터베이스의 엔드포인트입니다.
   - 포트 - 데이터베이스의 포트 번호입니다. MySQL의 기본 포트는 3306입니다.
   - 인증 유형 - 새로 만들기를 선택합니다.
   - 새 비밀 컨테이너 이름 - 데이터베이스 보안 인증을 위한 Secrets Manager 컨테이너의 이름입니다.
   - 사용자 이름 - 데이터베이스에 대한 관리 액세스 권한이 있는 사용자의 이름입니다.
   - 암호 - 사용자의 암호이며, 다음으로 인증 저장을 선택합니다.
   - 데이터베이스 이름 -**`content`**.
   - 테이블 이름 – **`documents`**.
   - IAM 역할 - 새 역할 생성을 선택한 다음, 역할의 이름을 입력합니다.
7. 열 구성에 다음을 입력합니다.
   - 문서 ID 열 이름 - **`ID`**
   - 문서 제목 열 이름 - **`Title`**
   - 문서 데이터 열 이름 - **`Body`**
8. 열 변경 감지에 다음을 입력합니다.
   - 변경 감지 열 - **`LastUpdate`**
9. VPC 및 보안 그룹 구성에 다음을 제공합니다.
   - Virtual Private Cloud(VPC)에서 VPC를 선택합니다.
   - 서브넷에서, VPC에서 생성한 프라이빗 서브넷을 선택합니다.
   - VPC 보안 그룹에서, MySQL 데이터베이스용 VPC에서 만든 인바운드 규칙과 아웃바운드 규칙을 모두 포함하는 보안 그룹을 선택합니다. 예: **`DataSourceSecurityGroup`**.
10. 동기화 실행 일정 설정에서 온디맨드 실행을 선택하고 다음을 선택합니다.
11. 데이터 소스 필드 매핑에서 다음을 선택합니다.
12. 데이터 소스의 구성을 검토하여 올바른지 확인합니다. 모든 내용이 정확하다고 생각되면 생성을 선택합니다.

AWS IAM Identity Center ID 소스 시작하기 (콘솔)

AWS IAM Identity Center ID 소스에는 사용자 및 그룹에 대한 정보가 포함됩니다. 이는 사용자 또는 해당 그룹의 문서 액세스 권한을 기반으로 여러 사용자에 대한 검색 결과를 필터링하는 사용자 컨텍스트 Amazon Kendra 필터링을 설정하는 데 유용합니다.

IAM Identity Center ID 소스를 생성하려면 AWS Organizations에서 IAM Identity Center를 활성화하고 조직을 생성해야 합니다. IAM Identity Center를 활성화하고 처음으로 조직을 생성하면 자동으로 Identity Center 디렉터리가 ID 소스로 기본 설정됩니다. Active Directory(Amazon 관리형 또는 자체 관리형) 또는 외부 ID 공급자로 ID 소스로 변경할 수 있습니다. 이에 대한 올바른 지침을 따라야 합니다. IAM Identity Center ID 소스 변경을 참조하세요. 조직당 Identity Center 한 개만 가질 수 있습니다.

사용자와 그룹에 서로 다른 수준의 문서 액세스 권한을 할당하려면 문서를 인덱스로 수집할 때 액세스 제어 목록에 사용자와 그룹을 포함해야 합니다. 이렇게 하면 사용자와 그룹이 액세스 수준에 따라 문서를 검색할 수 있습니다. Amazon Kendra 쿼리를 실행할 때 사용자 ID는 IAM Identity Center의 사용자 이름과 정확히 일치해야 합니다.

또한 IAM Identity Center를 사용하는 데 필요한 권한을 부여해야 합니다. Amazon Kendra 자세한 정보는 IAM Identity Center에 대한 IAM 역할을 참조하세요.

IAM Identity Center ID 소스 설정하기

1. IAM Identity Center 콘솔을 엽니다.
2. IAM ID 센터 활성화를 선택한 다음 조직 생성을 AWS 선택합니다.

   Identity Center 디렉터리는 기본적으로 생성되며 조직과 연결된 이메일 주소를 확인하기 위한 이메일이 사용자에게 전송됩니다.
3. AWS 조직에 그룹을 추가하려면 탐색 창에서 그룹을 선택합니다.
4. 그룹 페이지에서 대화 상자에 그룹 이름과 설명을 입력하고 그룹 생성을 선택합니다. 생성을 선택하세요.
5. 조직에 사용자를 추가하려면 탐색 창에서 사용자를 선택합니다.
6. 사용자 페이지에서 사용자 추가를 선택합니다. 사용자 세부 정보에서 모든 필수 필드를 지정합니다. 암호에 대해 사용자에게 이메일 보내기를 선택합니다. 다음을 선택합니다.
7. 그룹에 사용자를 추가하려면 그룹을 선택하고 그룹을 선택합니다.
8. 세부 정보 페이지에서 그룹 멤버의 사용자 추가를 선택합니다.
9. 그룹에 사용자 추가 페이지에서 그룹 멤버로 추가할 사용자를 선택합니다. 그룹에 추가할 사용자를 여러 명 선택할 수 있습니다.

10. 사용자 및 그룹 목록을 IAM Identity Center와 동기화하려면 ID 소스를 Active Directory 또는 외부 ID 공급자로 변경합니다.

    Identity Center 디렉터리는 기본 ID 소스이며 공급자가 관리하는 자체 목록이 없는 경우 이 소스를 사용하여 사용자 및 그룹을 수동으로 추가해야 합니다. ID 소스를 변경하려면 이에 대한 올바른 지침을 따라야 합니다. IAM Identity Center ID 소스 변경을 참조하세요.

> **Note**
>
> Active Directory 또는 외부 ID 공급자를 ID 소스로 사용하는 경우, 교차 도메인 ID 관리 시스템(SCIM) 프로토콜을 지정할 때 사용자의 이메일 주소를 IAM Identity Center 사용자 이름에 매핑해야 합니다. 자세한 내용은 IAM Identity Center 활성화용 SCIM에 대한 IAM Identity Center 설명서를 참조하세요.

IAM Identity Center ID 소스를 설정한 후에는 인덱스를 생성하거나 편집할 때 콘솔에서 이를 활성화할 수 있습니다. 인덱스 설정의 사용자 액세스 제어로 이동하여 IAM Identity Center에서 사용자 그룹 정보를 가져올 수 있도록 설정을 편집하세요.

UserGroupResolutionConfiguration객체를 사용하여 IAM ID 센터를 활성화할 수도 있습니다. `UserGroupResolutionMode`AWS_SSOas를 제공하고`sso:ListDirectoryAssociations`, `sso-directory:SearchUserssso-directory:ListGroupsForUser`, `sso-directory:DescribeGroups` 를 호출할 권한을 부여하는 IAM 역할을 생성합니다.

> **Warning**
>
> Amazon Kendra 현재 AWS 조직 구성원 `UserGroupResolutionConfiguration` 계정과 함께 IAM Identity Center ID 소스를 사용하는 것은 지원되지 않습니다. `UserGroupResolutionConfiguration`을 사용하려면 조직의 관리 계정에서 인덱스를 생성해야 합니다.

다음은 사용자 컨텍스트에 따라 검색 결과를 필터링하기 위해 `UserGroupResolutionConfiguration` 및 사용자 액세스 제어를 사용하여 데이터 소스를 설정하는 방법에 대한 개요입니다. 이는 인덱스에 대한 인덱스와 IAM 역할을 이미 생성했다고 가정합니다. API를 사용하여 색인을 만들고 IAM 역할을 제공합니다. CreateIndex

`UserGroupResolutionConfiguration` 및 사용자 컨텍스트 필터링을 사용하여 데이터 소스 설정

1. IAM Identity Center ID 소스에 액세스할 수 있는 권한을 부여하는 IAM 역할을 생성합니다.
2. 모드를 `UserGroupResolutionConfiguration`로 설정하고 IAM Identity Center를 사용하도록 인덱스를 UpdateIndex업데이트하도록 호출하여 구성합니다. `AWS_SSO`
3. 토큰 기반 사용자 액세스 제어를 사용하여 사용자 컨텍스트에 따라 검색 결과를 필터링하려면 `USER_TOKEN` 호출할 때 UserContextPolicy로 설정하십시오. `UpdateIndex` 그렇지 않으면 대부분의 Amazon Kendra 데이터 소스 커넥터에 대해 각 문서의 액세스 제어 목록을 크롤링합니다. `UserContext`에서 사용자 및 그룹 정보를 제공하여 쿼리 API에서 사용자 컨텍스트에 따라 검색 결과를 필터링할 수도 있습니다. 를 사용하여 사용자를 그룹에 매핑할 수도 PutPrincipalMapping있으므로 쿼리를 실행할 때 사용자 ID만 제공하면 됩니다.
4. 데이터 소스에 액세스할 수 있는 권한을 부여하는 IAM 역할을 만듭니다.
5. 데이터 소스를 구성합니다. 데이터 소스에 연결하는 데 필요한 연결 정보를 제공해야 합니다.
6. CreateDataSourceAPI를 사용하여 데이터 소스를 생성합니다. `TemplateConfiguration`, 인덱스의 ID, 데이터 소스의 IAM 역할, 데이터 소스 유형을 포함하는 `DataSourceConfiguration` 객체를 제공하고 데이터 소스에 이름을 지정합니다. 데이터 소스를 업데이트할 수도 있습니다.

IAM Identity Center ID 소스 변경하기

> **Warning**
>
> IAM Identity Center 설정에서 ID 소스를 변경하면 사용자 및 그룹 정보의 보존에 영향을 미칠 수 있습니다. 이를 안전하게 수행하려면 ID 소스 변경 고려 사항을 검토하는 것이 좋습니다. ID 소스를 변경하면 새 ID 소스 ID가 생성됩니다. 모드를 in으로 설정하기 전에 올바른 ID를 사용하고 있는지 확인하세요 UserGroupResolutionConfiguration. `AWS_SSO`

IAM Identity Center ID 소스 변경하기

1. IAM Identity Center> 콘솔을 엽니다.
2. 설정을 선택합니다.
3. 설정 페이지의 ID 소스에서 변경을 선택합니다.
4. ID 소스 변경 페이지에서 원하는 ID 소스를 선택하고 다음을 선택합니다.

인덱스 생성

콘솔을 사용하거나 CreateIndexAPI를 호출하여 색인을 생성할 수 있습니다. API와 함께 AWS Command Line Interface (AWS CLI) 또는 SDK를 사용할 수 있습니다. 인덱스를 만든 후 인덱스에 직접 또는 데이터 소스에서 문서를 추가할 수 있습니다.

인덱스를 생성하려면 인덱스가 액세스할 수 있는 () 역할의 Amazon 리소스 이름 AWS Identity and Access Management (ARNIAM) 을 제공해야 합니다. CloudWatch 자세한 내용은 인덱스에 대한 IAM 역할을 참조하세요.

다음 탭은 AWS Management Console,,, Python 및 Java SDK를 사용하기 위한 코드 예제를 사용하여 색인을 생성하는 절차를 제공합니다. AWS CLI

Console

인덱스를 만들려면

1. AWS 관리 콘솔에 로그인하고 https://console.aws.amazon.com/kendra/ 에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 섹션에서 인덱스 생성을 선택합니다.
3. 인덱스 세부정보 지정 페이지에서 인덱스 이름과 설명을 입력합니다.
4. IAM 역할에서 IAM 역할을 제공하십시오. 역할을 찾으려면 계정에서 “kendra”라는 단어가 포함된 역할 중에서 선택하거나 다른 역할의 이름을 입력하세요. 역할에 필요한 권한에 대한 자세한 내용은 인덱스에 대한 IAM 역할을 참조하세요.
5. 다음을 선택합니다.
6. 사용자 액세스 제어 구성 페이지에서 다음을 선택합니다. 인덱스를 만든 후 액세스 제어에 토큰을 사용하도록 인덱스를 업데이트할 수 있습니다. 자세한 내용은 문서에 대한 액세스 제어를 참조하세요.
7. 프로비저닝 세부 정보 페이지에서 생성을 선택합니다.
8. 인덱스를 생성하는 데 시간이 걸릴 수 있습니다. 인덱스 목록을 확인하여 인덱스 생성 진행 상황을 확인하세요. 인덱스 상태가 `ACTIVE`가 되면 인덱스를 사용할 준비가 된 것입니다.

AWS CLI

색인을 만들려면

1. 다음 명령을 사용하여 인덱스를 생성합니다. 는 작업을 실행할 Amazon Kendra 수 있는 IAM 역할의 Amazon 리소스 이름 (ARN) `role-arn` 이어야 합니다. 자세한 내용은 IAM 역할을 참조하세요.

   이 명령은 Linux 및 macOS용으로 형식이 지정됩니다. Windows를 사용하는 경우 Unix 줄 연속 문자(\)를 캐럿(^)으로 바꿉니다.

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

2. 인덱스를 생성하는 데 시간이 걸릴 수 있습니다. 인덱스 상태를 확인하려면 `create-index`에서 반환한 인덱스 ID를 다음 명령과 함께 사용하세요. 인덱스 상태가 `ACTIVE`가 되면 인덱스를 사용할 준비가 된 것입니다.

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

Python

인덱스를 만들려면

- 다음 코드 예제에 다음 변수의 값을 입력합니다.
  - `description` - 생성 중인 인덱스에 대한 설명입니다. 이는 선택 사항입니다.
  - `index_name` - 생성 중인 인덱스의 이름입니다.
  - `role_arn`—API를 실행할 수 Amazon Kendra 있는 역할의 Amazon 리소스 이름 (ARN). 자세한 내용은 IAM 역할을 참조하세요.

```
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

인덱스를 만들려면

- 다음 코드 예제에 다음 변수의 값을 입력합니다.
  - `description` - 생성 중인 인덱스에 대한 설명입니다. 이는 선택 사항입니다.
  - `index_name` - 생성 중인 인덱스의 이름입니다.
  - `role_arn`—API를 실행할 수 Amazon Kendra 있는 역할의 Amazon 리소스 이름 (ARN). 자세한 내용은 IAM 역할을 참조하세요.

```
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.");
    }
}
```

인덱스를 만든 후 인덱스에 문서를 추가합니다. 직접 추가하거나 정기적으로 인덱스를 업데이트하는 데이터 소스를 만들 수 있습니다.

주제

- 일괄 업로드를 사용하여 인덱스에 직접 문서 추가
- 인덱스에 자주 묻는 질문(FAQ) 추가
- 사용자 지정 문서 필드 만들기
- 토큰을 사용하여 문서에 대한 사용자 액세스 제어

일괄 업로드를 사용하여 인덱스에 직접 문서 추가

API를 사용하여 색인에 직접 문서를 추가할 수 있습니다. BatchPutDocument 콘솔을 사용하여 문서를 직접 추가할 수 없습니다. 콘솔을 사용하는 경우 데이터 소스에 연결하여 인덱스에 문서를 추가합니다.

문서는 S3 버킷에서 추가하거나 이진 데이터로 제공할 수 있습니다. 에서 지원하는 문서 유형 목록은 문서 유형을 Amazon Kendra 참조하십시오.

BatchPutDocument를 사용하여 인덱스에 문서를 추가하는 것은 비동기식 작업입니다. API를 호출한 후 BatchPutDocument API를 사용하여 문서 색인 생성 진행 상황을 모니터링합니다. BatchGetDocumentStatus 문서 ID 목록을 사용하여 BatchGetDocumentStatus API를 호출하면 문서 상태가 반환됩니다. 문서 상태가 INDEXED 또는 FAILED이면 문서 처리가 완료된 것입니다. 상태가 FAILED인 경우 BatchGetDocumentStatus API는 문서를 인덱싱할 수 없는 이유를 반환합니다.

문서 수집 프로세스 중에 콘텐츠와 문서 메타데이터나 필드 또는 속성을 변경하려면 Amazon Kendra 사용자 지정 문서 보강을 참조하세요. 사용자 지정 데이터 소스를 사용하려는 경우 BatchPutDocument API를 사용하여 제출하는 각 문서에는 데이터 소스 ID와 실행 ID가 속성 또는 필드로 필요합니다. 자세한 내용은 사용자 지정 데이터 소스의 필수 속성을 참조하세요.

> **Note**
>
> 각 문서 ID는 색인별로 고유해야 합니다. 고유한 ID로 문서를 인덱싱하는 데이터 소스를 만든 다음 BatchPutDocument API를 사용하여 동일한 문서를 인덱싱할 수 없으며 그 반대의 경우도 마찬가지입니다. 데이터 소스를 삭제한 다음 BatchPutDocument API를 사용하여 동일한 문서를 인덱싱하거나 그 반대로 할 수 있습니다. 동일한 문서 세트에 BatchPutDocument 및 BatchDeleteDocument API를 Amazon Kendra 데이터 소스 커넥터와 함께 사용하면 데이터와 일치하지 않을 수 있습니다. 대신, Amazon Kendra 사용자 지정 데이터 소스 커넥터를 사용하는 것이 좋습니다.

다음 개발자 가이드 문서는 문서를 인덱스에 직접 추가하는 방법을 보여줍니다.

주제

- BatchPutDocumentAPI로 문서 추가
- S3 버킷에서 문서 추가

BatchPutDocumentAPI로 문서 추가

다음 예제는 호출을 BatchPutDocument통해 색인에 텍스트 덩어리를 추가합니다. BatchPutDocumentAPI를 사용하여 색인에 직접 문서를 추가할 수 있습니다. 에서 지원하는 문서 유형 목록은 문서 유형을 Amazon Kendra 참조하십시오.

AWS CLI 및 SDK를 사용하여 색인을 만드는 예제는 색인 만들기를 참조하십시오. CLI 및 SDK를 설정하려면 Amazon Kendra 설정을 참조하세요.

> **Note**
>
> 인덱스에 추가된 파일은 UTF-8 인코딩된 바이트 스트림에 있어야 합니다.

다음 예제에서는 UTF-8 인코딩 형식의 텍스트가 인덱스에 추가됩니다.

CLI

AWS Command Line Interface에서는 다음 명령을 사용합니다. 이 명령은 Linux 및 macOS용으로 형식이 지정됩니다. Windows를 사용하는 경우 Unix 줄 연속 문자(\)를 캐럿(^)으로 바꿉니다.

```
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));
    }
}
```

S3 버킷에서 문서 추가

BatchPutDocumentAPI를 사용하여 Amazon S3 버킷에서 색인에 직접 문서를 추가할 수 있습니다. 동일한 호출에 문서를 10개까지 추가할 수 있습니다. S3 버킷을 사용하는 경우 문서가 포함된 버킷에 액세스할 수 있는 권한이 있는 IAM 역할을 제공해야 합니다. `RoleArn` 파라미터에 역할을 지정합니다.

BatchPutDocumentAPI를 사용하여 Amazon S3 버킷에서 문서를 추가하는 것은 일회성 작업입니다. 인덱스를 버킷 콘텐츠와 동기화된 상태로 유지하려면 Amazon S3 데이터 소스를 생성하세요. 자세한 내용은 Amazon S3 데이터 소스를 참조하세요.

AWS CLI 및 SDK를 사용하여 인덱스를 만드는 예제는 인덱스 만들기를 참조하십시오. CLI 및 SDK를 설정하려면 Amazon Kendra 설정을 참조하세요. S3 버킷 생성에 대한 자세한 내용은 Amazon Simple Storage Service 설명서를 참조하세요.

다음 예제에서는 `BatchPutDocument` API를 사용하여 두 개의 Microsoft Word 문서를 인덱스에 추가합니다.

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));
    }
}
```

인덱스에 자주 묻는 질문(FAQ) 추가

콘솔 또는 API를 사용하여 자주 묻는 질문 (FAQ) 을 색인에 직접 추가할 수 있습니다. CreateFaq FAQ를 인덱스에 추가하는 것은 비동기식 작업입니다. FAQ 데이터는 버킷에 저장하는 파일에 저장합니다. Amazon Simple Storage Service CSV 또는 JSON 파일을 FAQ 입력으로 사용할 수 있습니다.

- 기본 CSV - 각 행에 질문, 답변 및 선택적 소스 URI가 포함된 CSV 파일입니다.
- 사용자 지정 CSV - FAQ 응답을 패싯, 표시 또는 정렬하는 데 사용할 수 있는 사용자 지정 필드/속성에 대한 질문, 답변 및 헤더가 포함된 CSV 파일입니다. 액세스 제어 필드를 정의하여 FAQ 응답을 볼 수 있는 특정 사용자 및 그룹으로 FAQ 응답을 제한할 수도 있습니다.
- JSON - 질문, 답변, 사용자 지정 필드/속성을 포함하는 JSON 파일로, FAQ 응답을 패싯, 표시 또는 정렬하는 데 사용할 수 있습니다. 액세스 제어 필드를 정의하여 FAQ 응답을 볼 수 있는 특정 사용자 및 그룹으로 FAQ 응답을 제한할 수도 있습니다.

예를 들어, 다음은 미국 워싱턴주 스포캔과 미국 미주리주 마운틴뷰에 있는 무료 진료소에 대한 질문에 대한 답변을 제공하는 기본 CSV 파일입니다.

```
How many free clinics are in Spokane WA?, 13
How many free clinics are there in Mountain View Missouri?, 7
```

> **Note**
>
> FAQ 파일은 UTF-8 인코딩 파일이어야 합니다.

주제

FAQ 파일의 인덱스 필드 생성

사용자 지정 CSV 또는 JSON 파일을 입력용으로 사용하는 경우 FAQ 질문에 대한 사용자 지정 필드를 선언할 수 있습니다. 예를 들어 각 FAQ 질문에 비즈니스 부서를 할당하는 사용자 지정 필드를 만들 수 있습니다. 예를 들어 FAQ가 응답으로 반환되면 부서를 패싯으로 사용하여 검색 범위를 “HR” 또는 “재무”로만 좁힐 수 있습니다.

사용자 지정 필드는 인덱스 필드에 매핑되어야 합니다. 콘솔에서 패싯 정의 페이지를 사용하여 인덱스 필드를 만들 수 있습니다. API를 사용할 때는 먼저 API를 사용하여 인덱스 필드를 만들어야 합니다. UpdateIndex

FAQ 파일의 필드/속성 유형은 관련 인덱스 필드의 유형과 일치해야 합니다. 예를 들어, “부서” 필드는 STRING_LIST 유형 필드입니다. 따라서 FAQ 파일에 부서 필드 값을 문자열 목록으로 제공해야 합니다. 콘솔의 Facet 정의 페이지를 사용하거나 DescribeIndexAPI를 사용하여 인덱스 필드의 유형을 확인할 수 있습니다.

사용자 지정 속성에 매핑되는 인덱스 필드를 만들 때 해당 필드를 표시 가능, 패싯 가능 또는 정렬 가능으로 표시할 수 있습니다. 사용자 지정 속성을 검색 가능하게 만들 수는 없습니다.

사용자 지정 속성 외에도 사용자 지정 CSV 또는 JSON 파일에서 Amazon Kendra 예약된 필드 또는 공통 필드를 사용할 수 있습니다. 자세한 내용은 문서 속성 또는 필드를 참조하세요.

기본 CSV 파일

FAQ에 간단한 구조를 사용하려는 경우 기본 CSV 파일을 사용하세요. 기본 CSV 파일의 각 행에는 질문, 답변, 추가 정보가 포함된 문서를 가리키는 선택적 소스 URI 등 2개 또는 3개의 필드가 있습니다.

파일의 내용은 RFC 4180 CSV(쉼표로 구분된 값) 파일의 공통 형식 및 MIME 유형을 따라야 합니다.

다음은 기본 CSV 형식의 FAQ 파일입니다.

```
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
```

사용자 지정 CSV 파일

FAQ 질문에 사용자 지정 필드/속성을 추가하려면 사용자 지정 CSV 파일을 사용하세요. 사용자 지정 CSV 파일의 경우 CSV 파일의 헤더 행을 사용하여 추가 속성을 정의합니다.

CSV 파일에는 다음 두 개의 필수 필드가 포함되어야 합니다.

- _question - 자주 묻는 질문
- _answer - 자주 묻는 질문에 대한 답변

파일에는 Amazon Kendra 예약된 필드와 사용자 지정 필드가 모두 포함될 수 있습니다. 다음은 CSV 파일의 예제입니다.

```
_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
```

사용자 지정 파일의 내용은 RFC 4180 CSV(쉼표로 구분된 값) 파일의 공통 형식 및 MIME 유형을 따라야 합니다.

다음은 사용자 지정 필드 유형 목록입니다.

- 날짜 - ISO 8601로 인코딩된 날짜와 시간 값입니다.

  예를 들어, 2012-03-25T12:30:10+01:00은 중부 유럽 시간대로 2012년 3월 25일 오후 12시 30분 10초에 대한 ISO 8601 날짜/시간 형식입니다.
- Long - 숫자(예: 1234).
- 문자열 - 문자열 값. 문자열에 쉼표가 포함된 경우 전체 값을 큰따옴표 (“) 로 묶으세요(예: `"custom attribute, and more"`).
- 문자열 목록 - 문자열 값의 목록입니다. 따옴표(“)로 묶인 쉼표로 구분된 목록으로 값을 나열합니다(예: `"item1, item2, item3"`). 목록에 항목이 하나뿐인 경우 따옴표를 생략할 수 있습니다(예: `item1`).

사용자 지정 CSV 파일에는 사용자 액세스 제어 필드가 포함될 수 있습니다. 이 필드를 사용하여 FAQ에 대한 액세스를 특정 사용자 및 그룹으로 제한할 수 있습니다. 사용자 컨텍스트를 기준으로 필터링하려면 사용자가 쿼리에 사용자 및 그룹 정보를 제공해야 합니다. 그렇지 않으면 모든 관련 FAQ가 반환됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

다음은 FAQ의 사용자 컨텍스트 필터 목록입니다.

- `_acl_user_allow` - 허용 목록에 있는 사용자는 쿼리 응답에서 FAQ를 볼 수 있습니다. 이 FAQ는 다른 사용자에게 반환되지 않습니다.
- `_acl_user_deny` - 거부 목록에 있는 사용자는 쿼리 응답에서 FAQ를 볼 수 없습니다. 이 FAQ는 쿼리와 관련된 경우 다른 모든 사용자에게 반환됩니다.
- `_acl_group_allow` - 허용된 그룹의 구성원인 사용자는 쿼리 응답에서 FAQ를 볼 수 있습니다. 다른 그룹의 구성원인 사용자에게는 FAQ가 반환되지 않습니다.

- _acl_group_deny - 거부된 그룹의 구성원인 사용자는 쿼리 응답에서 FAQ를 볼 수 없습니다. 이 FAQ는 쿼리와 관련된 경우 다른 그룹들에 반환됩니다.

따옴표로 묶인 쉼표로 구분된 목록으로 허용 및 거부 목록의 값을 제공합니다(예: "user1,user2,user3"). 사용자나 그룹을 허용 목록 또는 거부 목록에 포함할 수 있지만, 동일한 사용자를 둘 모두에 포함할 수는 없습니다(개인은 허용되지만 그룹은 거부되는 경우). 둘 모두에 사용자나 그룹을 포함하면 오류가 발생합니다.

다음은 사용자 컨텍스트 정보가 포함된 사용자 지정 CSV 파일의 예입니다.

```
_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
```

JSON 파일

JSON 파일을 사용하여 인덱스에 대한 질문, 답변, 필드를 제공할 수 있습니다. Amazon Kendra 예약된 필드 또는 사용자 지정 필드를 FAQ에 추가할 수 있습니다.

다음은 JSON 파일에 대한 스키마입니다.

```
{
    "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
```

```
    ]
}
```

다음 예제 JSON 파일은 두 개의 FAQ 문서를 보여줍니다. 문서 중 하나에는 필수 질문과 답변만 있습니다. 다른 문서에는 추가 필드 및 사용자 컨텍스트 또는 액세스 제어 정보도 포함되어 있습니다.

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

다음은 사용자 지정 필드 유형 목록입니다.

- 날짜 - ISO 8601로 인코딩된 날짜와 시간 값이 있는 JSON 문자열 값입니다. 예를 들어, 2012-03-25T12:30:10+01:00은 중부 유럽 시간대로 2012년 3월 25일 오후 12시 30분 10초에 대한 ISO 8601 날짜/시간 형식입니다.

- Long - JSON 숫자 값(예: 1234)입니다.
- 문자열 - JSON 문자열 값(예: "custom attribute")입니다.
- 문자열 목록 - 문자열 값의 JSON 배열(예: ["item1,item2,item3"])입니다.

JSON 파일에는 사용자 액세스 제어 필드가 포함될 수 있습니다. 이 필드를 사용하여 FAQ에 대한 액세스를 특정 사용자 및 그룹으로 제한할 수 있습니다. 사용자 컨텍스트를 기준으로 필터링하려면 사용자가 쿼리에 사용자 및 그룹 정보를 제공해야 합니다. 그렇지 않으면 모든 관련 FAQ가 반환됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

사용자나 그룹을 허용 목록 또는 거부 목록에 포함할 수 있지만, 동일한 사용자를 둘 모두에 포함할 수는 없습니다(개인은 허용되지만 그룹은 거부되는 경우). 둘 모두에 사용자나 그룹을 포함하면 오류가 발생합니다.

다음은 JSON FAQ에 대한 사용자 액세스 제어를 포함하는 예입니다.

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

FAQ 파일 사용

FAQ 입력 파일을 S3 버킷에 저장한 후 콘솔 또는 CreateFaq API를 사용하여 질문과 답변을 인덱스에 추가합니다. FAQ를 업데이트하려면 FAQ를 삭제하고 다시 생성하세요. DeleteFaq API를 사용하여 FAQ를 삭제합니다.

원본 파일이 포함된 S3 버킷에 액세스할 수 있는 IAM 역할을 제공해야 합니다. 콘솔 또는 RoleArn 파라미터에 역할을 지정합니다. 다음은 FAQ 파일을 인덱스에 추가하는 예입니다.

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));

    }
}
```

영어 이외의 언어로 된 FAQ 파일

지원되는 언어로 FAQ를 인덱싱할 수 있습니다. Amazon Kendra 언어를 지정하지 않은 경우 기본적으로 FAQ를 영어로 인덱싱합니다. CreateFaq오퍼레이션을 호출할 때 언어 코드를 지정하거나 FAQ 메타데이터에 FAQ의 언어 코드를 필드로 포함할 수 있습니다. 메타데이터 필드에 지정된 FAQ의 메타데이터에 언어 코드가 없는 경우 `CreateFAQ` 작업을 호출할 때 지정된 언어 코드를 사용하여 FAQ를 인덱싱합니다. 콘솔에서 지원되는 언어로 FAQ 문서를 인덱싱하려면 FAQ로 이동하여 FAQ 추가를 선택합니다. 언어 드롭다운에서 언어를 선택합니다.

사용자 지정 문서 필드 만들기

Amazon Kendra 색인에서 문서의 사용자 지정 속성 또는 필드를 생성할 수 있습니다. 예를 들어, “HR”, “영업” 및 “제조” 값을 사용하여 “부서”라는 사용자 지정 필드 또는 속성을 만들 수 있습니다. 이러한 사용자 지정 필드 또는 속성을 Amazon Kendra 색인에 매핑하면 이를 사용하여 “HR” 부서 속성 등을 기준으로 문서를 포함하도록 검색 결과를 필터링할 수 있습니다.

사용자 지정 필드나 속성을 사용하려면 먼저 인덱스에 필드를 생성해야 합니다. 콘솔을 사용하여 데이터 소스 필드 매핑을 편집하여 사용자 지정 필드를 추가하거나 UpdateIndexAPI를 사용하여 인덱스 필드를 생성합니다. 필드를 생성한 후에는 필드 데이터 유형을 변경할 수 없습니다.

대부분의 데이터 소스에서는 외부 데이터 소스의 필드를 Amazon Kendra의 해당 필드에 매핑합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요. S3 데이터 소스의 경우 JSON 메타데이터 파일을 사용하여 사용자 지정 필드 또는 속성을 생성할 수 있습니다.

최대 500개의 사용자 지정 필드 또는 속성을 만들 수 있습니다.

Amazon Kendra 예약된 필드나 공통 필드를 사용할 수도 있습니다. 자세한 내용은 문서 속성 또는 필드를 참조하세요.

주제

- 사용자 지정 문서 필드 업데이트

사용자 지정 문서 필드 업데이트

`UpdateIndex` API를 사용하면 `DocumentMetadataConfigurationUpdates` 파라미터를 사용하여 사용자 지정 필드 또는 속성을 추가할 수 있습니다.

다음 JSON 예제는 `DocumentMetadataConfigurationUpdates`를 사용하여 “Department”라는 필드를 인덱스에 추가합니다.

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

다음 섹션에는 Amazon S3 데이터 소스의 BatchPutDocument및 를 사용하여 사용자 지정 속성 또는 필드를 추가하는 예제가 포함되어 있습니다.

주제

- BatchPutDocument API를 사용하여 사용자 지정 속성 또는 필드 추가
- Amazon S3 데이터 소스에 사용자 지정 속성 또는 필드 추가

BatchPutDocument API를 사용하여 사용자 지정 속성 또는 필드 추가

BatchPutDocumentAPI를 사용하여 색인에 문서를 추가할 때는 사용자 지정 필드 또는 속성을 색인의 일부로 지정합니다`Attributes`. API를 호출할 때 여러 필드 또는 속성을 추가할 수 있습니다. 최대 500개의 사용자 지정 필드 또는 속성을 만들 수 있습니다. 다음 예제는 문서에 “부서”를 추가하는 사용자 지정 필드 또는 속성입니다.

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

```
    }
```

Amazon S3 데이터 소스에 사용자 지정 속성 또는 필드 추가

S3 버킷을 인덱스의 데이터 소스로 사용하는 경우, 관련 메타데이터 파일과 함께 문서에 메타데이터를 추가합니다. 문서와 병렬인 디렉터리 구조에 메타데이터 JSON 파일을 배치합니다. 자세한 내용은 S3 문서 메타데이터를 참조하세요.

`Attributes` JSON 구조에서 사용자 지정 필드 또는 속성을 지정합니다. 최대 500개의 사용자 지정 필드 또는 속성을 만들 수 있습니다. 예를 들어, 다음 예제에서는 `Attributes`를 사용하여 사용자 지정 필드 또는 속성 3개와 예약된 필드 1개를 정의합니다.

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

다음 단계는 Amazon S3 데이터 소스에 사용자 지정 속성을 추가하는 방법을 안내합니다.

주제

- 1단계: 아마존 켄드라 인덱스 생성
- 2단계: 색인을 업데이트하여 사용자 지정 문서 필드 추가
- 3단계: Amazon S3 데이터 소스를 생성하고 데이터 소스 필드를 사용자 지정 속성에 매핑

1단계: 아마존 켄드라 인덱스 생성

다음 단계에 따라 Amazon Kendra 인덱스 생성 인덱스를 생성하십시오.

2단계: 색인을 업데이트하여 사용자 지정 문서 필드 추가

색인을 만든 후 색인에 필드를 추가합니다. 다음 절차는 콘솔과 CLI를 사용하여 인덱스에 필드를 추가하는 방법을 보여줍니다.

Console

인덱스 필드를 만들려면

1. 색인을 생성했는지 확인하세요.

2. 그런 다음 왼쪽 탐색 메뉴의 데이터 관리에서 패싯 정의를 선택합니다.
3. 인덱스 필드 설정 가이드의 인덱스 필드에서 필드 추가를 선택하여 사용자 지정 필드를 추가합니다.
4. 인덱스 필드 추가 대화 상자에서 다음을 수행합니다.

   - 필드 이름 - 필드 이름을 추가합니다.
   - 데이터 유형 - 문자열, 문자열 목록 또는 날짜 중에서 데이터 유형을 선택합니다.
   - 사용 유형 - 팩터블, 검색 가능, 표시 가능, 정렬 가능 등 사용 유형을 선택합니다.

     그런 다음 추가를 선택합니다.

   매핑하려는 다른 필드에 대해 마지막 단계를 반복합니다.

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
        }
    }
```

```
...
]"
```

3단계: Amazon S3 데이터 소스를 생성하고 데이터 소스 필드를 사용자 지정 속성에 매핑

Amazon S3 데이터 소스를 생성하고 여기에 필드를 매핑하려면 의 지침을 따르십시오Amazon S3.

API를 사용하는 경우 API를 사용할 `configuration` 때 아래에 있는 `fieldMappings` 속성을 사용하십시오. CreateDataSource

데이터 소스 필드가 매핑되는 방식에 대한 개요는 을 참조하십시오데이터 소스 필드 매핑.

토큰을 사용하여 문서에 대한 사용자 액세스 제어

인덱스의 특정 문서에 액세스하거나 검색 결과에서 특정 문서를 볼 수 있는 사용자 또는 그룹을 제어할 수 있습니다. 이를 사용자 컨텍스트 필터링이라고 합니다. 문서에 대한 액세스를 제어할 수 있는 이점이 있는 일종의 개인화된 검색입니다. 예를 들어 회사 포털에서 정보를 검색하는 모든 팀이 일급 기밀 회사 문서에 액세스해야 하는 것은 아니며, 이러한 문서가 모든 사용자에게 관련된 것도 아닙니다. 일급 기밀 문서에 대한 액세스 권한을 받은 특정 사용자 또는 팀 그룹만 검색 결과에서 이러한 문서를 볼 수 있습니다.

Amazon Kendra는 다음 토큰 유형을 사용하여 토큰 기반 사용자 액세스 제어를 활성화합니다.

- Open ID
- 공유 암호가 있는 JWT
- 퍼블릭 키가 있는 JWT
- JSON

Amazon Kendra는 검색 애플리케이션을 위한 매우 안전한 엔터프라이즈 검색을 제공합니다. 검색 결과는 조직의 보안 모델을 반영합니다. 고객은 사용자가 검색 애플리케이션에 액세스할 수 있도록 인증하고 권한을 부여할 책임이 있습니다. 검색 시 Amazon Kendra 서비스는 고객의 검색 애플리케이션에서 제공한 사용자 ID와 크롤링/인덱싱 시간 동안 Amazon Kendra 커넥터가 수집한 문서 액세스 제어 목록(ACL)을 기반으로 검색 결과를 필터링합니다. 검색 결과에는 원본 문서 리포지토리를 가리키는 URL과 짧은 발췌문이 반환됩니다. 원본 리포지토리에서는 여전히 전체 문서에 대한 액세스를 강제합니다.

주제

OpenID 사용

액세스 제어에 OpenID 토큰을 사용하도록 Amazon Kendra 인덱스를 구성하려면 OpenID 공급자의 JWKS (JSON 웹 키 세트) URL이 필요합니다. 대부분의 경우 JWKS URL은 다음과 같은 형식입니다 (OpenID 검색을 따르는 경우). `https://domain-name/.well_known/jwks.json`

다음 예제는 인덱스를 생성할 때 사용자 액세스 제어를 위해 OpenID 토큰을 사용하는 방법을 보여줍니다.

Console

1. 새 인덱스 생성을 시작하려면 인덱스 생성을 선택합니다.
2. 인덱스 세부정보 지정 페이지에서 인덱스에 이름과 설명을 입력합니다.
3. IAM 역할에서 역할을 선택하거나 새 역할 생성을 선택하고 역할 이름을 지정하여 새 역할을 생성합니다. IAM 역할의 접두사는 "- “입니다. AmazonKendra
4. 다른 모든 필드의 기본값은 그대로 둡니다. 다음을 선택합니다.
5. 사용자 액세스 제어 구성 페이지의 액세스 제어 설정에서 예를 선택하여 액세스 제어에 토큰을 사용합니다.
6. 토큰 구성에서 OpenID를 토큰 유형으로 선택합니다.
7. 서명 키 URL을 지정합니다. URL은 JSON 웹 키 집합을 가리켜야 합니다.
8. 선택 사항 고급 구성에서:
    a. ACL 검사에 사용할 사용자 이름을 지정합니다.
    b. ACL 검사에 사용할 그룹을 하나 이상 지정합니다.
    c. 토큰 발행자를 검증할 발행자를 지정합니다.
    d. 클라이언트 ID를 지정합니다. JWT의 대상과 일치하는 정규 표현식을 지정해야 합니다.
9. 프로비저닝 세부 정보 페이지에서 개발자 에디션을 선택합니다.
10. 생성을 선택하여 인덱스를 생성합니다.
11. 인덱스가 생성될 때까지 기다리세요. Amazon Kendra 인덱스의 하드웨어를 프로비저닝합니다. 이 작업에는 다소 시간이 걸릴 수 있습니다.

CLI

JSON 입력 파일을 AWS CLI 사용하여 색인을 만들려면 먼저 원하는 매개 변수를 사용하여 JSON 파일을 만드십시오.

```
{
    "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"
}
```

기본 사용자 및 그룹 필드 이름을 재정의할 수 있습니다. `UserNameAttributeField`의 기본값은 “user”입니다. `GroupAttributeField`의 기본값은 "groups"입니다.

다음으로, 입력 파일을 사용하여 `create-index`를 호출합니다. 예를 들어 JSON 파일 이름이 `create-index-openid.json`인 경우 다음을 사용할 수 있습니다.

```
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'
)
```

공유 암호가 있는 JSON 웹 토큰(JWT) 사용

다음 예는 인덱스를 생성할 때 사용자 액세스 제어를 위한 공유 암호 토큰과 함께 JSON 웹 토큰 (JWT)을 사용하는 방법을 보여줍니다.

Console

1. 새 인덱스 생성을 시작하려면 인덱스 생성을 선택합니다.
2. 인덱스 세부정보 지정 페이지에서 인덱스에 이름과 설명을 입력합니다.
3. IAM 역할에서 역할을 선택하거나 새 역할 생성을 선택하고 역할 이름을 지정하여 새 역할을 생성합니다. IAM 역할 접두사는 "AmazonKendra- “입니다.
4. 다른 모든 필드의 기본값은 그대로 둡니다. 다음을 선택합니다.
5. 사용자 액세스 제어 구성 페이지의 액세스 제어 설정에서 예를 선택하여 액세스 제어에 토큰을 사용합니다.
6. 토큰 구성에서 토큰 유형으로 공유 암호와 JWT를 선택합니다.
7. 공유 암호 서명 파라미터에서 암호 유형을 선택합니다. 기존의 AWS Secrets Manager 공유 암호를 사용하거나 새 공유 암호를 생성할 수 있습니다.

   새 공유 암호를 생성하려면 새로 만들기를 선택한 후 다음 단계를 수행합니다.

   a. 새 AWS Secrets Manager 비밀번호에서 시크릿 이름을 지정합니다. 퍼블릭 키를 저장하면 `AmazonKendra-` 접두사가 추가됩니다.

   b. 키 ID를 지정합니다. 키 ID는 토큰의 JSON 웹 서명을 보호하는 데 어떤 키가 사용되었는지 나타내는 힌트입니다.

   c. 토큰의 서명 알고리즘을 선택합니다. ID 토큰을 보호하는 데 사용되는 암호화 알고리즘입니다. RSA에 대한 자세한 내용은 RSA 암호화를 참조하세요.

d. base64 URL로 인코딩된 암호를 입력하여 공유 암호를 지정합니다. 암호 생성을 선택하여 암호가 자동으로 생성되도록 할 수도 있습니다. 암호가 base64 URL로 인코딩된 암호인지 확인해야 합니다.

e. (선택 사항) 공유 암호가 유효한 시기를 지정합니다. 암호가 유효한 날짜와 시간(시작, 종료, 또는 둘 모두)을 지정할 수 있습니다. 암호는 지정한 기간 동안 유효합니다.

f. 암호 저장을 선택하여 새 암호를 저장합니다.

8. (선택 사항) 고급 구성에서:

a. ACL 검사에 사용할 사용자 이름을 지정합니다.

b. ACL 검사에 사용할 그룹을 하나 이상 지정합니다.

c. 토큰 발행자를 검증할 발행자를 지정합니다.

d. 클레임 ID를 지정합니다. JWT의 대상과 일치하는 정규 표현식을 지정해야 합니다.

9. 프로비저닝 세부 정보 페이지에서 개발자 에디션을 선택합니다.

10. 생성을 선택하여 인덱스를 생성합니다.

11. 색인이 생성될 때까지 기다리세요. Amazon Kendra 인덱스의 하드웨어를 프로비저닝합니다. 이 작업에는 다소 시간이 걸릴 수 있습니다.

CLI

공유 암호가 들어 있는 JWT 토큰을 사용할 수 있습니다 AWS Secrets Manager. 암호는 base64 URL로 인코딩된 암호여야 합니다. Secrets Manager ARN이 필요하며 Amazon Kendra 역할에 리소스에 `GetSecretValue` 대한 액세스 권한이 있어야 합니다. Secrets Manager 를 사용하여 Secrets Manager AWS KMS리소스를 암호화하는 경우 역할에 암호 해독 작업에 대한 액세스 권한도 있어야 합니다.

JSON 입력 파일을 AWS CLI 사용하여 색인을 만들려면 먼저 원하는 매개 변수를 사용하여 JSON 파일을 생성합니다.

```
{
    "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"
}
```

기본 사용자 및 그룹 필드 이름을 재정의할 수 있습니다. `UserNameAttributeField`의 기본값은 “user”입니다. `GroupAttributeField`의 기본값은 "groups"입니다.

다음으로, 입력 파일을 사용하여 `create-index`를 호출합니다. 예를 들어 JSON 파일 이름이 `create-index-openid.json`인 경우 다음을 사용할 수 있습니다.

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

시크릿의 형식은 다음과 같아야 합니다. 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"
    }
  ]
}
```

JWT에 대한 자세한 내용은 jwt.io를 참조하세요.

Python

공유 암호가 들어 있는 JWT 토큰을 사용할 수 있습니다 AWS Secrets Manager. 암호는 base64 URL로 인코딩된 암호여야 합니다. Secrets Manager ARN이 필요하며 Amazon Kendra 역할에

리소스에 GetSecretValue 대한 액세스 권한이 있어야 합니다. Secrets Manager 를 사용하여 Secrets Manager AWS KMS리소스를 암호화하는 경우 역할에 암호 해독 작업에 대한 액세스 권한도 있어야 합니다.

```
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'
)
```

JSON 웹 토큰(JWT)을 퍼블릭 키와 함께 사용

다음 예는 인덱스를 생성할 때 사용자 액세스 제어를 위한 공개 키와 함께 JSON 웹 토큰 (JWT) 을 사용하는 방법을 보여줍니다. JWT에 대한 자세한 내용은 jwt.io를 참조하세요.

Console

1. 새 인덱스 생성을 시작하려면 인덱스 생성을 선택합니다.
2. 인덱스 세부정보 지정 페이지에서 인덱스에 이름과 설명을 입력합니다.
3. IAM 역할에서 역할을 선택하거나 새 역할 생성을 선택하고 역할 이름을 지정하여 새 역할을 생성합니다. IAM 역할 접두사는 "AmazonKendra- “입니다.
4. 다른 모든 필드의 기본값은 그대로 둡니다. 다음을 선택합니다.
5. 사용자 액세스 제어 구성 페이지의 액세스 제어 설정에서 예를 선택하여 액세스 제어에 토큰을 사용합니다.
6. 토큰 구성에서 토큰 유형으로 퍼블릭 키와 JWT를 선택합니다.

7. 퍼블릭 키 서명 파라미터에서 암호 유형을 선택합니다. 기존의 AWS Secrets Manager 암호를 사용하거나 새 암호를 생성할 수 있습니다.

   새 암호를 생성하려면 새로 만들기를 선택한 후 다음 단계를 수행합니다.

   a. 새 AWS Secrets Manager 비밀번호에서 시크릿 이름을 지정합니다. 퍼블릭 키를 저장하면 `AmazonKendra-` 접두사가 추가됩니다.
   b. 키 ID를 지정합니다. 키 ID는 토큰의 JSON 웹 서명을 보호하는 데 어떤 키가 사용되었는지 나타내는 힌트입니다.
   c. 토큰의 서명 알고리즘을 선택합니다. ID 토큰을 보호하는 데 사용되는 암호화 알고리즘입니다. RSA에 대한 자세한 내용은 RSA 암호화를 참조하세요.
   d. 인증서 속성에서 선택적 인증서 체인을 지정합니다. 인증서 체인은 인증서 목록으로 구성됩니다. 서버의 인증서로 시작하여 루트 인증서로 종료됩니다.
   e. 선택 사항 엄지 또는 검지 지문을 지정합니다. 모든 인증서 데이터와 서명을 대상으로 계산된 인증서의 해시여야 합니다.
   f. 지수를 지정합니다. RSA 퍼블릭 키의 지수 값입니다. 이 값은 Base64urlUInt 인코딩 값으로 표현됩니다.
   g. 모듈러스를 지정합니다. RSA 퍼블릭 키의 지수 값입니다. 이 값은 Base64urlUInt 인코딩 값으로 표현됩니다.
   h. 키 저장을 선택하여 새 키를 저장합니다.
8. 선택 사항 고급 구성에서:

   a. ACL 검사에 사용할 사용자 이름을 지정합니다.
   b. ACL 검사에 사용할 그룹을 하나 이상 지정합니다.
   c. 토큰 발행자를 검증할 발행자를 지정합니다.
   d. 클라이언트 ID를 지정합니다. JWT의 대상과 일치하는 정규 표현식을 지정해야 합니다.
9. 프로비저닝 세부 정보 페이지에서 개발자 에디션을 선택합니다.
10. 생성을 선택하여 인덱스를 생성합니다.
11. 색인이 생성될 때까지 기다리세요. Amazon Kendra 인덱스의 하드웨어를 프로비저닝합니다. 이 작업에는 다소 시간이 걸릴 수 있습니다.

CLI

AWS Secrets Manager내부에 퍼블릭 키가 있는 JWT를 사용할 수 있습니다. Secrets Manager ARN이 필요하며 Amazon Kendra 역할에 리소스에 `GetSecretValue` 대한 액세스 권한이 있어야

합니다. Secrets Manager 를 사용하여 Secrets Manager AWS KMS리소스를 암호화하는 경우 역할에 암호 해독 작업에 대한 액세스 권한도 있어야 합니다.

JSON 입력 파일을 AWS CLI 사용하여 색인을 만들려면 먼저 원하는 매개변수를 사용하여 JSON 파일을 생성합니다.

```
{
    "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"
}
```

기본 사용자 및 그룹 필드 이름을 재정의할 수 있습니다. `UserNameAttributeField`의 기본값은 “user”입니다. `GroupAttributeField`의 기본값은 "groups"입니다.

다음으로, 입력 파일을 사용하여 `create-index`를 호출합니다. 예를 들어 JSON 파일 이름이 `create-index-openid.json`인 경우 다음을 사용할 수 있습니다.

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

시크릿의 형식은 다음과 같아야 합니다. 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"
      ]
    }
  ]
}
```

JWT에 대한 자세한 내용은 jwt.io를 참조하세요.

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'
)
```

JSON 사용

다음 예시는 인덱스를 생성할 때 사용자 액세스 제어에 JSON을 사용하는 방법을 보여줍니다.

> **Warning**
>
> JSON 토큰은 검증되지 않은 페이로드입니다. 이는 Amazon Kendra 요청이 브라우저가 아닌 신뢰할 수 있는 서버에서 오는 경우에만 사용해야 합니다.

Console

1. 새 인덱스 생성을 시작하려면 인덱스 생성을 선택합니다.
2. 인덱스 세부정보 지정 페이지에서 인덱스에 이름과 설명을 입력합니다.
3. IAM 역할에서 역할을 선택하거나 새 역할 생성을 선택하고 역할 이름을 지정하여 새 역할을 생성합니다. IAM 역할의 접두사는 "AmazonKendra- “입니다.
4. 다른 모든 필드의 기본값은 그대로 둡니다. 다음을 선택합니다.
5. 사용자 액세스 제어 구성 페이지의 액세스 제어 설정에서 예를 선택하여 액세스 제어에 토큰을 사용합니다.
6. 토큰 구성에서 토큰 유형으로 JSON을 선택합니다.
7. ACL 검사에 사용할 사용자 이름을 지정합니다.
8. ACL 검사에 사용할 그룹을 하나 이상 지정합니다.
9. 다음을 선택합니다.
10. 프로비저닝 세부 정보 페이지에서 개발자 에디션을 선택합니다.
11. 생성을 선택하여 인덱스를 생성합니다.
12. 색인이 생성될 때까지 기다리세요. Amazon Kendra 인덱스의 하드웨어를 프로비저닝합니다. 이 작업에는 다소 시간이 걸릴 수 있습니다.

CLI

JSON 입력 파일을 AWS CLI 사용하여 색인을 만들려면 먼저 원하는 매개 변수를 사용하여 JSON 파일을 만드십시오.

```
{
    "Name": "user-context",
    "Edition": "ENTERPRISE_EDITION",
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",
    "UserTokenConfigurations": [
        {
            "JsonTokenTypeConfiguration": {
```

```
                "UserNameAttributeField": "user",
                "GroupAttributeField": "group"
            }
        }
    ],
    "UserContextPolicy": "USER_TOKEN"
}
```

다음으로, 입력 파일을 사용하여 `create-index`를 호출합니다. 예를 들어 JSON 파일 이름이 `create-index-openid.json`인 경우 다음을 사용할 수 있습니다.

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

Open ID를 사용하지 않는 경우 JSON 형식으로 토큰을 보내주시면 됩니다. AWS IAM Identity Center그럴 경우 JSON 토큰의 어떤 필드에 사용자 이름이 포함되고 어떤 필드에 그룹이 포함되는지 지정해야 합니다. 그룹 필드 값은 JSON 문자열 배열이어야 합니다. 예를 들어 SAML을 사용하는 경우 토큰은 다음과 같을 것입니다.

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

`TokenConfiguration`은 사용자 이름과 그룹 필드 이름을 지정합니다.

```
{
    "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'
)
```

데이터 소스 커넥터 생성

문서에 연결하고 문서를 색인화하는 Amazon Kendra 데 사용할 데이터 소스 커넥터를 만들 수 있습니다. Amazon Kendra Microsoft SharePoint, Google Drive 및 기타 여러 공급자에 연결할 수 있습니다. 데이터 소스 커넥터를 만들 때는 소스 리포지토리에 연결하는 데 필요한 구성 정보를 제공합니다 Amazon Kendra . 인덱스에 직접 문서를 추가하는 것과 달리, 데이터 소스를 정기적으로 스캔하여 인덱스를 업데이트할 수 있습니다.

예를 들어 Amazon S3 버킷에 세금 문서 저장소가 저장되어 있다고 가정해 보겠습니다. 수시로, 기존 문서가 변경되고 새 문서가 리포지토리에 추가됩니다. 리포지토리를 데이터 Amazon Kendra 소스로 추가하는 경우 데이터 소스와 인덱스 간에 주기적인 동기화를 설정하여 색인을 최신 상태로 유지할 수 있습니다.

콘솔이나 API를 사용하여 색인을 수동으로 업데이트하도록 선택할 수 있습니다. StartDataSourceSyncJob 그렇지 않으면, 인덱스를 업데이트하고 데이터 소스와 동기화하도록 일정을 설정합니다.

인덱스에는 데이터 소스가 둘 이상 있을 수 있습니다. 각 데이터 소스에는 고유한 업데이트 일정이 있을 수 있습니다. 예를 들어 작업 문서의 인덱스를 매일 또는 매시간 업데이트하고 아카이브가 변경될 때마다 보관된 문서를 수동으로 업데이트할 수 있습니다.

문서 수집 프로세스 중에 문서 메타데이터나 속성 및 콘텐츠를 변경하려면 Amazon Kendra 사용자 지정 문서 보강을 참조하세요.

> **Note**
>
> 각 문서 ID는 색인별로 고유해야 합니다. 고유한 ID로 문서를 인덱싱하는 데이터 소스를 만든 다음 `BatchPutDocument` API를 사용하여 동일한 문서를 인덱싱할 수 없으며 그 반대의 경우도 마찬가지입니다. 데이터 소스를 삭제한 다음 `BatchPutDocument` API를 사용하여 동일한 문서를 인덱싱하거나 그 반대로 할 수 있습니다. 동일한 문서 세트에 대해 `BatchPutDocument` 및 `BatchDeleteDocument` API를 Amazon Kendra 데이터 소스 커넥터와 함께 사용하면 데이터가 일치하지 않을 수 있습니다. 대신, Amazon Kendra 사용자 지정 데이터 소스 커넥터를 사용하는 것이 좋습니다.

> **Note**
>
> 인덱스에 추가된 파일은 UTF-8 인코딩된 바이트 스트림에 있어야 합니다. 의 문서에 대한 자세한 내용은 문서를 참조하십시오. Amazon Kendra

업데이트 일정 설정

콘솔에서 정기적으로 업데이트하거나, 데이터 소스를 만들거나 업데이트할 때 `Schedule` 파라미터를 사용하여 데이터 소스를 구성합니다. 파라미터의 내용은 `cron` 형식 일정 문자열이나 필요에 따라 인덱스가 업데이트됨을 나타내는 빈 문자열을 포함하는 문자열입니다. cron 표현식의 형식에 대한 내용은 Amazon CloudWatch Events 사용 설명서의 규칙에 대한 스케줄 표현식을 참조하십시오. Amazon Kendra cron 표현식만 지원합니다. rate 표현식은 지원하지 않습니다.

언어 설정

지원되는 언어로 데이터 소스의 모든 문서를 인덱싱할 수 있습니다. CreateDataSource호출할 때 데이터 소스에 있는 모든 문서의 언어 코드를 지정합니다. 메타데이터 필드에 지정된 언어 코드가 없는 문서의 경우, 데이터 소스 수준에서 모든 문서에 지정된 언어 코드를 사용하여 문서가 인덱싱됩니다. 언어를 지정하지 않으면 Amazon Kendra 는 기본적으로 데이터 소스의 문서를 영어로 인덱싱합니다. 코드를 포함하여 지원되는 언어에 대한 자세한 내용은 영어 이외의 언어로 문서 추가를 참조하세요.

콘솔을 사용하여 지원되는 언어로 데이터 소스의 모든 문서를 인덱싱하세요. 데이터 소스로 이동하여 데이터 소스를 편집하거나, 데이터 소스 추가를 통해 새 데이터 소스를 추가합니다. 데이터 소스 세부 정보 지정 페이지의 언어 드롭다운에서 언어를 선택합니다. 업데이트를 선택하거나 구성 정보를 계속 입력하여 데이터 소스에 연결합니다.

데이터 소스 커넥터

이 섹션에서는 및 API를 Amazon Kendra 사용하여 Amazon Kendra 지원되는 데이터베이스 및 데이터 원본 리포지토리에 연결하는 방법을 보여줍니다. AWS Management Console Amazon Kendra

주제

- 데이터 소스 템플릿 스키마
- Adobe Experience Manager
- Alfresco
- Aurora (MySQL)

- Aurora (PostgreSQL)
- Amazon FSx (윈도우)
- Amazon FSx (NetApp ONTAP)
- Amazon RDS/Aurora
- Amazon RDS (마이크로소프트 SQL 서버)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon S3
- Amazon Kendra 웹 크롤러
- Amazon WorkDocs
- Box
- Confluence
- 사용자 지정 데이터 소스 커넥터
- Dropbox
- Drupal
- GitHub
- Gmail
- Google Drive
- IBM DB2
- Jira
- Microsoft Exchange
- 마이크로소프트 OneDrive
- 마이크로소프트 SharePoint
- Microsoft SQL Server
- Microsoft Teams
- Microsoft Yammer
- MySQL
- Oracle Database
- PostgreSQL

- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

데이터 소스 템플릿 스키마

다음은 템플릿이 지원되는 데이터 소스의 템플릿 스키마입니다.

주제

- Adobe Experience Manager 템플릿 스키마
- Amazon FSx (Windows) 템플릿 스키마
- Amazon FSx (NetApp ONTAP) 템플릿 스키마
- Alfresco 템플릿 스키마
- Aurora (MySQL) 템플릿 스키마
- Aurora (PostgreSQL) 템플릿 스키마
- Amazon RDS (Microsoft SQL Server) 템플릿 스키마
- Amazon RDS (MySQL) 템플릿 스키마
- Amazon RDS (오라클) 템플릿 스키마
- Amazon RDS (PostgreSQL) 템플릿 스키마
- Amazon S3 템플릿 스키마
- Amazon Kendra 웹 크롤러 템플릿 스키마
- Confluence 템플릿 스키마
- Dropbox 템플릿 스키마
- Drupal 템플릿 스키마
- GitHub 템플릿 스키마
- Gmail 템플릿 스키마
- Google Drive 템플릿 스키마
- IBM DB2 템플릿 스키마
- Microsoft Exchange 템플릿 스키마
- Microsoft OneDrive 템플릿 스키마

- Microsoft SharePoint 템플릿 스키마
- Microsoft SQL Server 템플릿 스키마
- Microsoft Teams 템플릿 스키마
- Microsoft Yammer 템플릿 스키마
- MySQL 템플릿 스키마
- Oracle Database 템플릿 스키마
- PostgreSQL 템플릿 스키마
- Salesforce 템플릿 스키마
- ServiceNow 템플릿 스키마
- 슬랙 템플릿 스키마
- Zendesk 템플릿 스키마

Adobe Experience Manager 템플릿 스키마

데이터 소스 스키마를 TemplateConfiguration 객체의 일부로 포함하는 JSON을 포함합니다. Adobe Experience Manager 호스트 URL, 인증 유형, 그리고 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 클라우드 서비스형 Adobe Experience Manager(AEM) 또는 AEM 온프레미스 중에서 무엇을 사용할지 여부를 제공합니다. 또한 데이터 소스 유형으로 AEM, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 CreateDataSource를 호출할 때 TEMPLATE을 Type으로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. 자세한 정보는 Adobe Experience Manager JSON 스키마을 참조하세요.

다음 표는 AEM JSON 스키마의 매개변수를 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
aemUrl	Adobe Experience Manager 호스트 URL. 예를 들어 AEM 온프레미스를 사용하는 경우 호스트 이름과 포트를 포함해야 합니다. https://hostname:port 클라우드 서비스형 AEM을 사용하는 경우 작성자 URL을 사용할 수 있습니다.

구성	설명
	https://author-xxxxxx-xxxxxxx.adobeaemcloud.com
authType	사용하는 인증 유형(`Basic` 또는 `OAuth2`).
deploymentType	사용하는 Adobe Experience Manager 유형(`CLOUD` 또는 `ON_PREMISE`).
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• 페이지를 방문하십시오 • asset	Adobe Experience Manager페이지 및 자산의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
timeZoneId	AEM On-Premise를 사용하고 서버의 시간대가 Amazon Kendra AEM 커넥터 또는 인덱스의 시간대와 다른 경우 AEM 커넥터 또는 색인에 맞춰 서버 시간대를 지정할 수 있습니다. AEM 온-프레미스의 기본 시간대는 AEM 커넥터 또는 인덱스의 시간대입니다. Amazon Kendra 클라우드 서비스형 AEM의 기본 시간대는 그리니치 표준시입니다.
• pageRootPaths • assetRootPaths	페이지 및 자산의 루트 경로 목록. 예를 들어 페이지의 루트 경로는 /content/sub일 수 있고 자산의 루트 경로는 /content/sub/asset1일 수 있습니다.
crawlAssets	자산을 크롤링할 `true`.

구성	설명
crawlPages	페이지를 크롤링할 `true`.
• pagePathInclusion패턴 • pageNameInclusion패턴 • assetPathInclusion패턴 • assetTypeInclusion패턴 • assetNameInclusion패턴	Adobe Experience Manager 데이터 소스에서 특정 페이지 및 자산을 포함하는 정규식 패턴 목록. 패턴과 일치하는 항목은 인덱스에 포함됩니다. 패턴과 일치하지 않는 항목은 인덱스에서 제외됩니다. 페이지 또는 자산이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 콘텐츠는 인덱스에 포함되지 않습니다.
• pagePathExclusion패턴 • pageNameExclusion패턴 • assetPathExclusion패턴 • assetTypeInclusion패턴 • assetNameInclusion패턴	Adobe Experience Manager 데이터 소스에서 특정 페이지 및 자산을 제외하는 정규식 패턴 목록. 패턴과 일치하는 항목은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 항목은 인덱스에 포함됩니다. 페이지 또는 자산이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 콘텐츠는 인덱스에 포함되지 않습니다.
pageComponents	인덱싱하려는 특정 페이지 구성 요소의 목록.
contentFragmentVariations	인덱싱하려는 Adobe Experience Manager 내용 조각의 저장된 특정 변형의 이름 목록.
type	데이터 소스의 유형. `AEM`을 데이터 소스 유형으로 지정합니다.
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 사용자 및 그룹의 ID/주체 정보를 특정 문서에 대한 액세스 권한과 동기화합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

구성	설명
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	Adobe Experience Manager에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 암호의 Amazon 리소스 이름 (ARN). 이러한 키-값 쌍에 대한 자세한 내용은 Adobe Experience Manager의 연결 지침을 참조하십시오.
version	현재 지원되는 이 템플릿의 버전.

Adobe Experience Manager JSON 스키마

```
{
  "$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 (Windows) 템플릿 스키마

데이터 소스 스키마를 TemplateConfiguration 객체의 일부로 포함하는 JSON을 포함합니다. 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 파일 시스템 ID를 제공합니다. 또한 데이터 원본의 유형FSX, 인증 자격 증명의 암호 및 기타 필요한 구성을 지정해야 합니다. 그런 다음 CreateDataSource를 호출할 때 TEMPLATE을 Type으로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Amazon FSx (윈도우) JSON 스키마를 참조하세요.

다음 표에서는 Amazon FSx (Windows) JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
fileSystemId	Amazon FSx 파일 시스템의 식별자입니다. 파일 시스템 ID는 Amazon FSx 콘솔의 파일 시스템 대시보드에서 찾을 수 있습니다.
fileSystemType	Amazon FSx 파일 시스템 유형. 파일 시스템 Windows File Server 유형으로 사용하려면 지정하십시오`WINDOWS`.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
모두	Amazon FSx 데이터 원본에 있는 파일의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
isCrawlAcl	`true`ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 크롤링하려면 ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
inclusionPatterns	Amazon FSx 데이터 원본에 특정 파일을 포함하기 위한 정규 표현식 패턴 목록입니다. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다.

구성	설명
	파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
exclusionPatterns	Amazon FSx 데이터 원본에서 특정 파일을 제외하기 위한 정규 표현식 패턴 목록입니다. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 제외 및 포함 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 사용자 및 그룹의 ID/주체 정보를 특정 문서에 대한 액세스 권한과 동기화하는 데 사용됩니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • `FORCED_FULL_CRAWL` 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • `FULL_CRAWL` 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

구성	설명
type	데이터 소스의 유형. Windows 파일 시스템 데이터 원본의 경우 지정하십시오FSX.

Amazon FSx (윈도우) JSON 스키마

```
{
  "$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 (NetApp ONTAP) 템플릿 스키마

데이터 소스 스키마를 TemplateConfiguration 객체의 일부로 포함하는 JSON을 포함합니다. 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 파일 시스템 ID와 SVM (스토리지 가상 머신) 을 제공

합니다. 또한 데이터 소스의 유형FSXONTAP, 인증 자격 증명의 암호 및 기타 필요한 구성을 지정해야 합니다. 그런 다음 CreateDataSource를 호출할 때 TEMPLATE을 Type으로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Amazon FSx (NetApp ONTAP) JSON 스키마를 참조하세요.

다음 표는 Amazon FSx (NetApp ONTAP) JSON 스키마의 매개 변수를 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
fileSystemId	Amazon FSx 파일 시스템의 식별자입니다. 파일 시스템 ID는 Amazon FSx 콘솔의 파일 시스템 대시보드에서 찾을 수 있습니다. Amazon FSx 콘솔에서 NetApp ONTAP용 파일 시스템을 생성하는 방법에 대한 자세한 내용은 FSx for ONTAP 사용 설명서의 NetAppONTAP 시작 안내서를 참조하십시오.
fileSystemType	Amazon FSx 파일 시스템 유형. 파일 시스템 NetApp ONTAP 유형으로 사용하려면 지정하십시오ONTAP.
SVMid	Amazon FSx 파일 시스템에서 사용되는 스토리지 가상 머신 (SVM) 의 식별자입니다. NetApp ONTAP Amazon FSx 콘솔의 파일 시스템 대시보드로 이동하여 파일 시스템 ID를 선택한 다음 스토리지 가상 시스템을 선택하면 SVM ID를 찾을 수 있습니다. Amazon FSx 콘솔에서 파일 시스템을 생성하는 방법에 대한 자세한 내용은 사용 FSx for ONTAP 설명서의 NetApp ONTAP 시작 안내서를 참조하십시오. NetApp ONTAP
프로토콜 유형	Windows용 CIFS (공용 인터넷 파일 시스템) 프로토콜을 사용하는지, Linux용 네트워크 파일 시스템 (NFS) 프로토콜을 사용하는지 여부.

구성	설명
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
파일	Amazon FSx 데이터 원본에 있는 파일의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요. 데이터 원본 필드 이름은 파일의 사용자 지정 메타데이터에 있어야 합니다.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
crawlAcl	`true`ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 크롤링합니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
inclusionPatterns	Amazon FSx 데이터 원본에 특정 파일을 포함하기 위한 정규 표현식 패턴 목록입니다. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.

구성	설명
exclusionPatterns	Amazon FSx 데이터 원본에서 특정 파일을 제외하기 위한 정규 표현식 패턴 목록입니다. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 제외 및 포함 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
type	데이터 소스의 유형. NetApp ONTAP파일 시스템 데이터 원본의 경우 지정하십시오FSXONTAP.
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

구성	설명

secretArn	파일 시스템에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 암호의 Amazon 리소스 이름 (ARN). Amazon FSx 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다.
	
	`{ "username": " user@corp.example.com ", "password": " password" }`
	
	Amazon FSx 파일 시스템에 NFS 프로토콜을 사용하는 경우 암호는 다음 키와 함께 JSON 구조에 저장됩니다.
	
	`{ "leftId": "left ID", "rightId": " right ID", "preSharedKey": " pre-shared key " }`

Amazon FSx (NetApp ONTAP) JSON 스키마

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "fileSystemId": {
              "type": "string",
                "pattern": "^(fs-[0-9a-f]{8,21})$"
            },
```

```
            "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"
  ]
}
```

Alfresco 템플릿 스키마

데이터 소스 스키마를 TemplateConfiguration 객체의 일부로 포함하는 JSON을 포함합니다. Alfresco 사이트 ID, 리포지토리 URL, 사용자 인터페이스 URL, 인증 유형, 클라우드 또는 온프레미스 사용 여부, 크롤링하려는 콘텐츠 유형을 제공합니다. 이 정보는 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 제공합니다. 또한 데이터 소스 유형으로 `ALFRESCO`, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 CreateDataSource를 호출할 때 `TEMPLATE`을 `Type`으로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Alfresco JSON 스키마를 참조하세요.

다음 표에는 Alfresco JSON 스키마의 매개 변수가 설명되어 있습니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
siteId	Alfresco 사이트의 식별자.
repoUrl	Alfresco 리포지토리의 URL. 리포지토리 URL은 Alfresco 관리자로부터 받을 수 있습니다. 예를 들어 Alfresco Cloud(PaaS)를 사용하는 경우 리포지토리 URL은 https://company.alfrescocloud.com과 같을 수 있습니다. 또는 Alfresco 온프레미스를 사용하는 경우 리포지토리 URL은 https://company-alfresco-instance.company-domain.suffix:port와 같을 수 있습니다.
webAppUrl	Alfresco 사용자 인터페이스의 URL. Alfresco 사용자 인터페이스 URL은 Alfresco 관리자로부터 받을 수 있습니다. 예를 들어 사용자 인터페이스 URL은 https://example.com과 같을 수 있습니다.
repositoryAdditionalProperties	리포지토리/데이터 소스 엔드포인트와 연결하기 위한 추가 속성.
authType	사용하는 인증 유형(`OAuth2` 또는 `Basic`).

구성	설명
type (deployment)	사용하는 Alfresco 유형(PAAS 또는 ON-PREM).
crawlType	크롤링하려는 콘텐츠의 유형, 즉 ASPECT(Alfresco에서 'Aspect'로 표시된 콘텐츠), SITE_ID(특정 Alfresco 사이트 내 콘텐츠) 또는 ALL_SITES (모든 Alfresco 사이트의 콘텐츠).
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• 문서 • 설명	Alfresco 문서 및 주석의 속성 또는 필드 이름을 인덱스 필드 이름에 매핑하는 객체 목록입니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
aspectName	인덱싱하려는 특정 'Aspect'의 이름.
aspectProperties	인덱싱하려는 특정 'Aspect' 콘텐츠 속성의 목록.
enableFineGrained제어	'Aspects'를 크롤링할 true.
isCrawlComment	true댓글을 크롤링하려면.
• inclusionFileName패턴 • inclusionFileType패턴 • inclusionFilePath패턴	Alfresco 데이터 소스에서 특정 파일을 포함하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.

구성	설명
• exclusionFileName패턴 • exclusionFileType패턴 • exclusionFilePath패턴	Alfresco 데이터 소스에서 특정 파일을 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
type	데이터 소스의 유형. ALFRESCO을 데이터 소스 유형으로 지정합니다.

secretArn	연결에 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀의 Amazon 리소스 이름 (ARN). Alfresco 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다.
	
	기본 인증을 사용하는 경우:
	
	`{ "username": " user name", "password": " password" }`
	
	OAuth 2.0 인증을 사용하는 경우
	
	`{ "clientId": " client ID", "clientSecret": " client secret", "tokenUrl": " token URL" }`

구성	설명
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
enableIdentityCrawler	true Amazon Kendra의 ID 크롤러를 사용하여 특정 문서에 대한 액세스 권한이 있는 사용자 및 그룹의 ID/주체 정보를 동기화하는 데 사용합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
version	현재 지원되는 이 템플릿의 버전.

Alfresco JSON 스키마

```
{
  "$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 (MySQL) 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration객체의 일부로 포함합니다. 데이터 소스로 JDBC, 데이터 소스 유형으로 mysql, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Aurora (MySQL) JSON 스키마를 참조하세요.

다음 표는 Aurora (MySQL) JSON 스키마의 매개 변수를 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DBType - 사용하는 Java 데이터베이스의 유형 (,,`mysql`, `db2` 또는) `postgresql` `oracle` `sqlserver` • dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니

구성	설명
	다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. `JDBC`을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

Aurora (MySQL) JSON 스키마

```
{
  "$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 (PostgreSQL) 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 객체의 일부로 포함합니다. TemplateConfiguration 데이터 소스로 JDBC, 데이터 소스 유형으로 postgresql, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Aurora (PostgreSQL) JSON 스키마를 참조하세요.

다음 표에서는 Aurora (PostgreSQL) JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DbType - 사용하는 자바 데이터베이스 유형 (,,, 또는) `mysql` `db2` `postgresql` `oracle` `sqlserver` • dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니

구성	설명
	다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. JDBC을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

Aurora (PostgreSQL) JSON 스키마

```
{
  "$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 (Microsoft SQL Server) 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration개체의 일부로 포함합니다. 데이터 소스로 JDBC, 데이터 소스 유형으로 sqlserver, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Amazon RDS (마이크로소프트 SQL 서버) JSON 스키마를 참조하세요.

다음 표에서는 Amazon RDS (Microsoft SQL Server) JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DBType - 사용하는 Java 데이터베이스의 유형 (,, `mysql` `db2`, `postgresql` 또는) `oracle` `sqlserver` • dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니

구성	설명
	다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. JDBC을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

Amazon RDS (마이크로소프트 SQL 서버) JSON 스키마

```
{
  "$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 (MySQL) 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration객체의 일부로 포함합니다. 데이터 소스로 JDBC, 데이터 소스 유형으로 mysql, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Amazon RDS (MySQL) JSON 스키마를 참조하세요.

다음 표는 Amazon RDS (MySQL) JSON 스키마의 매개 변수를 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DBType - 사용하는 Java 데이터베이스의 유형 (,,`mysql`, `db2` 또는) `postgresql` `oracle` `sqlserver` • dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니

구성	설명
	다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. JDBC을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

Amazon RDS (MySQL) JSON 스키마

```
{
  "$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 (오라클) 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration객체의 일부로 포함합니다. 데이터 소스로 JDBC, 데이터 소스 유형으로 oracle, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Amazon RDS (Oracle) JSON 스키마를 참조하세요.

다음 표에서는 Amazon RDS (Oracle) JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DBType - 사용하는 Java 데이터베이스 유형 (,, `mysql`, `db2`, `postgresql` 또는) `oracle` `sqlserver` • dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니

구성	설명
	다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. JDBC을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

Amazon RDS (Oracle) JSON 스키마

```
{
  "$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 (PostgreSQL) 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 객체의 일부로 포함합니다. TemplateConfiguration 데이터 소스로 JDBC, 데이터 소스 유형으로 postgresql, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Amazon RDS (PostgreSQL) JSON 스키마를 참조하세요.

다음 표에서는 Amazon RDS (PostgreSQL) JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DbType - 사용하는 자바 데이터베이스 유형 (,,, 또는) `mysql` `db2` `postgresql` `oracle` `sqlserver` • dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니

구성	설명
	다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. JDBC을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

Amazon RDS (PostgreSQL) JSON 스키마

```
{
  "$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 템플릿 스키마

데이터 소스 스키마를 템플릿 구성의 일부로 포함하는 JSON을 포함합니다. S3 버킷의 이름을 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 제공합니다. 또한 데이터 소스 유형으로 S3, 기타 필수 구성을 지정합니다. 그런 다음 호출 TEMPLATE Type 시점으로 지정합니다 CreateDataSource.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. S3 JSON 스키마를 참조하세요.

다음 표에서는 Amazon S3 JSON 스키마의 파라미터에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
BucketName	Amazon S3 버킷 이름.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
• inclusionPatterns • exclusionPatterns • inclusionPrefixes • exclusionPrefixes	Amazon S3 데이터 원본에 특정 파일을 포함하거나 제외하기 위한 정규 표현식 패턴 목록입니다. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
aclConfigurationFile경로	Amazon Kendra 인덱스의 문서에 대한 액세스를 제어하는 파일 경로.
metadataFilesPrefix	버킷 내 메타데이터 파일 위치.
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠

구성	설명
	변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
type	데이터 소스의 유형. S3을 데이터 소스 유형으로 지정합니다.
version	지원되는 템플릿의 버전.

S3 JSON 스키마

```
{
  "$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 웹 크롤러 템플릿 스키마

데이터 소스 스키마를 TemplateConfiguration 객체의 일부로 포함하는 JSON을 포함합니다.

연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 시드 또는 시작 지점 URL을 제공하거나 사이트맵 URL을 제공할 수 있습니다. 모든 URL을 수동으로 나열하는 대신 시드 URL 목록의 텍스트 파일 또는 사이트맵 XML 파일을 저장하는 Amazon S3 버킷의 경로를 제공할 수 있습니다. 이 파일은 S3에서 ZIP 파일로 묶을 수 있습니다.

또한 데이터 소스 유형을 로 지정하고`WEBCRAWLERV2`, 웹 사이트에 인증이 필요한 경우 웹 사이트 인증 자격 증명 및 인증 유형, 기타 필요한 구성을 지정합니다.

그런 다음 CreateDataSource를 호출할 때 `TEMPLATE`을 `Type`으로 지정합니다.

> **Important**
>
> Web Crawler v2.0 커넥터 생성은 에서 지원되지 않습니다. AWS CloudFormation지원이 필요한 경우 웹 크롤러 v1.0 커넥터를 사용하십시오. AWS CloudFormation

인덱싱할 웹 사이트를 선택할 때 Amazon 이용 정책과 기타 모든 Amazon 약관을 준수해야 합니다. Amazon Kendra 웹 크롤러는 자신의 웹 페이지 또는 인덱싱할 권한이 있는 웹 페이지를 인덱싱할 때만 사용해야 한다는 점을 기억하십시오. Amazon Kendra Web Crawler가 웹 사이트를 인덱싱하지 못하게 하는 방법을 알아보려면 Amazon Kendra 웹 크롤러용 *robots.txt* 파일 구성을 참조하세요.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Amazon Kendra 웹 크롤러 JSON 스키마를 참조하세요.

다음 표에서는 Amazon Kendra Web Crawler JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
siteMapUrls	크롤링하려는 웹 사이트의 사이트 맵 URL의 목록입니다. 최대 3개의 사이트 맵 URL을 나열할 수 있습니다.

구성	설명
s3 SeedUrl	시드 또는 시작 지점 URL의 목록을 저장하는 텍스트 파일의 S3 경로입니다. 예를 들어 s3://bucket-name/directory/입니다. 텍스트 파일의 각 URL은 별도의 줄에 형식을 지정해야 합니다. 최대 100개의 시드 URL을 한 파일에 나열할 수 있습니다.
s3 SiteMapUrl	사이트맵 XML 파일의 S3 경로. 예를 들어 s3://bucket-name/directory/입니다. 최대 3개의 사이트 맵 XML 파일을 나열할 수 있습니다. 여러 사이트맵 파일을 ZIP 파일로 묶어 Amazon S3 버킷에 저장할 수 있습니다.
seedUrlConnections	크롤링하려는 웹사이트의 시드 또는 시작점 URL 목록. 최대 100개의 시드 URL을 나열할 수 있습니다.
seedUrl	시드 또는 시작 지점 URL.
authentication	웹 사이트에 동일한 인증이 필요한 경우의 인증 유형이며, 그렇지 않으면 `NoAuthentication` 를 지정합니다.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• webPage • attachment	웹 페이지 및 웹 페이지 파일의 속성 또는 필드 이름을 매핑하여 필드 이름을 Amazon Kendra 인덱싱하는 객체 목록입니다. 예를 들어, HTML 웹 페이지 제목 태그를 `_document_title` 인덱스 필드에 매핑할 수 있습니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
rateLimit	1분간 웹 사이트 호스트당 크롤링되는 최대 URL 수입니다.
maxFileSize	크롤링할 웹 페이지나 첨부 파일의 최대 크기(MB 단위)입니다.
crawlDepth	시드 URL에서 크롤링할 레벨 수. 예를 들어 시드 URL 페이지는 깊이 1이고 이 페이지에서 크롤링되는 모든 하이퍼링크는 깊이 2입니다.
maxLinksPerUrl	웹 사이트를 크롤링할 때 포함시킬 웹 페이지의 최대 URL 수입니다. 이 수는 웹 페이지를 기준으로 합니다. 웹 사이트의 웹 페이지가 크롤링되면 웹 페이지가 링크하는 모든 URL도 크롤링됩니다. 웹 페이지의 URL은 표시되는 순서대로 크롤링됩니다.

구성	설명
crawlSubDomain	`true`: 하위 도메인이 있는 웹 사이트 도메인을 크롤링합니다. 예를 들어, 시드 URL이 “abc.example.com“이면" “a.abc.example.com “ 및 “b.abc.example.com“도 크롤링됩니다. `crawlSubDomain` 또는 `crawlAllDomain` 로 `true` 설정하지 않으면 크롤링하려는 웹 사이트의 Amazon Kendra 도메인만 크롤링합니다.
crawlAllDomain	`true`: 웹 페이지가 연결되는 하위 도메인과 기타 도메인이 있는 웹 사이트 도메인을 크롤링합니다. `crawlSubDomain` 또는 `crawlAllDomain` 로 설정하지 않으면 크롤링하려는 `true` 웹 사이트의 Amazon Kendra 도메인만 크롤링합니다.
honorRobots	`true`: 크롤링하려는 웹 사이트의 robots.txt 지침을 준수합니다. 이러한 지침은 Amazon Kendra Web Crawler가 웹 사이트를 크롤링하는 방식(특정 콘텐츠만 Amazon Kendra 크롤링할 수 있는지 또는 어떤 콘텐츠도 크롤링하지 않는지 여부) 을 제어합니다.
crawlAttachments	`true`: 웹 페이지가 링크된 파일을 크롤링합니다.
• 포함 URL CrawlPatterns • 포함 URL IndexPatterns	특정 URL을 크롤링하고 해당 URL 웹 페이지의 하이퍼링크를 인덱싱하는 것을 포함하는 정규 표현식 패턴 목록. 패턴과 일치하는 URL은 인덱스에 포함됩니다. 패턴과 일치하지 않는 URL은 인덱스에서 제외됩니다. URL이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되며 URL/웹사이트의 웹 페이지는 인덱스에 포함되지 않습니다.

구성	설명
• 제외 URL CrawlPatterns • 제외 URL IndexPatterns	특정 URL을 크롤링하고 해당 URL 웹 페이지의 하이퍼링크를 인덱싱하는 것을 제외하는 정규 표현식 패턴 목록. 패턴과 일치하는 URL은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 URL은 인덱스에 포함됩니다. URL이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되며 URL/웹사이트의 웹 페이지는 인덱스에 포함되지 않습니다.
inclusionFileIndex패턴	특정 웹 페이지 파일을 포함하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
exclusionFileIndex패턴	특정 웹 페이지 파일을 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
proxy	웹 프록시를 통해 내부 웹 사이트에 연결하는 데 필요한 구성 정보입니다.
host	내부 웹사이트에 연결하는 데 사용하려는 프록시 서버의 호스트 이름. 예를 들어, https://a.example.com/page1.html의 호스트 이름은 “a.example.com”입니다.
포트	내부 웹사이트에 연결하는 데 사용하려는 프록시 서버의 포트 번호. 예를 들어, 443은 HTTPS의 표준 포트입니다.

구성	설명
secretArn (proxy)	웹 사이트 호스트에 연결하는 데 웹 프록시 자격 증명이 필요한 경우 자격 증명을 저장하는 AWS Secrets Manager 암호를 만들 수 있습니다. 암호의 Amazon 리소스 이름(ARN)을 제공합니다.
type	데이터 소스의 유형. `WEBCRAWLERV2` 을 데이터 소스 유형으로 지정합니다.

구성	설명
secretArn	웹 사이트에서 웹 사이트에 액세스하기 위해 인증이 필요한 경우 사용되는 AWS Secrets Manager 암호의 Amazon 리소스 이름 (ARN). JSON 키-값 쌍이 포함된 암호에 웹 사이트의 보안 인증을 저장합니다. 기본 또는 NTML/Kerberos를 사용하는 경우 사용자 이름 및 암호를 입력합니다. 암호의 JSON 키는 `userName` 및 `password`여야 합니다. NTLM 인증 프로토콜에는 암호 해싱이 포함되고 Kerberos 인증 프로토콜에는 암호 암호화가 포함됩니다. SAML 또는 양식 인증을 사용하는 경우 사용자 이름과 암호를 입력하고, 사용자 이름 필드에 XPath(SAML을 사용하는 경우 사용자 이름 버튼), 암호 필드와 버튼에 XPaths, 로그인 페이지 URL을 입력합니다. 암호의 JSON 키는 `userName`, `password`, `userNameFieldXpath` , `userNameButtonXpath` , `passwordFieldXpath` , `passwordButtonXpath` , `loginPageUrl` 이어야 합니다. 웹 브라우저의 개발자 도구를 사용하여 요소의 XPaths(XML 경로 언어)를 찾을 수 있습니다. XPaths는 일반적으로 다음 형식을 따릅니다. `//tagname[@Attribute='Value']` Amazon Kendra 또한 시크릿에 포함된 엔드포인트 정보 (시드 URL) 가 데이터 소스 엔드포인트 구성 세부 정보에 지정된 엔드포인트 정보와 동일한지 확인합니다.
version	현재 지원되는 이 템플릿의 버전.

Amazon Kendra 웹 크롤러 JSON 스키마

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

Confluence 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 개체의 일부로 포함합니다. TemplateConfiguration 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 Confluence 호스트 URL, 호스팅 방법 및 인증 유형을 제공합니다. 또한 데이터 소스 유형으로 CONFLUENCEV2, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Confluence JSON 스키마를 참조하세요.

다음 표에서는 Confluence JSON 스키마의 파라미터에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
hostUrl	Confluence 인스턴스의 URL. 예: *https://example.confluence.com* .
type	Confluence 인스턴스의 호스팅 방법(SAAS 및 ON_PREM)
authType	Confluence 인스턴스의 인증 방법(예: Basic, OAuth2 또는 Personal-token)

구성	설명
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• space • 페이지를 방문하십시오 • blog • 설명 • attachment	Confluence 스페이스, 페이지, 블로그, 댓글 및 첨부 파일의 속성 또는 필드 이름을 인덱스 필드 이름에 매핑하는 객체 목록입니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요. Confluence 데이터 소스 필드 이름은 Confluence 사용자 지정 메타데이터에 있어야 합니다.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
isCrawlAcl	`true`ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 크롤링하십시오. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
fieldForUserId	사용자 ID로 사용자 이메일을 사용할지 `email` 여부를 지정하십시오. `email`기본적으로 사용되며 현재 지원되는 유일한 사용자 ID 유형입니다.

구성	설명
• inclusionSpaceKey필터 • exclusionSpaceKey필터 • pageTitleRegEX • blogTitleRegEX • commentTitleRegEX • attachmentTitleRegEX • inclusionFileType패턴 • exclusionFileType패턴 • inclusionUrlPatterns • exclusionUrlPatterns	Confluence 데이터 소스에서 특정 파일을 포함 및/또는 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
proxyHost	사용하는 웹 프록시의 호스트 이름 (`http://`또는 `https://` 프로토콜 제외).
proxyPort	호스트 URL 전송 프로토콜에서 사용하는 포트 번호입니다. 이 값은 0~65535의 숫자 값이어야 합니다.
• isCrawlPersonal스페이스 • isCrawlArchived우주 • isCrawlArchived페이지 • isCrawlPage • isCrawlBlog • isCrawlPage코멘트 • isCrawlPage첨부파일 • isCrawlBlog코멘트 • isCrawlBlog첨부파일	`true`Confluence 개인 공간, 페이지, 블로그, 페이지 댓글, 페이지 첨부 파일, 블로그 댓글, 블로그 첨부 파일의 파일을 크롤링할 수 있습니다.

구성	설명
maxFileSizeInMegaBytes	크롤링할 수 있는 파일 크기 제한을 MB 단위로 지정하십시오. Amazon Kendra Amazon Kendra 정의한 크기 제한 내에 있는 파일만 크롤링합니다. 기본 파일 크기는 50MB입니다. 최대 파일 크기는 0MB보다 크고 50MB보다 작거나 같아야 합니다.
type	데이터 소스의 유형. `CONFLUENCEV2` 을 데이터 소스 유형으로 지정합니다.
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 특정 문서에 대한 액세스 권한이 있는 사용자 및 그룹의 ID/주체 정보를 동기화하는 데 사용합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • `FORCED_FULL_CRAWL` 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • `FULL_CRAWL` 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

구성	설명
secretARN	Confluence에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 시크릿의 Amazon 리소스 이름 (ARN). 이러한 키-값 쌍에 대한 자세한 내용은 Confluence의 연결 지침을 참조하십시오.
version	현재 지원되는 이 템플릿의 버전.

Confluence JSON 스키마

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

Dropbox 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 개체의 일부로 포함합니다. TemplateConfiguration 보안 인증을 저장하는 암호의 일부로 Dropbox 앱 키, 앱 암호, 액세스 토큰을 제공합니다. 또한 데이터 소스 유형으로 DROPBOX를 지정하고, 사용하려는 액세스 토큰 유형(임시 또는 영구) 및 기타 필요한 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Dropbox JSON 스키마를 참조하세요.

다음 표에는 Dropbox JSON 스키마의 매개 변수가 설명되어 있습니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보. 이 데이터 소스는 repositoryEndpointMetadata 에 엔드포인트를 지정하지 않습니다. 대신 사용자가

구성	설명
	제공하는 AWS Secrets Manager 비밀번호에 연결 정보가 포함됩니다. `secretArn`
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• 파일 • paper • papert • shortcut	Dropbox 파일, Dropbox Paper의 속성 또는 필드 이름을 매핑하고 필드 이름을 Amazon Kendra 인덱싱하기 위한 바로가기를 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
syncMode	데이터 원본 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 명시하세요. 다음 중에서 선택할 수 있습니다. • `FORCED_FULL_CRAWL` 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • `FULL_CRAWL` 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • `CHANGE_LOG` 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

구성	설명
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 특정 문서에 대한 액세스 권한이 있는 사용자 및 그룹의 ID/주체 정보를 동기화하는 데 사용합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
secretARN	Dropbox에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀번호의 Amazon 리소스 이름 (ARN). 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "appKey": "Dropbox app key", "appSecret": " Dropbox app secret", "accesstoken": " temporary access token or refresh access token" }`
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
isCrawlAcl	`true`ACL이 있고 액세스 제어에 사용하려는 경우 문서의 액세스 제어 목록 (ACL) 정보를 크롤링하기 위해서입니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

구성	설명
• inclusionFileName패턴 • inclusionFileType패턴	Dropbox 데이터 소스에서 특정 파일 이름 및 형식을 포함하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
• exclusionFileName패턴 • exclusionFileType패턴	Dropbox 데이터 소스에서 특정 파일 이름 및 형식을 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 제외 및 포함 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
• crawlFile • crawlPaper • crawlPapert • crawlShortcut	`true`드롭박스에 있는 파일, 드롭박스 페이퍼 문서, 드롭박스 페이퍼 템플릿, 드롭박스에 저장된 웹페이지 바로가기를 크롤링할 수 있습니다.
type	데이터 소스의 유형. `DROPBOX`을 데이터 소스 유형으로 지정합니다.
tokenType	액세스 토큰 유형(영구 또는 임시 액세스 토큰)을 지정합니다. 4시간 후에 만료되는 일회용 액세스 토큰을 사용하는 것보다 Dropbox에서 만료되지 않는 새로 고침 액세스 토큰을 만드는 것이 좋습니다. Dropbox 개발자 콘솔에서 앱과 새로 고침 액세스 토큰을 만들고 암호에 액세스 토큰을 입력합니다.
version	현재 지원되는 이 템플릿의 버전.

Dropbox JSON 스키마

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

Drupal 템플릿 스키마

데이터 소스 스키마가 포함된 JSON을 개체의 일부로 포함합니다. TemplateConfiguration 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 Drupal 호스트 URL과 인증 유형을 제공합니다. 또한 데이터 소스 유형으로 DRUPAL, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Drupal JSON 스키마를 참조하세요.

다음 표는 Drupal JSON 스키마의 매개 변수를 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
hostUrl	Drupal 웹 사이트의 호스트 URL. <hostname ><drupalsitename>예를 들어 *https:///####*.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보.
• content • 설명 • attachment	Drupal 파일의 속성 또는 필드 이름을 매핑하는 객체 목록. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요. Drupal 데이터 소스

구성	설명
	필드 이름이 Drupal 사용자 지정 메타데이터에 있어야 합니다.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
• inclusionFileName패턴 • articleTitleInclusion패턴 • pageTitleInclusion패턴 • customContentTitleInclusionPatterns • basicBlockTitleInclusionPatterns • customBlockTitleInclusionPatterns	Drupal 데이터 소스에서 특정 파일을 포함하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
• exclusionFileName패턴 • articleTitleExclusion패턴 • pageTitleExclusion패턴 • customContentTitleExclusionPatterns • basicBlockTitleExclusionPatterns • customBlockTitleExclusionPatterns	Drupal 데이터 소스에서 특정 파일을 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 제외 및 포함 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
contentDefinitions • contentType • fieldDefinition • isCrawlComments • isCrawlFiles • isCrawlArticle • isCrawlBasic페이지 • isCrawlBasic블록 • isCrawlCustomContentTypesList	크롤링할 콘텐츠 유형, 선택한 콘텐츠 유형에 대한 설명 및 첨부 파일을 크롤링할지 여부를 지정합니다.
type	데이터 소스의 유형. DRUPAL을 데이터 소스 유형으로 지정합니다.

구성	설명
authType	사용하는 인증 유형(BASIC-AUTH 또는 OAUTH2).
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
enableIdentityCrawler	true Amazon Kendra의 ID 크롤러를 사용하여 특정 문서에 대한 액세스 권한이 있는 사용자 및 그룹의 ID/주체 정보를 동기화하는 데 사용합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

구성	설명

secretARN	Drupal에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀의 Amazon 리소스 이름 (ARN). 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다.
	
	기본 인증을 사용하는 경우:
	
	`{ "username": "user name", "passwords": "password" }`
	
	OAuth 2.0 인증을 사용하는 경우:
	
	`{ "username": "user name", "password": "password" , "clientId": "client id", "clientSecret": "client secret" }`

version	현재 지원되는 이 템플릿의 버전.

Drupal JSON 스키마

```
{
 "$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 템플릿 스키마

데이터 소스 스키마를 TemplateConfiguration개체의 일부로 포함하는 JSON을 포함합니다. 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 GitHub 호스트 URL, 조직 이름, GitHub 클라우드 또는 GitHub 온프레미스 사용 여부를 제공합니다. 또한 데이터 소스 유형으로 GITHUB, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. GitHub JSON 스키마를 참조하세요.

다음 표에서는 GitHub JSON 스키마의 파라미터에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.

구성	설명
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
type	유형을 SAAS 또는 ON_PREMISE 로 지정합니다.
hostUrl	GitHub 호스트 URL. 예를 들어, GitHub SaaS/엔터프라이즈 클라우드를 사용하는 경우:. https://api.github.com 또는 GitHub 온프레미스/엔터프라이즈 서버를 사용하는 경우:. https://on-prem-host-url/api/v3/
organizationName	GitHub 데스크톱에 로그인하고 프로필 사진 드롭다운에서 내 조직으로 이동하면 조직 이름을 찾을 수 있습니다.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• GH리포지토리 • GH 커밋 • ghIssueDocument • ghIssueComment • ghIssueAttachment • GPR 문서 • GPR 논평 • GPR 첨부 파일	GitHub 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.

구성	설명
isCrawlAcl	`true`문서에 대한 액세스 제어 목록 (ACL) 정보를 크롤링하려면 (ACL이 있고 액세스 제어에 사용하려는 경우) ACL은 사용자와 그룹이 액세스하고 검색할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
fieldForUserId	ACL 크롤링에 사용할 사용자 ID 유형을 지정합니다. 사용자 ID로 사용자 이메일을 사용할지, 아니면 `username` 사용자 ID로 사용자 이름을 사용할지를 지정합니다. `email` 옵션을 지정하지 않는 경우 기본적으로 `email` 이 옵션이 사용됩니다.
리포지토리 필터	인덱싱하려는 특정 리포지토리 이름 및 브랜치 이름 목록.
크롤/리포지토리	`true`리포지토리를 크롤링하는 방법.
crawlRepositoryDocuments	`true`리포지토리 문서를 크롤링합니다.
크롤링 문제	`true`크롤링 이슈를 위해서요.
crawlIssueComment	`true`이슈 댓글을 크롤링하기 위해서요.
crawlIssueComment첨부 파일	`true`댓글 첨부 파일을 크롤링하려면
crawlPullRequest	`true`풀 리퀘스트를 크롤링하기 위해서요.
crawlPullRequest코멘트	`true`풀 리퀘스트 댓글을 크롤링하려면
crawlPullRequestCommentAttachment	`true`풀 리퀘스트 댓글 첨부 파일을 크롤링하기 위해서입니다.

구성	설명
• inclusionFolderName패턴 • inclusionFileType패턴 • inclusionFileName패턴	GitHub데이터 원본에 특정 콘텐츠를 포함하기 위한 정규 표현식 패턴 목록입니다. 패턴과 일치하는 콘텐츠는 인덱스에 포함됩니다. 패턴과 일치하지 않는 콘텐츠는 인덱스에서 제외됩니다. 포함 패턴과 제외 패턴 모두에 일치하는 콘텐츠가 있는 경우 제외 패턴이 우선하며 콘텐츠는 색인에 포함되지 않습니다.
• exclusionFolderName패턴 • exclusionFileType패턴 • exclusionFileName패턴	GitHub데이터 원본에서 특정 콘텐츠를 제외하기 위한 정규 표현식 패턴 목록입니다. 패턴과 일치하는 콘텐츠는 인덱스에서 제외됩니다. 패턴과 일치하지 않는 콘텐츠는 인덱스에 포함됩니다. 포함 패턴과 제외 패턴 모두에 일치하는 콘텐츠가 있는 경우 제외 패턴이 우선하며 콘텐츠는 색인에 포함되지 않습니다.
type	데이터 소스의 유형. `GITHUB`을 데이터 소스 유형으로 지정합니다.
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 사용자 및 그룹의 ID/주체 정보를 특정 문서에 대한 액세스 권한과 동기화하는 데 사용합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

구성	설명
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화 될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화 될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	연결에 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀의 Amazon 리소스 이름 (ARN). GitHub 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "personalToken": " token" }`
version	현재 지원되는 이 템플릿의 버전입니다.

GitHub JSON 스키마

다음은 GitHub JSON 스키마입니다.

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

Gmail 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 개체의 TemplateConfiguration일부로 포함합니다. 데이터 소스 유형으로 GMAIL, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Gmail JSON 스키마를 참조하세요.

다음 표에는 Gmail JSON 스키마의 매개변수가 설명되어 있습니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보. 이 데이터 소스는 repositoryEndpointMetadata 에 엔드포인트를 지정하지 않습니다. 대신 사용자가 제공하는 AWS Secrets Manager 비밀번호에 연결 정보가 포함됩니다. secretArn
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
• message • attachments	Gmail 메일과 첨부파일의 속성 또는 필드 이름을 Amazon Kendra 색인 필드 이름에 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

구성	설명
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
• inclusionLabelName패턴 • exclusionLabelName패턴 • inclusionAttachmentType패턴 • exclusionAttachmentType패턴 • inclusionAttachmentName패턴 • exclusionAttachmentName패턴 • inclusionSubjectFilter • exclusionSubjectFilter • isSubjectAnd • inclusionFromFilter • exclusionFromFilter • inclusionToFilter • exclusionToFilter • inclusionCcFilter • exclusionCcFilter • inclusionBccFilter • exclusionBccFilter	Gmail 데이터 소스에서 특정 주제 이름의 메시지를 포함하거나 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
beforeDateFilter	특정 날짜 이전에 포함할 메시지와 첨부 파일을 지정합니다.
afterDateFilter	특정 날짜 이후에 포함할 메시지와 첨부 파일을 지정합니다.
isCrawlAttachment	첨부 파일을 크롤링할지 여부를 선택하는 부울 값. 메시지는 자동으로 크롤링됩니다.
type	데이터 소스의 유형. `GMAIL`을 데이터 소스 유형으로 지정합니다.

구성	설명
shouldCrawlDraft메시지	초안 메시지를 크롤링할지 여부를 선택하는 부울 값.
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. **Important** 영구 삭제된 Gmail 메시지를 업데이트하는 API가 없으므로 새 콘텐츠, 수정 콘텐츠, 삭제된 콘텐츠가 동기화됩니다. • Gmail에서 영구 삭제된 메일은 색인에서 삭제되지 않습니다. Amazon Kendra • Gmail 이메일 라벨의 변경사항을 동기화하지 않습니다. Gmail 데이터 소스 라벨 변경사항과 영구 삭제된 이메일 메시지를 Amazon Kendra 색인에 동기화하려면 정기적으로 전체 크롤링을 실행해야 합니다.

구성	설명
secretARN	Gmail에 연결하는 데 필요한 키-값 페어가 포함된 Secrets Manager 암호의 Amazon 리소스 이름(ARN). 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "adminAccountEmailId": " service account email", "clientEmailId": " user account email", "privateKey": " private key" }`
version	현재 지원되는 템플릿의 버전.

Gmail JSON 스키마

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

Google Drive 템플릿 스키마

데이터 소스 스키마가 포함된 JSON을 개체의 일부로 포함하세요. TemplateConfiguration 데이터 소스 유형으로 GOOGLEDRIVE2, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Google Drive JSON 스키마를 참조하세요.

다음 표에는 Google 드라이브 JSON 스키마의 매개변수가 설명되어 있습니다.

구성	설명
connectionConfiguration	데이터 소스에 대한 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보. 이 데이터 소스는 엔드포인트를 지정하지 않습니다. 인증 유형: serviceAccount 및 OAuth2를 선택합니다. 연결 정보는 사용자가 제공하는 AWS Secrets Manager 비밀번호에 포함됩니다. secretArn
authType	사용 사례를 기반으로 serviceAccount 또는 OAuth2 중에서 선택합니다.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• 파일 • 설명	Amazon Kendra 인덱스 필드 이름에 Google Drive의 속성 또는 필드 이름을 매핑하는 객체의 목록. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.

구성	설명
• maxFileSizeInMegaBytes	Amazon Kendra 크롤링해야 하는 파일 크기 제한을 MB 단위로 지정합니다.
• iscrawlComment	`true`Google 드라이브 데이터 소스의 댓글을 크롤링하려면
• isCrawlMyDriveAndSharedWithMe	`true`Google 드라이브 데이터 소스의 드라이브를 크롤링하여 내 MyDrive 드라이브와 공유하기
• isCrawlShared드라이브	`true`Google 드라이브 데이터 소스의 공유 드라이브를 크롤링하기.
isCrawlAcl	`true`ACL이 있고 액세스 제어에 사용하려는 경우 문서의 ACL (액세스 제어 목록) 정보를 크롤링할 수 있습니다. ACL은 사용자와 그룹이 액세스하고 검색할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
• excludeUserAccounts • excludeSharedDrives • excludeMimeTypes • exclusionFileType패턴 • exclusionFileName패턴 • exclusionFilePath필터	Google Drive 데이터 소스에서 특정 파일을 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
• includeUserAccounts • includeSharedDrives • includeMimeTypes • inclusionFileType패턴 • inclusionFileName패턴 • inclusionFilePath필터	Google Drive 데이터 소스에서 특정 파일을 포함하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.

구성	설명
type	데이터 소스의 유형. `GOOGLEDRIVEV2` 을 데이터 소스 유형으로 지정합니다.
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 사용자 및 그룹의 ID/주체 정보를 특정 문서에 대한 액세스 권한과 동기화합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • `FORCED_FULL_CRAWL` 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • `FULL_CRAWL` 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • `CHANGE_LOG` 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

구성	설명

secretARN	Google 드라이브에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀의 Amazon 리소스 이름 (ARN). 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다.
	
	Google 서비스 계정 인증을 사용하는 경우:
	
	`{ "clientEmail": " user account email", "adminAccountEmail": " service account email", "privateKey": " private key" }`
	
	OAuth 2.0 인증을 사용하는 경우:
	
	`{ "clientID": " OAuth client ID", "clientSecret": " client secret", "refreshToken": " refresh token" }`

version	현재 지원되는 이 템플릿의 버전.

Google Drive JSON 스키마

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

IBM DB2 템플릿 스키마

데이터 소스 스키마가 포함된 JSON을 객체의 일부로 포함합니다. TemplateConfiguration 데이터 소스로 JDBC, 데이터 소스 유형으로 db2, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. IBM DB2 JSON 스키마를 참조하세요.

다음 표에는 IBM DB2 JSON 스키마의 매개변수가 설명되어 있습니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DBType - 사용하는 자바 데이터베이스 유형 (,,mysql, db2 또는). postgresql oracle sqlserver • dbHost - 데이터베이스 호스트 이름.

구성	설명
	• dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.

구성	설명
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. `JDBC`을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

IBM DB2 JSON 스키마

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

Microsoft Exchange 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration개체의 일부로 포함합니다. 테넌트 ID는 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 제공합니다. 또한 데이터 소스 유형으로 MSEXCHANGE, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Microsoft Exchange JSON 스키마를 참조하세요.

다음 표에서는 Microsoft Exchange JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
tenantId	Microsoft 365 테넌트 ID. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• 이메일 • attachment • calendar • contacts • notes	Microsoft Exchange 데이터 원본의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드에 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
inclusionPatterns	Microsoft Exchange 데이터 소스에서 특정 파일을 포함하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
exclusionPatterns	Microsoft Exchange 데이터 소스에서 특정 파일을 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 제외 및 포함 패턴과 모두 일치하는 경우 제외 패

구성	설명
	턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
• inclusionUsersList • inclusionUsersFile이름 • inclusionDomainUsers	Microsoft Exchange 데이터 소스에서 특정 사용자 및 사용자 파일을 포함하는 정규식 패턴 목록. 패턴과 일치하는 사용자는 인덱스에 포함됩니다. 패턴과 일치하지 않는 사용자는 인덱스에서 제외됩니다. 사용자가 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 사용자는 인덱스에 포함되지 않습니다.
• exclusionUsersList • exclusionUsersFile이름 • exclusionDomainUsers	Microsoft Exchange 데이터 소스에서 특정 사용자 및 사용자 파일을 제외하는 정규식 패턴 목록. 패턴과 일치하는 사용자는 인덱스에서 제외됩니다. 패턴과 일치하지 않는 사용자는 인덱스에 포함됩니다. 사용자가 제외 및 포함 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 사용자는 인덱스에 포함되지 않습니다.
s3bucketName	사용할 S3 버킷의 이름.
• crawlCalendar • crawlNotes • crawlContacts • crawlFolderAcl	`true`Microsoft Exchange 데이터 원본의 이러한 유형의 콘텐츠를 크롤링하고 제어 정보에 액세스할 수 있습니다.
startCalendarDate소요 시간	일정 콘텐츠의 특정 시작 날짜/시간을 구성할 수 있습니다.
endCalendarDate시간	일정 콘텐츠의 특정 종료 날짜/시간을 구성할 수 있습니다.
subject	메일 콘텐츠의 특정 제목 줄을 구성할 수 있습니다.

구성	설명
emailFrom	'보낸 사람' 또는 보낸 사람 메일 콘텐츠에 대해 특정 이메일을 구성할 수 있습니다.
emailTo	'받는 사람' 또는 받는 사람 메일 콘텐츠에 대해 특정 이메일을 구성할 수 있습니다.
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
type	데이터 소스의 유형. MSEXCHANGE 을 데이터 소스 유형으로 지정합니다.
secretARN	Microsoft Exchange에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 입니다. 여기에는 Azure 포털에서 OAuth 애플리케이션을 만들 때 생성되는 클라이언트 ID와 클라이언트 암호가 포함됩니다.

구성	설명
version	현재 지원되는 이 템플릿의 버전.

Microsoft Exchange JSON 스키마

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "tenantId": {
              "type": "string",
              "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]
{12}$",
              "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"
  ]
}
```

Microsoft OneDrive 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration개체의 일부로 포함합니다. 테넌트 ID는 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 제공합니다. 또한 데이터 소스 유형으로 ONEDRIVEV2, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. 마이크로소프트 OneDrive JSON 스키마를 참조하세요.

다음 표에서는 Microsoft OneDrive JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
tenantId	Microsoft 365 테넌트 ID. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.

구성	설명
파일	Microsoft OneDrive 파일의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
• userNameFilter • userFilterPath • inclusionFileType패턴 • exclusionFileType패턴 • inclusionFileName패턴 • exclusionFileName패턴 • inclusionFilePath패턴 • exclusionFilePath패턴 • inclusionOneNoteSectionNamePatterns • exclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns • exclusionOneNotepageNamePatterns	특정 파일, OneNote 섹션, OneNote 페이지를 인덱싱하도록 선택하고 사용자 이름을 기준으로 필터링할 수 있습니다.
isUserNameOnS3	`true`: Amazon S3에 저장된 파일의 사용자 이름 목록을 제공합니다.
type	데이터 소스의 유형. `ONEDRIVEV2` 을 데이터 소스 유형으로 지정합니다.

구성	설명
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 사용자 및 그룹의 ID/주체 정보를 특정 문서에 대한 액세스 권한과 동기화합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
type	데이터 소스의 유형. `ONEDRIVEV2` 을 데이터 소스 유형으로 지정합니다.
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • `FORCED_FULL_CRAWL` 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • `FULL_CRAWL` 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • `CHANGE_LOG` 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

구성	설명
secretARN	Microsoft에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 시크릿의 Amazon 리소스 이름 (ARN) 입니다. OneDrive 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "clientId": " client ID", "clientSecret": " client secret" }`
version	현재 지원되는 이 템플릿의 버전.

마이크로소프트 OneDrive JSON 스키마

```
{
 "$schema": "http://json-schema.org/draft-04/schema#",
 "type": "object",
 "properties": {
  "connectionConfiguration": {
   "type": "object",
   "properties": {
    "repositoryEndpointMetadata": {
     "type": "object",
     "properties": {
      "tenantId": {
       "type": "string",
       "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}$",
       "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"
  ]
}
```

Microsoft SharePoint 템플릿 스키마

데이터 소스 스키마를 TemplateConfiguration개체의 일부로 포함하는 JSON을 포함합니다. SharePoint 사이트 URL/URL, 도메인 및 필요한 경우 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 테넌트 ID를 제공합니다. 또한 데이터 소스 유형으로 SHAREPOINTV2, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 호출할 TEMPLATE 때 유형으로 지정합니다. CreateDataSource

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. SharePoint JSON 스키마를 참조하세요.

다음 표에서는 Microsoft SharePoint JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보
tenantId	SharePoint 계정의 테넌트 ID.
도메인	SharePoint 계정의 도메인.
siteUrls	SharePoint 계정의 호스트 URL.
repositoryAdditionalProperties	리포지토리/데이터 소스 엔드포인트와 연결하기 위한 추가 속성.
s3bucketName	Azure AD 자체 서명 Amazon S3 X.509 인증서를 저장하는 버킷의 이름입니다.
s3certificateName	버킷에 저장된 Azure AD 자체 서명 X.509 인증서의 이름입니다. Amazon S3
authType	사용하는 인증 유형 (예:,,, `OAuth2` `OAuth2Certificate` `OAuth2App` , `Basic` 또는) `OAuth2_RefreshToken` `NTLM` `Kerberos`
version	사용하는 SharePoint 버전 (`Server`또는 `Online` 여부)
onPremVersion	사용하는 SharePoint 서버 버전 (예: `2013` `20162019`, 또는`SubscriptionEdition`)
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• 이벤트 • 페이지를 방문하십시오 • 파일	SharePoint콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는

구성	설명
• link • attachment • 설명	개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
• eventTitleFilterRegEx • pageTitleFilterRegEx • linkTitleFilterRegEx • inclusionFilePath • exclusionFilePath • inclusionFileType패턴 • exclusionFileType패턴 • inclusionFileName패턴 • exclusionFileName패턴 • inclusionOneNoteSectionNamePatterns • exclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns • exclusionOneNotePageNamePatterns	데이터 원본에 특정 콘텐츠를 포함/제외하기 위한 정규 표현식 패턴 목록입니다. SharePoint 포함 패턴과 일치하는 콘텐츠 항목이 색인에 포함됩니다. 포함 패턴과 일치하지 않는 콘텐츠 항목은 색인에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
• crawlFiles • crawlPages • crawlEvents • crawlComments • crawlLinks • crawlAttachments	`true`이러한 유형의 콘텐츠를 크롤링하기 위해서입니다.

구성	설명
crawlAcl	`true`ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 크롤링합니다. ACL은 사용자와 그룹이 액세스하고 검색할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
fieldForUserId	사용자 ID로 사용자 이메일을 사용할지, 아니면 `userPrincipalName` 사용자 ID로 사용자 이름을 사용할지를 지정하십시오. `email` 옵션을 지정하지 않는 경우 기본적으로 `email` 이 옵션이 사용됩니다.
aclConfiguration	`ACLWithLDAPEmailFmt` `ACLWithManualEmailFmt` , 또는 중 하나를 지정합니다`ACLWithUsernameFmtM` .
emailDomain	이메일의 도메인. 예: *“amazon.com”* .
• isCrawlLocalGroupMapping • isCrawlAdGroupMapping	`true`그룹 매핑 정보를 크롤링하기 위해서입니다.
proxyHost	http://또는 https://프로토콜을 사용하지 않고 사용하는 웹 프록시의 호스트 이름.
proxyPort	호스트 URL 전송 프로토콜에서 사용하는 포트 번호입니다. 이 값은 0~65535의 숫자 값이어야 합니다.
type	`SHAREPOINTV2` 를 데이터 소스 유형으로 지정합니다.

구성	설명
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 사용자 및 그룹의 ID/주체 정보를 특정 문서에 대한 액세스 권한과 동기화합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • `FORCED_FULL_CRAWL` 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • `FULL_CRAWL` 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • `CHANGE_LOG` 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretARN	연결에 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀의 Amazon 리소스 이름 (ARN). SharePoint 이러한 키-값 쌍에 대한 자세한 내용은 온라인 및 서버의 연결 지침을 참조하십시오. SharePoint SharePoint

구성	설명
version	현재 지원되는 이 템플릿의 버전.

SharePoint JSON 스키마

```
{
 "$schema": "http://json-schema.org/draft-04/schema#",
 "type": "object",
 "properties": {
  "connectionConfiguration": {
   "type": "object",
   "properties": {
    "repositoryEndpointMetadata": {
     "type": "object",
     "properties": {
      "tenantId": {
       "type": "string",
       "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}$",
       "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"
  ]
}
```

Microsoft SQL Server 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 개체의 TemplateConfiguration일부로 포함합니다. 데이터 소스로 JDBC, 데이터 소스 유형으로 sqlserver, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Microsoft SQL Server JSON 스키마를 참조하세요.

다음 표에서는 마이크로소프트 SQL Server JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DBType - 사용하는 Java 데이터베이스의 유형 (,,, 또는) `mysql` `db2` `postgresql` `oracle` `sqlserver` • dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니

구성	설명
	다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. `JDBC`을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

Microsoft SQL Server JSON 스키마

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

Microsoft Teams 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration개체의 일부로 포함합니다. 테넌트 ID는 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 제공합니다. 또한 데이터 소스 유형으로 MSTEAMS, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Microsoft Teams JSON 스키마를 참조하세요.

다음 표는 Microsoft Teams JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
tenantId	Microsoft 365 테넌트 ID. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• chatMessage • chatAttachment • channelPost • channelWiki • channelAttachment • meetingChat • meetingFile • meetingNote • calendarMeeting	Microsoft Teams 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
paymentModel	Microsoft Teams 데이터 소스와 함께 사용할 결제 모델 유형을 지정합니다. 모델 A 결제 모델은 보안 준수가 필요한 라이선싱 및 결제 모델로 제한됩니다. 모델 B 결제 모델은 보안 준수가 필요하지 않은 라이선싱 및 결제 모델에 적합합니다.
• inclusionTeamName필터 • inclusionChannelName필터 • inclusionFileName패턴	Microsoft Teams 데이터 소스에서 특정 콘텐츠를 포함하는 정규식 패턴 목록. 패턴과 일치하는 콘텐츠는 인덱스에 포함됩니다. 패턴과 일치

구성	설명
• inclusionFileType패턴 • inclusionUserEmail필터 • inclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns	하지 않는 콘텐츠는 인덱스에서 제외됩니다. 콘텐츠가 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 콘텐츠는 인덱스에 포함되지 않습니다.
• exclusionTeamName필터 • exclusionChannelName필터 • exclusionFileName패턴 • exclusionFileType패턴 • exclusionUserEmail필터 • exclusionOneNoteSectionNamePatterns • exclusionOneNotePageNamePatterns	Microsoft Teams 데이터 소스에서 특정 콘텐츠를 제외하는 정규식 패턴 목록. 패턴과 일치하는 콘텐츠는 인덱스에서 제외됩니다. 패턴과 일치하지 않는 콘텐츠는 인덱스에 포함됩니다. 콘텐츠가 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 콘텐츠는 인덱스에 포함되지 않습니다.
• isCrawlChat메시지 • isCrawlChat첨부파일 • isCrawlChannel포스트 • isCrawlChannel첨부 • isCrawlChannel위키 • isCrawlCalendar미팅 • isCrawlMeeting채팅 • isCrawlMeeting파일 • isCrawlMeeting노트	`true`Microsoft Teams 데이터 원본에서 이러한 유형의 콘텐츠를 크롤링할 수 있습니다.
startCalendarDate소요 시간	일정 콘텐츠의 특정 시작 날짜/시간을 구성할 수 있습니다.
endCalendarDate시간	일정 콘텐츠의 특정 종료 날짜/시간을 구성할 수 있습니다.
type	데이터 소스의 유형. `MSTEAMS`을 데이터 소스 유형으로 지정합니다.

구성	설명
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 사용자 및 그룹의 ID/주체 정보를 특정 문서에 대한 액세스 권한과 동기화합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • `FORCED_FULL_CRAWL` 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • `FULL_CRAWL` 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • `CHANGE_LOG` 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	Microsoft Teams에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 시크릿의 Amazon 리소스 이름 (ARN). 여기에는 Azure 포털에서 OAuth 애플리케이션을 만들 때 생성되는 클라이언트 ID와 클라이언트 암호가 포함됩니다.

구성	설명
version	현재 지원되는 이 템플릿의 버전.

Microsoft Teams JSON 스키마

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "tenantId": {
              "type": "string",
              "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]
{12}$",
              "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"
  ]
}
```

Microsoft Yammer 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 객체의 일부로 포함합니다. TemplateConfiguration 데이터 소스 유형으로 YAMMER, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 CreateDataSource호출할 때 TEMPLATE Type으로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다.

다음 표에서는 Microsoft Yammer JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스에 대한 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보. 이 데이터 소스는 repositoryEndpointMetadata 에 엔드포인트를 지정하지 않습니다. 대신 사용자가 제공하는 AWS Secrets Manager 암호에 연결 정보가 포함됩니다. secretArn
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.

구성	설명
• community • 사용자 • message • attachment	Amazon Kendra 인덱스 필드 이름에 Microsoft Yammer 콘텐츠의 속성 또는 필드 이름을 매핑하는 객체의 목록. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
inclusionPatterns	Microsoft Yammer 데이터 소스에서 특정 파일을 포함하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
exclusionPatterns	Microsoft Yammer 데이터 소스에서 특정 파일을 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 제외 및 포함 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
sinceDate	Microsoft Yammer 커넥터가 특정 `sinceDate` 기반으로 콘텐츠를 크롤링하도록 `sinceDate` 파라미터를 구성하도록 선택할 수 있습니다.
communityNameFilter	특정 커뮤니티 콘텐츠를 인덱싱하도록 선택할 수 있습니다.
• isCrawlMessage • isCrawlAttachment • isCrawlPrivate메시지	`true`메시지, 메시지 첨부 파일, 개인 메시지를 크롤링합니다.
type	`YAMMER`을 데이터 소스 유형으로 지정합니다.

구성	설명
secretARN	Microsoft Yammer에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 시크릿의 Amazon 리소스 이름 (ARN) 입니다. 여기에는 Azure 포털에서 OAuth 애플리케이션을 만들 때 생성되는 Microsoft Yammer 사용자 이름과 암호, 클라이언트 ID와 클라이언트 암호가 포함됩니다.
useChangeLog	`true`Microsoft Yammer 변경 로그를 사용하여 색인에서 업데이트가 필요한 문서를 결정할 수 있습니다.
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 다음 중에서 선택할 수 있습니다. • `FORCED_FULL_CRAWL` 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • `FULL_CRAWL` 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • `CHANGE_LOG` 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

구성	설명
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 특정 문서에 대한 액세스 권한이 있는 사용자 및 그룹의 ID/주체 정보를 동기화하는 데 사용합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

Microsoft Yammer JSON 스키마

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

MySQL 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 개체의 일부로 포함시키십시오. TemplateConfiguration 데이터 소스로 JDBC, 데이터 소스 유형으로 mysql, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. MySQL JSON 스키마를 참조하세요.

다음 표는 MySQL JSON 스키마의 매개 변수를 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DBType - 사용하는 자바 데이터베이스 유형 (,,mysql, db2 또는). postgresql oracle sqlserver

구성	설명
	• dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.

구성	설명
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. `JDBC`을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

MySQL JSON 스키마

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

Oracle Database 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration개체의 일부로 포함합니다. 데이터 소스로 JDBC, 데이터 소스 유형으로 oracle, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Oracle Database JSON 스키마를 참조하세요.

다음 표에서는 Oracle 데이터베이스 JSON 스키마의 매개변수를 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DBType - 사용하는 Java 데이터베이스 유형 (,, `mysql`, `db2`, `postgresql` 또는) `oracle` `sqlserver` • dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니

구성	설명
	다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. JDBC을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

Oracle Database JSON 스키마

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

PostgreSQL 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration개체의 일부로 포함합니다. 데이터 소스로 JDBC, 데이터 소스 유형으로 postgresql, 보안 인증 정보의 암호, 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. PostgreSQL JSON 스키마를 참조하세요.

다음 표에서는 PostgreSQL JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스 연결을 위한 필수 구성 정보. • DBType - 사용하는 자바 데이터베이스 유형 (,,, 또는). `mysql` `db2` `postgresql` `oracle` `sqlserver` • dbHost - 데이터베이스 호스트 이름. • dbPort - 데이터베이스 포트. • dbInstance - 데이터베이스 인스턴스.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다. 데이터 소스 유형과 비밀 ARN을 지정합니다.
문서	데이터베이스 콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션. 데이터베이스 데이터 소스에 특정 콘텐츠를 포함하거나 제외하는 데 사용합니다.
primaryKeys	데이터베이스 테이블의 기본 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
titleColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
bodyColumn	데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
sqlQuery	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니

구성	설명
	다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
timestampColumn	타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
timestampFormat	콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.
timezone	콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
changeDetectingColumns	콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
allowedUsersColumns	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
allowedGroupsColumn	콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
sourceURIColumn	인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
isSslEnabled	SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
type	데이터 소스의 유형. JDBC을 데이터 소스 유형으로 지정합니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정하세요. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
secretArn	데이터베이스에 연결하는 데 필요한 사용자 이름 및 암호가 포함된 Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)입니다. 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "user name": "database user name", "password": " password" }`
version	현재 지원되는 템플릿의 버전.

PostgreSQL JSON 스키마

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

Salesforce 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 TemplateConfiguration개체의 일부로 포함합니다. Salesforce 호스트 URL은 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 제공합니다. 또한 데이터 소스 유형으로 SALESFORCEV2, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Salesforce JSON 스키마를 참조하세요.

다음 표에서는 Salesforce JSON 스키마의 매개 변수에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
hostUrl	인덱싱할 Salesforce 인스턴스의 URL.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• account • contact • campaign • case • product • lead • contract • partner • profile • idea • pricebook • task • 솔루션 • attachment • 사용자 • 문서 • knowledgeArticles • 그룹 • opportunity • chatter • customEntity	Salesforce 엔티티의 속성 또는 필드 이름을 인덱스 필드 이름에 매핑하는 개체 목록입니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

구성	설명
secretARN	Salesforce에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀의 Amazon 리소스 이름 (ARN). 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "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	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.

구성	설명
• accountFilter • contactFilter • caseFilter • campaignFilter • contractFilter • groupFilter • leadFilter • productFilter • opportunityFilter • partnerFilter • pricebookFilter • ideaFilter • profileFilter • taskFilter • solutionFilter • userFilter • chatterFilter • documentFilter • knowledgeArticleFilter • customEntities	필터링할 항목을 지정하는 문자열 모음.

구성	설명
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	Salesforce 데이터 소스에서 특정 파일을 포함하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.

구성	설명
exclusionPatterns • exclusionDocumentFileTypePatterns • exclusionDocumentFileNamePatterns • exclusionAccountFileTypePatterns • exclusionCampaignFileTypePatterns • exclusionCampaignFileNamePatterns • exclusionCaseFileTypePatterns • exclusionCaseFileNamePatterns • exclusionContactFileTypePatterns • exclusionContractFileNamePatterns • exclusionLeadFileTypePatterns • exclusionLeadFileNamePatterns • exclusionOpportunityFileTypePatterns • exclusionOpportunityFileNamePatterns • exclusionSolutionFileTypePatterns • exclusionSolutionFileNamePatterns • exclusionTaskFileTypePatterns • exclusionTaskFileNamePatterns • exclusionGroupFileTypePatterns • exclusionGroupFileNamePatterns • exclusionChatterFileTypePatterns • exclusionChatterFileNamePatterns • exclusionCustomEntityFileTypePatterns • exclusionCustomEntityFileNamePatterns	Salesforce 데이터 소스에서 특정 파일을 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 제외 및 포함 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.

구성	설명
• isCrawlAccount • isCrawlContact • isCrawlCase • isCrawlCampaign • isCrawlProduct • isCrawlLead • isCrawlContract • isCrawlPartner • isCrawlProfile • isCrawlIdea • isCrawlPricebook • isCrawlDocument • crawlSharedDocument • isCrawlGroup • isCrawlOpportunity • isCrawlChatter • isCrawlUser • isCrawlSolution • isCrawlTask • isCrawlAccount첨부 파일 • isCrawlContact첨부파일 • isCrawlCase첨부파일 • isCrawlCampaign첨부파일 • isCrawlLead첨부파일 • isCrawlContract첨부파일 • isCrawlGroup첨부파일 • isCrawlOpportunity첨부파일 • isCrawlChatter첨부파일 • isCrawlSolution첨부파일	`true`Salesforce 계정에서 이러한 유형의 파일을 크롤링하려면

구성	설명

• isCrawlTask첨부파일
• isCrawlCustomEntityAttachments
• isCrawlKnowledge기사
  • isCrawlDraft
  • isCrawlPublish
  • isCrawlArchived	

type	데이터 소스의 유형. `SALESFORCEV2` 을 데이터 소스 유형으로 지정합니다.
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 사용자 및 그룹의 ID/주체 정보를 특정 문서에 대한 액세스 권한과 동기화합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

구성	설명
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
version	현재 지원되는 이 템플릿의 버전.

Salesforce JSON 스키마

```
{
  "$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 템플릿 스키마

데이터 소스 스키마를 TemplateConfiguration개체의 일부로 포함하는 JSON을 포함합니다. 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 ServiceNow 호스트 URL, 인증 유형 및 인스턴스 버전을 제공합니다. 또한 데이터 소스 유형으로 SERVICENOWV2, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type 호출할 때 TEMPLATE 로 지정합니다 CreateDataSource.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. ServiceNow JSON 스키마를 참조하세요.

다음 표에서는 ServiceNow JSON 스키마의 파라미터에 대해 설명합니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
hostUrl	ServiceNow 호스트 URL. 예: *your-domain.service-now.com* .
authType	사용하는 인증 유형(basicAuth 또는 OAuth2).
servicenowInstanceVersion	사용하는 ServiceNow 버전. Tokyo,, SandiegoRome, 중에서 선택할 수 있습니다Others.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
• knowledgeArticle • attachment • serviceCatalog • incident	ServiceNow참조 문서, 첨부 파일, 서비스 카탈로그 및 인시던트의 속성 또는 필드 이름을 Amazon Kendra 색인 필드 이름에 매핑하는 개체 목록입니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요. ServiceNow 데이터 원본 필드 이름은 ServiceNow 사용자 지정 메타데이터에 있어야 합니다.
additional properties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.

구성	설명
maxFileSizeInMegaBytes	Amazon Kendra가 크롤링할 파일 크기 제한을 MB 단위로 지정합니다. Amazon Kendra는 사용자가 정의한 크기 제한 내에 있는 파일만 크롤링합니다. 기본 파일 크기는 50MB입니다. 최대 파일 크기는 0MB보다 크고 50MB보다 작거나 같아야 합니다.
• knowledgeArticleFilter • incidentQueryFilter • serviceCatalogQuery필터 • knowledgeArticleTitleRegExp • serviceCatalogTitleRegExp • incidentTitleReg경험치 • inclusionFileType패턴 • exclusionFileType패턴 • inclusionFileName패턴 • exclusionFileName패턴 • incidentStateType	ServiceNow 데이터 원본에 특정 파일을 포함 및/또는 제외하기 위한 정규 표현식 패턴 목록입니다. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.

구성	설명
• isCrawlKnowledge기사 • isCrawlKnowledgeArticleAttachment • includePublicArticles전용 • isCrawlService카탈로그 • isCrawlServiceCatalogAttachment • isCrawlActiveServiceCatalog • isCrawlInactiveServiceCatalog • isCrawlIncident • isCrawlIncident첨부파일 • isCrawlActive인시던트 • isCrawlInactive인시던트 • ACL 적용하기 ForKnowledgeArticle • ACL 적용하기 ForServiceCatalog • ACL 적용하기 ForIncident	`true` ServiceNow 지식 문서, 서비스 카탈로그, 인시던트 및 첨부 파일을 크롤링합니다.
type	데이터 소스의 유형. `SERVICENOWV2` 을 데이터 소스 유형으로 지정합니다.
enableIdentityCrawler	`true` Amazon Kendra의 ID 크롤러를 사용하여 사용자 및 그룹의 ID/주체 정보를 특정 문서에 대한 액세스 권한과 동기화합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMapping API를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

구성	설명
syncMode	데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정하십시오. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

secretARN	연결에 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀의 Amazon 리소스 이름 (ARN). ServiceNow 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다.
	`{ "username": " user name", "password": " password" }`
	OAuth2 인증을 사용하는 경우 암호에는 다음 키가 있는 JSON 구조가 포함되어야 합니다.
	`{ "username": " user name", "password": " password", "clientId": " client id", "clientSecret": " client secret" }`

version	현재 지원되는 템플릿의 버전.

ServiceNow JSON 스키마

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "hostUrl": {
              "type": "string",
              "pattern": "^(?!(^(https?|ftp|file):\/\/))[a-z0-9-]+(.service-
now.com|.servicenowservices.com)$",
              "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"
  ]
}
```

슬랙 템플릿 스키마

데이터 소스 스키마를 객체의 TemplateConfiguration일부로 포함하는 JSON을 포함합니다. 호스트 URL은 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 제공합니다. 또한 데이터 소스 유형으로 SLACK, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. 슬랙 JSON 스키마를 참조하세요.

다음 표에는 Slack JSON 스키마의 파라미터가 설명되어 있습니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
팀 ID	슬랙 메인 페이지 URL에서 복사한 슬랙 팀 ID.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.
모두	Slack콘텐츠의 속성 또는 필드 이름을 Amazon Kendra 인덱스 필드 이름에 매핑하는 객체 목록입니다.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
inclusionPatterns	Slack데이터 원본에 특정 콘텐츠를 포함하기 위한 정규 표현식 패턴 목록입니다. 패턴과 일치하는 콘텐츠는 인덱스에 포함됩니다. 패턴과 일치하지 않는 콘텐츠는 인덱스에서 제외됩니다. 포

구성	설명
	함 패턴과 제외 패턴 모두에 일치하는 콘텐츠가 있는 경우 제외 패턴이 우선하며 콘텐츠는 색인에 포함되지 않습니다.
exclusionPatterns	데이터 원본에서 특정 콘텐츠를 제외하기 위한 정규 표현식 패턴 목록입니다. Slack 패턴과 일치하는 콘텐츠는 인덱스에서 제외됩니다. 패턴과 일치하지 않는 콘텐츠는 색인에 포함됩니다. 포함 패턴과 제외 패턴 모두에 일치하는 콘텐츠가 있는 경우 제외 패턴이 우선하며 해당 콘텐츠는 색인에 포함되지 않습니다.
crawlBotMessages	`true`봇 메시지를 크롤링하기 위해서입니다.
제외/보관됨	`true`보관된 메시지의 크롤링을 제외하려면
대화 유형	인덱싱하려는 대화 유형 (`PUBLIC_CHANNEL` , 및 `PRIVATE_CHANNEL` 여부). `GROUP_MESSAGE` `DIRECT_MESSAGE`
채널 필터	색인을 생성할지 여부에 관계없이 `private_channel` 인덱싱하려는 채널 유형입니다. `public_channel`
sinceDate	Slack커넥터가 특정 내용을 기반으로 콘텐츠를 크롤링하도록 `sinceDate` 매개 변수를 구성하도록 선택할 수 있습니다. `sinceDate`
룩백	커넥터가 마지막 Slack 커넥터 동기화 전까지 지정된 시간까지 업데이트 또는 삭제된 콘텐츠를 크롤링하도록 `lookBack` 매개변수를 구성할 수 있습니다.

구성	설명
syncMode	데이터 원본 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정합니다. 다음 중에서 선택할 수 있습니다. • FORCED_FULL_CRAWL 모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다. • FULL_CRAWL 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하려는 경우 Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다. • CHANGE_LOG 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
type	데이터 소스의 유형. SLACK을 데이터 소스 유형으로 지정합니다.
enableIdentityCrawler	true Amazon Kendra의 ID 크롤러를 사용하여 특정 문서에 대한 액세스 권한이 있는 사용자 및 그룹의 ID/주체 정보를 동기화하는 데 사용합니다. ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

구성	설명
secretArn	연결에 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀의 Amazon 리소스 이름 (ARN). Slack 비밀에는 다음 키가 있는 JSON 구조가 포함되어야 합니다. `{ "slackToken": " token" }`
version	현재 지원되는 이 템플릿의 버전입니다.

슬랙 JSON 스키마

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

Zendesk 템플릿 스키마

데이터 소스 스키마를 포함하는 JSON을 객체의 일부로 포함합니다. TemplateConfiguration 호스트 URL은 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 제공합니다. 또한 데이터 소스 유형으로 ZENDESK, 보안 인증 정보의 암호 및 기타 필수 구성을 지정합니다. 그런 다음 Type CreateDataSource호출할 때 TEMPLATE 로 지정합니다.

이 개발자 안내서에 제공된 템플릿을 사용할 수 있습니다. Zendesk JSON 스키마를 참조하세요.

다음 표에는 Zendesk JSON 스키마의 파라미터가 설명되어 있습니다.

구성	설명
connectionConfiguration	데이터 소스의 엔드포인트의 구성 정보.
repositoryEndpointMetadata	데이터 소스의 엔드포인트 정보.
hostURL	Zendesk 호스트 URL. 예: https://yoursubdomain.zendesk.com.
repositoryConfigurations	데이터 소스 콘텐츠의 구성 정보. 특정 유형의 콘텐츠 및 필드 매핑을 구성하는 경우를 예로 들 수 있습니다.

구성	설명
• ticket • ticketComment • ticketCommentAttachment • article • articleComment • articleAttachment • communityTopic • communityPostComment	Amazon Kendra 인덱스 필드 이름에 Zendesk 티켓의 속성 또는 필드 이름을 매핑하는 객체의 목록. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.
secretARN	Zendesk에 연결하는 데 필요한 키-값 쌍이 포함된 AWS Secrets Manager 비밀번호의 Amazon 리소스 이름 (ARN). 암호에는 호스트 URL, 클라이언트 ID, 클라이언트 암호, 사용자 이름 및 암호 키로 구성된 JSON 구조가 포함되어야 합니다.
additionalProperties	데이터 소스의 콘텐츠에 대한 추가 구성 옵션.
organizationNameFilter	특정 조직 내에 있는 티켓을 인덱싱하도록 선택할 수 있습니다.
sinceDate	Zendesk 커넥터가 특정 `sinceDate` 기반으로 콘텐츠를 크롤링하도록 `sinceDate` 파라미터를 구성하도록 선택할 수 있습니다.
inclusionPatterns	Zendesk 데이터 소스에서 특정 파일을 포함하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에 포함됩니다. 패턴과 일치하지 않는 파일은 인덱스에서 제외됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.

구성	설명
exclusionPatterns	Zendesk 데이터 소스에서 특정 파일을 제외하는 정규식 패턴 목록. 패턴과 일치하는 파일은 인덱스에서 제외됩니다. 패턴과 일치하지 않는 파일은 인덱스에 포함됩니다. 파일이 포함 및 제외 패턴과 모두 일치하는 경우 제외 패턴이 우선 적용되고 파일은 인덱스에 포함되지 않습니다.
• isCrawlTicket • isCrawlTicket댓글 달기 • isCrawlTicketCommentAttachment • isCrawlArticle • isCrawlArticle코멘트 • isCrawlArticle첨부파일 • isCrawlCommunity주제 • isCrawlCommunity포스트 • isCrawlCommunityPostComment	이러한 유형의 콘텐츠를 크롤링하려면 `true` "“를 입력하세요.
type	`ZENDESK`을 데이터 소스 유형으로 지정합니다.
useChangeLog	Zendesk 변경 로그를 사용하여 색인에서 업데이트가 필요한 문서를 결정하려면 `true` "“를 입력하세요. 변경 로그의 크기에 따라 Zendesk에서 문서를 스캔하는 것이 더 빠를 수 있습니다. Zendesk 데이터 소스를 인덱스와 처음으로 동기화하는 경우 모든 문서가 스캔됩니다.

Zendesk JSON 스키마

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
```

```
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "hostUrl": {
              "type": "string",
              "pattern": "https:.*"
            }
          },
          "required": [
            "hostUrl"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "ticket": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
```

```
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"

                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "ticketComment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
```

```
                    "indexFieldType",
                    "dataSourceFieldName"

                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "ticketCommentAttachment": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
```

```
            ]
          }
        }
      },
      "required": [
        "fieldMappings"
      ]
    },
    "article": {
      "type": "object",
      "properties": {
        "fieldMappings": {
          "type": "array",
          "items": {
            "anyOf": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "dd-MM-yyyy HH:mm:ss"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        }
      },
      "required": [
```

```
            "fieldMappings"
          ]
        },
        "communityPostComment": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "articleComment": {
          "type": "object",
```

```
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
                  {
                    "type": "object",
                    "properties": {
                      "indexFieldName": {
                        "type": "string"
                      },
                      "indexFieldType": {
                        "type": "string",
                        "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                      },
                      "dataSourceFieldName": {
                        "type": "string"
                      },
                      "dateFieldFormat": {
                        "type": "string",
                        "pattern": "dd-MM-yyyy HH:mm:ss"
                      }
                    },
                    "required": [
                      "indexFieldName",
                      "indexFieldType",
                      "dataSourceFieldName"
                    ]
                  }
                ]
              }
            }
          },
          "required": [
            "fieldMappings"
          ]
        },
        "articleAttachment": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": {
                "anyOf": [
```

```
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      },
      "communityTopic": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": {
              "anyOf": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
```

```
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "dd-MM-yyyy HH:mm:ss"
                    }
                  },
                  "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                  ]
                }
              ]
            }
          }
        },
        "required": [
          "fieldMappings"
        ]
      }
    }
  },
  "secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "organizationNameFilter": {
        "type": "array"
      },
      "sinceDate": {
        "type": "string",
        "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2} [0-9]{2}:[0-9]{2}:[0-9]{2}$"
      },
```

```
        "inclusionPatterns": {
          "type": "array"
        },
        "exclusionPatterns": {
          "type": "array"
        },
        "isCrawTicket": {
          "type": "string"
        },
        "isCrawTicketComment": {
          "type": "string"
        },
        "isCrawTicketCommentAttachment": {
          "type": "string"
        },
        "isCrawlArticle": {
          "type": "string"
        },
        "isCrawlArticleAttachment": {
          "type": "string"
        },
        "isCrawlArticleComment": {
          "type": "string"
        },
        "isCrawlCommunityTopic": {
          "type": "string"
        },
        "isCrawlCommunityPost": {
          "type": "string"
        },
        "isCrawlCommunityPostComment": {
          "type": "string"
        }
      }
    },
    "type": {
      "type": "string",
      "pattern": "ZENDESK"
    },
    "useChangeLog": {
      "type": "string",
      "enum": ["true", "false"]
    }
  },
```

```
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "additionalProperties": false,
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "useChangeLog",
    "secretArn",
    "type"
  ]
}
```

Adobe Experience Manager

Adobe Experience Manager는 웹 사이트 또는 모바일 앱 콘텐츠를 만드는 데 사용되는 콘텐츠 관리 시스템입니다. 를 Amazon Kendra 사용하여 페이지 Adobe Experience Manager 및 콘텐츠 자산에 연결하고 이를 인덱싱할 수 있습니다.

Amazon Kendra 클라우드 서비스 작성자 인스턴스 및 Adobe Experience Manager 온프레미스 작성자 및 게시 인스턴스로 Adobe Experience Manager (AEM) 을 지원합니다.

Amazon Kendra 콘솔 또는 API를 사용하여 Adobe Experience Manager 데이터 원본에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Adobe Experience Manager 데이터 소스 커넥터 문제 해결에 대한 자세한 내용은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침

지원되는 기능

Adobe Experience Manager 데이터 소스 커넥터에서 지원하는 기능은 다음과 같습니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- OAuth 2.0 및 기본 인증
- Virtual Private Cloud(VPC)

필수 조건

Adobe Experience Manager데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Adobe Experience Manager 및 AWS 계정에서 다음과 같이 변경하십시오.

Adobe Experience Manager에서 다음 사항을 갖추었는지 확인하세요.

- 관리자 권한이 있는 계정 또는 관리자 사용자에 액세스합니다.
- Adobe Experience Manager 호스트 URL을 복사했습니다.

> **Note**
>
> (온프레미스/서버) AWS Secrets Manager 는 에 포함된 엔드포인트 정보가 데이터 원본 구성 세부 정보에 지정된 엔드포인트 정보와 동일한지 Amazon Kendra 확인합니다. 이렇게 하면 사용자가 작업을 수행할 권한이 없지만 구성된 보안 암호에 액세스하여 작업을 수행하는 데 Amazon Kendra 를 프록시로 사용하는 보안 문제인 혼동된 대리자 문제를 방지하는 데 도움이 됩니다. 나중에 엔드포인트 정보를 변경하는 경우 새 보안 암호를 생성하여 이 정보를 동기화해야 합니다.

- 관리자 사용자 이름과 암호의 기본 인증 보안 인증 정보를 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 선택 사항: Adobe Experience Manager (AEM) 에서 OAuth 2.0 자격 증명을 클라우드 서비스 또는 AEM 온프레미스로 구성했습니다. AEM온프레미스를 사용하는 경우 보안 인증 정보에는 클라이언트 ID, 클라이언트 암호 및 프라이빗 키가 포함됩니다. 클라우드 서비스형 AEM을 사용하는 경우 보안 인증 정보에는 클라이언트 ID, 클라이언트 보안 암호, 프라이빗 키, 조직 ID, 기술 계정 ID 및 Adobe Identity Management System(IMS) 호스트가 포함됩니다. 클라우드 서비스형 AEM의 보안 인증 정보를 생성하는 방법에 대한 자세한 내용은 Adobe Experience Manager 설명서를 참조하세요. AEM 온프레미스의 경우 Adobe Granite OAuth 2.0 서버 구현(com.adobe.granite.oauth.server)은 AEM의 OAuth 2.0 서버 기능에 대한 지원을 제공합니다.
- Adobe Experience Manager 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Adobe Experience Manager 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Adobe Experience Manager 데이터 소스를 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Adobe Experience Manager데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Adobe Experience Manager 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 Adobe Experience Manager 을 Amazon Kendra참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Adobe Experience Manager

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 소스 추가 페이지에서 Adobe Experience Manager 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Adobe Experience Manager 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 색인을 기준으로 문서를 필터링할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. 소스 - AEM 온프레미스 또는 클라우드 서비스형 AEM을 선택합니다.

      Adobe Experience Manager 호스트 URL을 입력합니다. 예를 들어 AEM 온프레미스를 사용하는 경우 호스트 이름과 포트를 포함해야 합니다. https://hostname:port 클라우드 서

비스형 AEM을 사용하는 경우 작성자 URL을 사용할 수 있습니다. https://author-xxxxxx-xxxxxxx.adobeaemcloud.com

b. SSL 인증서 위치 - Amazon S3 버킷에 저장된 SSL 인증서의 경로를 입력합니다. 이를 사용하여 보안 SSL 연결로 AEM 온프레미스에 연결합니다.

c. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

d. 인증 - 기본 인증 또는 OAuth 2.0 인증을 선택합니다. 그런 다음 기존 AWS Secrets Manager 암호를 선택하거나 Adobe Experience Manager 자격 증명을 저장할 새 암호를 생성합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 비밀 창이 열립니다.

   기본 인증을 선택한 경우 암호 이름, Adobe Experience Manager 사이트 사용자 이름 및 암호를 입력합니다. 사용자는 관리자 권한이 있거나 관리자 사용자여야 합니다.

   OAuth 2.0 인증을 선택하고 AEM 온프레미스를 사용하는 경우 암호, 클라이언트 ID, 클라이언트 보안 암호 및 프라이빗 키의 이름을 입력합니다. 클라우드 서비스형 AEM을 사용하는 경우 보안 암호 이름, 클라이언트 ID, 클라이언트 보안 암호, 프라이빗 키, 조직 ID, 기술 계정 ID 및 Adobe Identity Management System(IMS) 호스트를 입력합니다.

   비밀번호를 저장하고 추가하세요.

e. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

f. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

g. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

h. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위 - 특정 콘텐츠 유형, 페이지 구성 요소 및 루트 경로를 크롤링하기 위한 제한을 설정하고 정규식 패턴을 사용하여 콘텐츠를 필터링합니다.

i. 콘텐츠 유형 - 페이지나 자산만 크롤링할지, 아니면 둘 다 크롤링할지 선택합니다.

ii. (선택 사항) 추가 구성 - 다음 설정을 구성합니다.

- 페이지 구성 요소 - 페이지 구성 요소의 특정 이름입니다. 페이지 구성 요소는 Adobe Experience Manager 템플릿 편집기와 함께 작동하도록 설계된 확장 가능한 페이지 구성 요소로, 템플릿 편집기로 페이지 머리글/바닥글 및 구조 구성 요소를 조합할 수 있습니다.
- 컨텐츠 조각 변형 - 컨텐츠 조각 변형의 특정 이름. 컨텐츠 조각을 사용하면 Adobe Experience Manager에서 페이지에 구애받지 않는 컨텐츠를 설계, 생성, 큐레이션 및 게시할 수 있습니다. 이를 통해 여러 위치/여러 채널에서 바로 사용할 수 있도록 콘텐츠를 준비할 수 있습니다.
- 루트 경로 - 특정 콘텐츠의 루트 경로입니다.
- Regex 패턴 - 특정 페이지 및 자산을 포함하거나 제외하기 위한 정규 표현식 패턴입니다.

b. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

c. 시간대 ID - AEM 온프레미스를 사용하고 서버의 시간대가 Amazon Kendra AEM 커넥터 또는 인덱스의 시간대와 다른 경우 AEM 커넥터 또는 인덱스에 맞춰 서버 시간대를 지정할 수 있습니다. AEM 온프레미스의 기본 시간대는 Amazon Kendra AEM 커넥터 또는 인덱스의 시간대입니다. 클라우드 서비스형 AEM의 기본 시간대는 그리니치 표준시입니다.

d. 빈도용 동기화 실행 일정 —데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑할 필드를 선택합니다. 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

b. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra 연결하려면 Adobe Experience Manager

TemplateConfiguration API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 AEM 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- AEM 호스트 URL - Adobe Experience Manager 호스트 URL을 지정합니다. 예를 들어 AEM 온프레미스를 사용하는 경우 호스트 이름과 포트를 포함해야 합니다. https://hostname:port 클라우드 서비스형 AEM을 사용하는 경우 작성자 URL을 사용할 수 있습니다. https://author-xxxxxx-xxxxxxx.adobeaemcloud.com

- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 인증 유형 - 사용하려는 인증 유형(Basic 또는 OAuth2)을 지정합니다.
- AEM 유형 - 사용할 Adobe Experience Manager 유형(CLOUD 또는 ON_PREMISE)을 지정합니다.
- 보안 암호 Amazon 리소스 이름(ARN) - AEM 온프레미스 또는 클라우드에 기본 인증을 사용하려는 경우 사용자 이름 및 암호의 인증 보안 인증 정보를 저장하는 보안 암호를 제공합니다. 비밀의 Amazon 리소스 이름 (ARN) 을 AWS Secrets Manager 제공합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "aemUrl": "Adobe Experience Manager On-Premise host URL",
    "username": "user name with admin permissions",
    "password": "password with admin permissions"
}
```

AEM 온프레미스에 OAuth 2.0 인증을 사용하려는 경우 보안 암호는 다음 키를 사용하여 JSON 구조에 저장됩니다.

```
{
    "aemUrl": "Adobe Experience Manager host URL",
    "clientId": "client ID",
    "clientSecret": "client secret",
    "privateKey": "private key"
}
```

클라우드 서비스형 AEM에 OAuth 2.0 인증을 사용하려는 경우 보안 암호는 다음 키를 사용하여 JSON 구조에 저장됩니다.

```
{
    "clientId": "client ID",
    "clientSecret": "client secret",
    "privateKey": "private key",
    "orgId": "organization ID",
    "technicalAccountId": "technical account ID",
    "imsHost": "Adobe Identity Management System (IMS) host"
}
```

- IAM 역할 - 역할을 `CreateDataSource` 호출하여 Secrets Manager 비밀에 액세스할 수 있는 권한을 제공하고 Adobe Experience Manager 커넥터 및 에 필요한 공개 API를 호출할 `RoleArn` 시기를 지정합니다. IAM Amazon Kendra자세한 내용은 Adobe Experience Manager 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 시간대 ID - AEM On-Premise를 사용하고 서버의 시간대가 Amazon Kendra AEM 커넥터 또는 인덱스의 시간대와 다른 경우 AEM 커넥터 또는 색인에 맞게 서버 시간대를 지정할 수 있습니다.

  AEM 온프레미스의 기본 시간대는 AEM 커넥터 또는 인덱스의 시간대입니다. Amazon Kendra 클라우드 서비스형 AEM의 기본 시간대는 그리니치 표준시입니다.

  지원되는 시간대 ID에 대한 자세한 내용은 Adobe Experience Manager JSON 스키마를 참조하세요.
- 포함 및 제외 필터 - 특정 페이지 및 자산을 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용

하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

- 필드 매핑 - Adobe Experience Manager 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 해당 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Adobe Experience Manager 템플릿 스키마를 참조하세요.

Alfresco

Alfresco는 고객이 콘텐츠를 저장하고 관리할 수 있도록 지원하는 콘텐츠 관리 서비스입니다. 를 Amazon Kendra 사용하여 Alfresco 문서 라이브러리, Wiki 및 블로그를 인덱싱할 수 있습니다.

Amazon Kendra Alfresco온-프레미스 및 Alfresco 클라우드 (서비스형 플랫폼) 를 지원합니다.

Amazon Kendra 콘솔 또는 TemplateConfigurationAPI를 Amazon Kendra 사용하여 Alfresco 데이터 소스에 연결할 수 있습니다.

Amazon Kendra Alfresco 데이터 소스 커넥터의 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Alfresco 데이터 소스 커넥터에서 지원하는 기능은 다음과 같습니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- OAuth 2.0 및 기본 인증
- Virtual Private Cloud(VPC)

필수 조건

Alfresco 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 and에서 다음과 같이 변경하십시오. Alfresco AWS 계정

Alfresco에서 다음 사항을 갖추었는지 확인하세요.

- Alfresco 리포지토리 URL과 웹 애플리케이션 URL을 복사했습니다. 특정 Alfresco 사이트만 인덱싱하려는 경우에는 사이트 ID도 복사하세요.
- 최소한 읽기 권한이 있는 사용자 이름과 암호가 포함된 Alfresco 보안 인증 정보를 기록해 두었습니다. OAuth 2.0 인증을 사용하려면 Alfresco 관리자 그룹에 사용자를 추가해야 합니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 선택 사항: OAuth 2.0 자격 증명을 구성했습니다. Alfresco 보안 인증 정보에는 클라이언트 ID, 클라이언트 보안 암호 및 토큰 URL이 포함됩니다. Alfresco 온프레미스용 클라이언트를 구성하는 방법에 대한 자세한 내용은 Alfresco 설명서를 참조하세요. Alfresco 클라우드(PaaS)를 사용하는 경우 Hyland 지원 팀에 문의하여 Alfresco OAuth 2.0 인증을 요청해야 합니다.
- Alfresco 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 AWS 계정있는지 확인하십시오.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형 및 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Alfresco 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Alfresco 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. Amazon Kendra API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Alfresco 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 수 있도록 Amazon Kendra Alfresco 데이터 원본의 필수 세부 정보를 제공해야 합니다. Alfresco를 아직 구성하지 않은 경우 을 참조하십시오. Amazon Kendra필수 조건

Console

연결하려면 Amazon Kendra Alfresco

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Alfresco 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Alfresco 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. Alfresco유형 Alfresco —온프레미스/서버 또는 Alfresco 클라우드 (Platform as a Service) 중 무엇을 사용할지 선택합니다.

   b. Alfresco 리포지토리 URL - 알프레스코 리포지토리 URL을 입력합니다. 예를 들어 Alfresco Cloud(PaaS)를 사용하는 경우 리포지토리 URL은 https://company.alfrescocloud.com과 같을 수 있습니다. 또는 Alfresco 온프레미스를 사용하는 경우 리포지토리 URL은 https://company-alfresco-instance.company-domain.suffix:port와 같을 수 있습니다.

   c. Alfresco 사용자 애플리케이션. URL - Alfresco 사용자 인터페이스 URL을 입력합니다. 리포지토리 URL은 Alfresco 관리자로부터 받을 수 있습니다. 예를 들어 사용자 인터페이스 URL은 https://example.com과 같을 수 있습니다.

   d. SSL 인증서 위치 - 버킷에 저장된 SSL 인증서의 경로를 입력합니다. Amazon S3 이를 사용하여 보안 SSL 연결로 Alfresco 온프레미스에 연결합니다.

   e. 권한 부여 - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니

다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

f. 인증 - 기본 인증 또는 OAuth 2.0 인증을 선택합니다. 그런 다음 기존 Secrets Manager 보안 암호를 선택하거나 Alfresco 보안 인증 정보를 저장할 새 보안 암호를 생성합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 암호 창이 열립니다.

기본 인증을 선택한 경우 암호 이름, Alfresco 사용자 이름 및 암호를 입력합니다.

OAuth 2.0 인증을 선택하고 보안 암호의 이름, 클라이언트 ID, 클라이언트 보안 암호 및 토큰 URL을 입력합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

i. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

j. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위 - 특정 콘텐츠를 크롤링하기 위한 제한을 설정하고 정규식 패턴을 사용하여 콘텐츠를 필터링합니다.

b. i. 콘텐츠—Alfresco에서 '속성'으로 표시된 콘텐츠, 특정 Alfresco 사이트 내의 콘텐츠 또는 모든 Alfresco 사이트의 콘텐츠를 크롤링할지 여부를 선택합니다.

ii. (선택 사항) 추가 구성 - 다음 설정을 구성합니다.

- 설명 포함 - Alfresco 문서 라이브러리 및 블로그에 설명을 포함하도록 선택합니다.
- 정규식 패턴 - 특정 페이지 및 자산을 포함하거나 제외하기 위한 정규 표현식 패턴입니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra 연결하려면 Alfresco

TemplateConfiguration API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 ALFRESCO 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- Alfresco 사이트 ID - Alfresco 사이트 ID를 지정합니다.
- Alfresco 리포지토리 URL - Alfresco 리포지토리 URL을 지정합니다. 리포지토리 URL은 Alfresco 관리자로부터 받을 수 있습니다. 예를 들어 Alfresco Cloud(PaaS)를 사용하는 경우 리포지토리 URL은 https://company.alfrescocloud.com과 같을 수 있습니다. 또는 Alfresco 온프레미스를 사용하는 경우 리포지토리 URL은 https://company-alfresco-instance.company-domain.suffix:port와 같을 수 있습니다.
- Alfresco 웹 애플리케이션 URL - Alfresco 사용자 인터페이스 URL을 지정합니다. 리포지토리 URL은 Alfresco 관리자로부터 받을 수 있습니다. 예를 들어 사용자 인터페이스 URL은 https://example.com과 같을 수 있습니다.
- 인증 유형 - 사용하려는 인증 유형(OAuth2 또는 Basic)을 지정합니다.
- Alfresco 유형 - PAAS(클라우드/서비스형 플랫폼) 또는 ON_PREM(온프레미스) 중에서 사용할 Alfresco 유형을 지정합니다.
- 보안 암호 Amazon 리소스 이름(ARN) - 기본 인증을 사용하려는 경우 사용자 이름 및 암호의 보안 인증 정보를 저장하는 보안 암호를 제공합니다. 비밀의 Amazon 리소스 이름 (ARN) 을 AWS Secrets Manager 제공합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "username": "user name",
    "password": "password"
}
```

OAuth 2.0 인증을 사용하려는 경우 보안 암호는 다음 키를 사용하여 JSON 구조에 저장됩니다.

```
{
    "clientId": "client ID",
    "clientSecret": "client secret",
    "tokenUrl": "token URL"
}
```

- IAM 역할 - CreateDataSource Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할에 제공하고 Alfresco 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 RoleArn 시기를 지정하십시오. Amazon Kendra자세한 내용은 Alfresco 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 콘텐츠 유형 - Alfresco에서 '속성'으로 표시된 콘텐츠, 특정 Alfresco 사이트 내의 콘텐츠 또는 모든 Alfresco 사이트의 콘텐츠 등, 크롤링하려는 콘텐츠 유형입니다. 특정 '속성' 콘텐츠를 나열할 수도 있습니다.
- 포함 및 제외 필터 - 특정 파일을 포함할지 또는 제외할지 여부를 지정합니다.

  > **Note**
  >
  > 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 Amazon Kendra 때 색인을 업데이트하는 방법을 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - `FORCED_FULL_CRAWL`모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - `FULL_CRAWL`데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 필드 매핑 - Alfresco 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 해당 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Alfresco 템플릿 스키마를 참조하세요.

자세히 알아보기

Alfresco 데이터 Amazon Kendra 원본과의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- 를 사용하여 콘텐츠를 지능적으로 검색합니다. AlfrescoAmazon Kendra

Aurora (MySQL)

Aurora 클라우드용으로 구축된 관계형 데이터베이스 관리 시스템 (RDBMS) 입니다. Aurora 사용자인 경우 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다Aurora (MySQL). Amazon Kendra Aurora (MySQL)데이터 소스 커넥터는 Aurora MySQL 3 및 Aurora 서버리스 MySQL 8.0을 지원합니다.

Amazon Kendra 콘솔과 API를 사용하여 Aurora (MySQL) 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Aurora (MySQL)데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Aurora (MySQL)데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Aurora (MySQL) 및 AWS 계정에서 다음과 같이 변경하십시오.

Aurora (MySQL)에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
>
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다. Amazon RDS 콘솔에서 이 정보를 찾을 수 있습니다.
- Aurora (MySQL) 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Aurora (MySQL) 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Aurora (MySQL) 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Aurora (MySQL)데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Aurora (MySQL) 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 Aurora (MySQL) 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Aurora (MySQL)

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 Aurora (MySQL)커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Aurora (MySQL)커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. 소스에서 다음 정보를 입력합니다.

b. 호스트 - 데이터베이스 호스트 URL(예: `http://instance URL.region.rds.amazonaws.com`)을 입력합니다.

c. 포트 - 데이터베이스 포트(예: 5432)를 입력합니다.

d. 인스턴스 - 데이터베이스 인스턴스를 입력합니다.

e. 인증에서 - 다음 정보를 입력합니다.

- AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 만들어 Aurora (MySQL) 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

  A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

     I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Aurora (MySQL) -'가 자동으로 추가됩니다.

     II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

  B. 저장을 선택합니다.

f. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

g. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

h. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. SQL 쿼리는 32KB 미만이어야 하며 세미콜론 (;) 을 포함하지 않아야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra Aurora (MySQL)

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 mySql로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. Aurora (MySQL) 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 CreateDataSource Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 RoleArn 시기를 지정합니

다. Aurora (MySQL) Amazon Kendra자세한 내용은 Aurora (MySQL) 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Aurora (MySQL) 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Aurora (MySQL) 템플릿 스키마를 참조하세요.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Aurora (PostgreSQL)

Aurora 클라우드용으로 구축된 관계형 데이터베이스 관리 시스템 (RDBMS) 입니다. Aurora 사용자인 경우 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다Aurora (PostgreSQL). Amazon Kendra Aurora (PostgreSQL)데이터 소스 커넥터는 Aurora PostgreSQL 1을 지원합니다.

Amazon Kendra 콘솔과 API를 사용하여 Aurora (PostgreSQL) 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Aurora (PostgreSQL)데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Aurora (PostgreSQL)데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Aurora (PostgreSQL) 및 AWS 계정에서 다음과 같이 변경하십시오.

Aurora (PostgreSQL)에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
>
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다.
- Aurora (PostgreSQL) 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Aurora (PostgreSQL) 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Aurora (PostgreSQL) 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Aurora (PostgreSQL)데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Aurora (PostgreSQL) 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 Aurora (PostgreSQL) 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Aurora (PostgreSQL)

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 Aurora (PostgreSQL)커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Aurora (PostgreSQL)커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. 소스에서 다음 정보를 입력합니다.
   b. 호스트 - 데이터베이스 호스트 URL(예: `http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`)을 입력합니다.

c. 포트 - 데이터베이스 포트(예: 5432)를 입력합니다.

d. 인스턴스 - 데이터베이스 인스턴스(예: postgres)를 입력합니다.

e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.

f. 인증에서 - 다음 정보를 입력합니다.

- AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Aurora (PostgreSQL) 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

  A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

     I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Aurora (PostgreSQL) -'가 자동으로 추가됩니다.

     II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

  B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. SQL 쿼리는 32KB 미만이어야 하며 세미콜론 (;) 을 포함하지 않아야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.

- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra Aurora (PostgreSQL)

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 postgresql로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. Aurora (PostgreSQL) 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 CreateDataSource Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 RoleArn 시기를 지정합니다. Aurora (PostgreSQL) Amazon Kendra자세한 내용은 Aurora (PostgreSQL) 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - CreateDataSource를 호출할 때 VpcConfiguration을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Aurora (PostgreSQL) 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Aurora (PostgreSQL) 템플릿 스키마를 참조하세요.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Amazon FSx (윈도우)

Amazon FSx (Windows) 는 공유 스토리지 기능을 제공하는 완전 관리형 클라우드 기반 파일 서버 시스템입니다. Amazon FSx (Windows) 사용자인 경우 Amazon FSx (Windows) 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

> **Note**
>
> Amazon Kendra 이제 업그레이드된 Amazon FSx (Windows) 커넥터를 지원합니다.
> 콘솔이 자동으로 업그레이드되었습니다. 콘솔에서 새로 만드는 모든 커넥터는 업그레이드된 아키텍처를 사용합니다. API를 사용하는 경우 이제 TemplateConfiguration개체 대신 개체를 사용하여 커넥터를 구성해야 합니다. `FSxConfiguration`
> 이전 콘솔과 API 아키텍처를 사용하여 구성된 커넥터는 구성된 대로 계속 작동합니다. 하지만 편집하거나 업데이트할 수는 없습니다. 커넥터 구성을 편집하거나 업데이트하려면 새 커넥터를 만들어야 합니다.

커넥터 워크플로를 업그레이드된 버전으로 마이그레이션하는 것이 좋습니다. 이전 아키텍처를 사용하여 구성된 커넥터에 대한 지원은 2024년 6월에 종료될 예정입니다.

Amazon Kendra 콘솔 또는 TemplateConfigurationAPI를 Amazon Kendra 사용하여 Amazon FSx (Windows) 데이터 원본에 연결할 수 있습니다.

Amazon Kendra Amazon FSx (Windows) 데이터 소스 커넥터의 문제를 해결하려면 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Amazon FSx (Windows) 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 사용자 ID 크롤링
- 포함 및 제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Amazon FSx (Windows) 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 (Windows) 및 의 세부 정보를 확인하세요. Amazon FSx AWS 계정

Amazon FSx (Windows) 의 경우 다음이 있는지 확인하십시오.

- 읽기 및 탑재 권한으로 Amazon FSx (Windows) 를 설정합니다.

- 파일 시스템 ID를 기록해 두었습니다. 파일 시스템 ID는 Amazon FSx (Windows) 콘솔의 파일 시스템 대시보드에서 찾을 수 있습니다.
- Amazon FSx (Windows) 파일 시스템이 Amazon VPC 있는 위치를 사용하여 가상 사설 클라우드를 구성했습니다.
- Active Directory사용자 계정의 Amazon FSx (Windows) 인증 자격 증명을 기록해 두었습니다. 여기에는 Active Directory 사용자 이름과 DNS 도메인 이름 (예: user@corp.example.com) 및 암호가 포함됩니다.

> **Note**
>
> 커넥터가 작동하는 데 필요한 자격 증명만 사용하십시오. 도메인 관리자와 같은 권한 있는 자격 증명은 사용하지 마십시오.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 각 문서가 Amazon FSx (Windows) 및 동일한 색인에 사용하려는 다른 데이터 소스에서 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Amazon FSx (Windows) 인증 자격 증명을 AWS Secrets Manager 암호에 저장하고 API를 사용하는 경우 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Amazon FSx (Windows) 데이터 원본을 연결할 Amazon Kendra 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Amazon FSx (Windows) 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Amazon FSx (Windows) 데이터 원본의 필수 세부 정보를 제공해야 합니다. 에 대해 아직 구성 Amazon FSx (Windows) 하지 않은 경우 을 Amazon Kendra참조하십시오필수 조건.

Console

Amazon FSx (Windows) 파일 시스템에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Amazon FSx (Windows) 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Amazon FSx (Windows) 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. Amazon FSx (Windows) 파일 시스템 ID - (Windows) 에서 가져온 기존 파일 시스템 ID를 드롭다운에서 선택합니다. Amazon FSx 또는 Amazon FSx (Windows) 파일 시스템을 만들 수도 있습니다. 파일 시스템 ID는 Amazon FSx (Windows) 콘솔의 파일 시스템 대시보드에서 찾을 수 있습니다.

b. 인증 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

c. 인증 - 기존 AWS Secrets Manager 암호를 선택하거나 파일 시스템 자격 증명을 저장할 새 암호를 생성합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

사용자 이름 및 암호의 인증 자격 증명을 저장하는 암호를 제공하십시오. 사용자 이름에는 DNS 도메인 이름이 포함되어야 합니다. 예: user@corp.example.com.

비밀번호를 저장하고 추가하세요.

d. 가상 사설 클라우드 (VPC) - (Windows) Amazon VPC 가 있는 곳을 선택해야 합니다. Amazon FSx VPC 서브넷과 보안 그룹을 포함합니다. 구성을 참조하십시오. Amazon VPC

e. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

f. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

   a. 동기화 범위, 정규식 패턴 - 정규 표현식 패턴을 추가하여 특정 파일을 포함하거나 제외합니다.

   b. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

      - 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
      - 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

   c. 동기화 실행 일정 - 빈도에서 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

   d. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. Amazon Kendra 생성된 파일의 기본 필드 중에서 색인에 매핑할 필드를 선택합니다. 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

   b. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon FSx (Windows) 파일 시스템에 Amazon Kendra 연결하려면

TemplateConfiguration API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 FSX 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 파일 시스템 ID - Amazon FSx (Windows) 파일 시스템의 식별자입니다. 파일 시스템 ID는 Amazon FSx (Windows) 콘솔의 파일 시스템 대시보드에서 찾을 수 있습니다.
- 파일 시스템 유형 - 파일 시스템 유형을 WINDOWS로 지정합니다.
- Virtual Private Cloud(VPC) - CreateDataSource를 호출할 때 VpcConfiguration을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.

  > **Note**
  >
  > Amazon FSx (Windows) Amazon VPC 가 있는 위치를 선택해야 합니다. VPC 서브넷과 보안 그룹을 포함합니다.

- 동기화 모드 - 데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 인덱스를 업데이트하는 방법을 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) —사용자 (Windows) 계정의 Amazon FSx 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공하십시오. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
```

```
    "username": "user@corp.example.com",
    "password": "password"
}
```

- IAM 역할 - 역할에 Secrets Manager 보안 액세스 권한을 제공하고 (Windows) 커넥터 및 Amazon FSx (Windows) 커넥터에 필요한 퍼블릭 API를 호출하도록 호출할 `RoleArn` 시기를 지정합니다. `CreateDataSource` IAM Amazon Kendra자세한 내용은 Amazon FSx (Windows) 데이터 원본의IAM 역할을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- 포함 및 제외 필터 - 특정 파일을 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 액세스 제어 목록 (ACL) - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 ACL 정보를 크롤링할지 여부를 지정합니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

> **Note**
>
> 사용자에 대한 사용자 컨텍스트 필터링을 테스트하려면 쿼리를 실행할 때 DNS 도메인 이름을 사용자 이름의 일부로 포함해야 합니다. Active Directory 도메인의 관리 권한이 있어야 합니다. 그룹 이름에 대한 사용자 컨텍스트 필터링을 테스트할 수도 있습니다.

- 필드 매핑 - Amazon FSx (Windows) 데이터 원본 필드를 인덱스 필드에 매핑하도록 선택합니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Amazon FSx (Windows) 템플릿 스키마를 참조하십시오.

자세히 알아보기

Amazon FSx (Windows) 데이터 Amazon Kendra 소스와의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- Amazon FSx (Windows) 용 Amazon Kendra 커넥터를 사용하여 Windows 파일 시스템의 비정형 데이터를 안전하게 검색합니다. Windows File Server

Amazon FSx (NetApp ONTAP)

Amazon FSx (NetApp ONTAP) 는 공유 스토리지 기능을 제공하는 완전 관리형 클라우드 기반 파일 서버 시스템입니다. Amazon FSx (NetApp ONTAP) 사용자인 경우 (ONTAP) 데이터 소스를 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다 Amazon FSx . NetApp

Amazon Kendra 콘솔 또는 API를 Amazon Kendra 사용하여 Amazon FSx (NetApp ONTAP) 데이터 소스에 연결할 수 있습니다. TemplateConfiguration

Amazon Kendra Amazon FSx (NetApp ONTAP) 데이터 소스 커넥터 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침

지원되는 기능

Amazon Kendra Amazon FSx (NetApp ONTAP) 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함 및 제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Amazon FSx (NetApp ONTAP) 데이터 소스를 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 (ONTAP) 및 의 세부 정보를 확인하십시오. Amazon FSx NetApp AWS 계정

Amazon FSx (NetApp ONTAP) 의 경우 다음이 있는지 확인하십시오.

- 읽기 및 마운트 권한으로 Amazon FSx (NetApp ONTAP) 을 설정합니다.
- 파일 시스템 ID를 기록했습니다. 파일 시스템 ID는 Amazon FSx (NetApp ONTAP) 콘솔의 파일 시스템 대시보드에서 찾을 수 있습니다.
- 파일 시스템에 사용되는 SVM (스토리지 가상 머신) ID를 기록했습니다. Amazon FSx (NetApp ONTAP) 콘솔의 파일 시스템 대시보드로 이동하여 파일 시스템 ID를 선택한 다음 스토리지 가상 시스템을 선택하면 SVM ID를 찾을 수 있습니다.
- Amazon FSx (NetApp ONTAP) 파일 시스템이 Amazon VPC 있는 위치를 사용하여 가상 사설 클라우드를 구성했습니다.
- 사용자 계정의 Amazon FSx (NetApp ONTAP) 인증 자격 증명을 기록해 두었습니다. Active Directory 여기에는 Active Directory 사용자 이름과 DNS 도메인 이름 (예: user@corp.example.com) 및 암호가 포함됩니다. (NetApp ONTAP) 파일 시스템에 NFS (네트워크 파일 시스템) 프로토콜을 사용하는 경우 인증 자격 증명에는 왼쪽 ID, 오른쪽 ID 및 사전 공유 키가 포함됩니다. Amazon FSx

> **Note**
>
> 커넥터가 작동하는 데 필요한 자격 증명만 사용하십시오. 도메인 관리자와 같은 권한 있는 자격 증명은 사용하지 마십시오.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 각 문서가 Amazon FSx (NetApp ONTAP) 및 동일한 인덱스에 사용할 다른 데이터 소스에서 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Amazon FSx (NetApp ONTAP) 인증 자격 증명을 AWS Secrets Manager 암호에 저장하고 API를 사용하는 경우 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Amazon FSx (NetApp ONTAP) 데이터 소스를 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 생성할 수 있습니다. Amazon Kendra API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Amazon FSx (NetApp ONTAP) 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Amazon FSx (NetApp ONTAP) 데이터 소스의 필수 세부 정보를 제공해야 합니다. 에 대해 Amazon FSx (NetApp ONTAP) 를 아직 구성하지 않은 경우 을 Amazon Kendra참조하십시오. 필수 조건

Console

Amazon FSx (NetApp ONTAP) 파일 시스템에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Amazon FSx (NetApp ONTAP) 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Amazon FSx (NetApp ONTAP) 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. 출처 - 파일 시스템 정보를 제공합니다.

- 파일 시스템 프로토콜 - Amazon FSx (NetApp ONTAP) 파일 시스템의 프로토콜을 선택합니다. 공용 인터넷 파일 시스템 (CIFS) 프로토콜 또는 Linux용 네트워크 파일 시스템 (NFS) 프로토콜을 선택할 수 있습니다.
- Amazon FSx (NetApp ONTAP) 파일 시스템 ID - (ONTAP) 에서 가져온 기존 파일 시스템 ID를 드롭다운에서 선택합니다. Amazon FSx NetApp 또는 (ONTAP) 파일 시스템을 생성합니다.Amazon FSx NetApp 파일 시스템 ID는 Amazon FSx (NetApp ONTAP) 콘솔의 파일 시스템 대시보드에서 찾을 수 있습니다.
- SVM ID (Amazon FSx (NetApp ONTAP) NetApp ONTAP 전용) - 사용자 (ONTAP) 의 스토리지 가상 시스템 (SVM) ID를 제공하십시오. Amazon FSx NetApp NetApp ONTAP Amazon FSx (NetApp ONTAP) 콘솔의 파일 시스템 대시보드로 이동하여 파일 시스템 ID를 선택하고 스토리지 가상 시스템을 선택하면 SVM ID를 찾을 수 있습니다.

b. 인증 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

c. 인증 - 기존 AWS Secrets Manager 암호를 선택하거나 파일 시스템 자격 증명을 저장할 새 암호를 생성합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

   사용자 이름과 암호의 인증 자격 증명을 저장하는 암호를 제공하십시오. 사용자 이름에는 DNS 도메인 이름이 포함되어야 합니다. 예: user@corp.example.com.

   Amazon FSx (NetApp ONTAP) 파일 시스템에 NFS 프로토콜을 사용하는 경우 왼쪽 ID, 오른쪽 ID 및 사전 공유 키의 인증 자격 증명을 저장하는 암호를 제공하십시오.

   암호를 저장하고 추가합니다.

d. 가상 사설 클라우드 (VPC) —(ONTAP) Amazon VPC 가 있는 위치를 선택해야 합니다. Amazon FSx NetApp VPC 서브넷과 보안 그룹을 포함합니다. 구성을 참조하십시오. Amazon VPC

e. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

f. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

   a. 동기화 범위, 정규식 패턴 - 정규 표현식 패턴을 추가하여 특정 파일을 포함하거나 제외합니다.

   b. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

      - 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
      - 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

   c. 동기화 실행 일정 - 빈도에서 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

   d. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. Amazon Kendra 생성된 파일의 기본 필드 중에서 색인에 매핑할 필드를 선택합니다. 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

   b. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon FSx (NetApp ONTAP) 파일 시스템에 Amazon Kendra 연결하려면

TemplateConfiguration API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 소스 - TemplateConfigurationJSON 스키마를 사용할 FSXONTAP 때와 같이 데이터 소스 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 파일 시스템 ID - Amazon FSx (NetApp ONTAP) 파일 시스템의 식별자입니다. 파일 시스템 ID는 Amazon FSx (NetApp ONTAP) 콘솔의 파일 시스템 대시보드에서 찾을 수 있습니다.
- SVM ID - 파일 시스템에 사용되는 스토리지 가상 시스템 (SVM) ID입니다. Amazon FSx (NetApp ONTAP) 콘솔의 파일 시스템 대시보드로 이동하여 파일 시스템 ID를 선택한 다음 스토리지 가상 시스템을 선택하면 SVM ID를 찾을 수 있습니다.
- 프로토콜 유형 —CIFS (공용 인터넷 파일 시스템) 프로토콜을 사용할지, Linux용 네트워크 파일 시스템 (NFS) 프로토콜을 사용할지를 지정합니다.
- 파일 시스템 유형 --파일 시스템 유형을 둘 중 하나로 지정합니다. FSXONTAP
- Virtual Private Cloud(VPC) - CreateDataSource를 호출할 때 VpcConfiguration을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.

> **Note**
>
> Amazon FSx (NetApp ONTAP) Amazon VPC 가 있는 위치를 선택해야 합니다. VPC 서브넷과 보안 그룹을 포함합니다.

- 비밀 Amazon 리소스 이름 (ARN) —사용자 (ONTAP) 계정의 Amazon FSx 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공하십시오. NetApp 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "username": "user@corp.example.com",
    "password": "password"
}
```

Amazon FSx (NetApp ONTAP) 파일 시스템에 NFS 프로토콜을 사용하는 경우 암호는 다음 키가 있는 JSON 구조에 저장됩니다.

```
{
    "leftId": "left ID",
    "rightId": "right ID",
    "preSharedKey": "pre-shared key"
}
```

- IAM 역할 —역할에 Secrets Manager 보안 액세스 권한을 제공하고 Amazon FSx (NetApp ONTAP) 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. `CreateDataSource` IAM Amazon Kendra자세한 내용은 Amazon FSx (NetApp ONTAP) 데이터 소스의IAM 역할을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - `FORCED_FULL_CRAWL`모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - `FULL_CRAWL`데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 포함 및 제외 필터 - 특정 파일을 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- ACL (액세스 제어 목록) - ACL이 있고 액세스 제어에 사용하려는 경우 문서의 ACL 정보를 크롤링할지 여부를 지정합니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

> **Note**
>
> 사용자에 대한 사용자 컨텍스트 필터링을 테스트하려면 쿼리를 실행할 때 DNS 도메인 이름을 사용자 이름의 일부로 포함해야 합니다. Active Directory 도메인의 관리 권한이 있어야 합니다. 그룹 이름에 대한 사용자 컨텍스트 필터링을 테스트할 수도 있습니다.

- 필드 매핑 - Amazon FSx (NetApp ONTAP) 데이터 소스 필드를 인덱스 필드에 매핑하도록 선택합니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성할 기타 중요한 JSON 키 목록은 Amazon FSx (NetApp ONTAP) 템플릿 스키마를 참조하십시오.

Amazon RDS/Aurora

데이터베이스 데이터 소스를 사용하여 데이터베이스에 저장된 문서를 인덱싱할 수 있습니다. 데이터베이스의 연결 정보를 제공한 후 문서를 Amazon Kendra 연결하고 색인을 생성합니다.

Amazon Kendra 지원되는 데이터베이스는 다음과 같습니다.

- Amazon Aurora MySQL
- Amazon Aurora PostgreSQL
- Amazon RDS MySQL용
- Amazon RDS PostgreSQL의 경우

> **Note**
>
> Serverless Aurora 데이터베이스는 지원되지 않습니다.

Important

이 Amazon RDS/Aurora 커넥터는 2023년 말 사용이 중단될 예정입니다.
Amazon Kendra 이제 새 데이터베이스 데이터 소스 커넥터를 지원합니다. 사용 환경을 개선하려면 사용 사례에 맞게 다음과 같은 새 커넥터를 선택하는 것이 좋습니다.

- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon RDS (MySQL)
- Amazon RDS (마이크로소프트 SQL 서버)
- Amazon RDS (오라클)
- Amazon RDS (PostgreSQL)
- IBM DB2
- Microsoft SQL Server
- MySQL
- Oracle Database
- PostgreSQL

Amazon Kendra 콘솔과 API를 Amazon Kendra 사용하여 데이터베이스 데이터 소스에 연결할 수 있습니다. DatabaseConfiguration

Amazon Kendra 데이터베이스 데이터 원본 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침

지원되는 기능

Amazon Kendra 데이터베이스 데이터 원본 커넥터는 다음 기능을 지원합니다.

- 필드 매핑

- 사용자 컨텍스트 필터링
- Virtual Private Cloud(VPC)

필수 조건

데이터베이스 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 데이터베이스와 AWS 계정에서 다음과 같이 변경하십시오.

데이터베이스에서 다음을 확인하세요.

- 데이터베이스 사용자 이름과 암호의 기본 보안 인증 정보를 기록했습니다.
- 호스트 이름, 포트 번호, 호스트 주소, 데이터베이스 이름, 문서 데이터가 포함된 데이터 테이블 이름을 복사했습니다. PostgreSQL의 경우 데이터 테이블은 공개 테이블 또는 공용 스키마여야 합니다.

> **Note**
>
> 호스트와 포트는 인터넷에서 데이터베이스 서버를 찾을 수 Amazon Kendra 있는 위치를 알려줍니다. 데이터베이스 이름과 테이블 이름은 데이터베이스 서버에서 문서 데이터를 찾을 수 Amazon Kendra 있는 위치를 알려줍니다.

- 문서 데이터가 포함된 데이터 테이블의 열 이름을 복사했습니다. 문서 ID, 문서 본문, 문서 변경 여부를 감지할 열(예: 마지막으로 업데이트한 열), 사용자 지정 인덱스 필드에 매핑되는 선택적 데이터 테이블 열을 포함해야 합니다. Amazon Kendra 예약된 필드 이름을 테이블 열에 매핑할 수도 있습니다.
- 데이터베이스 엔진 유형 정보 (예: MySQL 또는 다른 유형에 사용하는지 여부) Amazon RDS 를 복사했습니다.
- 데이터베이스 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형 및 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- 데이터베이스 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 데이터베이스 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

데이터베이스 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 데이터베이스 데이터 원본의 필수 세부 정보를 제공해야 합니다. 에 대한 Amazon Kendra데이터베이스를 아직 구성하지 않은 경우 을 참조하십시오필수 조건.

Console

데이터베이스에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.

4. 데이터 원본 추가 페이지에서 데이터베이스 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 데이터베이스 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. 엔드포인트 - DNS 호스트 이름, IPv4 주소 또는 IPv6 주소.
   b. 포트 - 포트 번호.
   c. 데이터베이스 - 데이터베이스 이름.
   d. 테이블 이름 - 테이블 이름.
   e. 인증 유형에서 기존 인증과 신규 인증 정보 중 하나를 선택하여 데이터베이스 보안 인증 정보를 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.
      - AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.
        A. 보안 암호 이름 - 보안 암호의 이름. 비밀 이름에는 접두사 AmazonKendra '-database-'가 자동으로 추가됩니다.
        B. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스 계정의 보안 인증 값을 입력합니다.
        C. 인증 저장을 선택합니다.
   f. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

> **Note**
>
> 프라이빗 서브넷을 사용해야 합니다. RDS 인스턴스가 VPC의 퍼블릭 서브넷에 있는 경우, 퍼블릭 서브넷의 NAT 게이트웨이에 대한 아웃바운드 액세스 권한이 있는 프라이빗 서브넷을 생성할 수 있습니다. VPC 구성에 제공되는 서브넷은 미국 서부(오레곤), 미국 동부(버니지아), EU(아일랜드)에 있어야 합니다.

g. IAM 역할 - 기존 역할을 선택하거나 새 IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 IAM 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

h. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 사용 사례에 따라 Aurora MySQL, MySQL, Aurora PostgreSQL, PostgreSQL 중에서 선택합니다.

b. SQL 식별자를 큰따옴표로 묶기 - SQL 식별자를 큰따옴표로 묶으려면 이 옵션을 선택합니다. 예: “columnName”

c. ACL 열 및 변경 감지 열 - 변경 감지에 사용할 열 (예: 마지막 업데이트 열) 과 액세스 제어 목록을 구성합니다. Amazon Kendra

d. 동기화 실행 일정에서 빈도에 대해 —데이터 원본과 Amazon Kendra 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. Amazon Kendra 기본 필드 매핑 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다. `document_id` 및 `document_body`에 대한 데이터베이스 열 값을 추가해야 합니다.

b. 사용자 지정 필드 매핑 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

데이터베이스에 Amazon Kendra 연결하려면

DatabaseConfigurationAPI를 다음과 같이 지정해야 합니다.

- ColumnConfiguration—인덱스가 데이터베이스에서 문서 정보를 가져와야 하는 위치에 대한 정보. 자세한 내용은 ColumnConfiguration를 참조하세요. `DocumentDataColumnName`(문서 본문 또는 본문) 및 `DocumentIdColumnName`, `ChangeDetectingColumn`(예: 최근 업데이트된 열) 필드를 지정해야 합니다. `DocumentIdColumnName` 필드에 매핑된 열은 정수 열이어야 합니다. 다음 예제에서는 데이터베이스 데이터 소스의 단순한 열 구성을 보여줍니다.

```
"ColumnConfiguration": {
    "ChangeDetectingColumns": [
        "LastUpdateDate",
        "LastUpdateTime"
    ],
    "DocumentDataColumnName": "TextColumn",
    "DocumentIdColumnName": "IdentifierColumn",
    "DocoumentTitleColumnName": "TitleColumn",
    "FieldMappings": [
        {
            "DataSourceFieldName": "AbstractColumn",
            "IndexFieldName": "Abstract"
        }
    ]
}
```

- ConnectionConfiguration—데이터베이스에 연결하는 데 필요한 구성 정보. 자세한 내용은 ConnectionConfiguration를 참조하세요.
- DatabaseEngineType—데이터베이스를 실행하는 데이터베이스 엔진 유형. 의 `DatabaseHost` 필드는 데이터베이스의 Amazon Relational Database Service (Amazon RDS) 인스턴스 `ConnectionConfiguration` 엔드포인트여야 합니다. 클러스터 엔드포인트를 사용하지 마세요.

- 비밀 Amazon 리소스 이름 (ARN) - 데이터베이스 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. 보안 암호에는 다음 키가 있는 JSON 구조가 저장됩니다.

```
{
    "username": "user name",
    "password": "password"
}
```

다음 예제에서는 보안 암호 ARN을 포함한 데이터베이스 구성을 보여줍니다.

```
"DatabaseConfiguration": {
"ConnectionConfiguration": {
"DatabaseHost": "host.subdomain.domain.tld",
        "DatabaseName": "DocumentDatabase",
        "DatabasePort": 3306,
        "SecretArn": "arn:aws:secretmanager:region:account ID:secret/secret name",
        "TableName": "DocumentTable"
    }
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 — 역할에 Secrets Manager 보안 액세스 권한을 제공하고 데이터베이스 IAM 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. `CreateDataSource` Amazon Kendra자세한 내용은 데이터베이스 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - 데이터 소스 구성의 일부로 `VpcConfiguration`를 지정합니다. VPC 사용을 위한 Amazon Kendra 구성을 참조하세요.

> **Note**
>
> 프라이빗 서브넷만 사용해야 합니다. RDS 인스턴스가 VPC의 퍼블릭 서브넷에 있는 경우, 퍼블릭 서브넷의 NAT 게이트웨이에 대한 아웃바운드 액세스 권한이 있는 프라이빗 서브넷을 생성할 수 있습니다. VPC 구성에 제공되는 서브넷은 미국 서부(오레곤), 미국 동부(버니지아), EU(아일랜드)에 있어야 합니다.

- 필드 매핑 - 데이터베이스 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

Amazon RDS (마이크로소프트 SQL 서버)

SQL Server는 Microsoft에서 개발한 데이터베이스 관리 시스템입니다. Amazon RDS SQL Server의 경우 클라우드에서 SQL Server 배포를 쉽게 설정, 운영 및 확장할 수 있습니다. Amazon RDS (Microsoft SQL Server) 사용자인 경우 Amazon RDS (Microsoft SQL Server) 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다. Amazon Kendra JDBC 데이터 소스 커넥터는 마이크로소프트 SQL 서버 2019를 지원합니다.

Amazon Kendra 콘솔과 TemplateConfigurationAPI를 사용하여 Amazon RDS (Microsoft SQL Server) 데이터 원본에 연결할 Amazon Kendra 수 있습니다.

Amazon Kendra Amazon RDS (Microsoft SQL Server) 데이터 소스 커넥터의 문제를 해결하려면 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능

- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

를 사용하여 Amazon RDS (Microsoft SQL Server) 데이터 원본을 Amazon Kendra 인덱싱하려면 먼저 Amazon RDS (Microsoft SQL Server) 및 AWS 계정에서 이러한 변경을 수행하십시오.

Amazon RDS (Microsoft SQL Server) 에서 다음 항목이 있는지 확인합니다.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
> 가장 좋은 방법은 읽기 Amazon Kendra 전용 데이터베이스 자격 증명을 제공하는 것입니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다.
- 각 문서가 Amazon RDS (Microsoft SQL Server) 및 동일한 인덱스에 사용하려는 다른 데이터 원본에서 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음 사항이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Amazon RDS (Microsoft SQL Server) 인증 자격 증명을 AWS Secrets Manager 시크릿에 저장했고, API를 사용하는 경우 시크릿의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Amazon RDS (Microsoft SQL Server) 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Amazon RDS (Microsoft SQL Server) 데이터 원본에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Amazon RDS (Microsoft SQL Server) 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 Amazon RDS (Microsoft SQL Server) 를 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon RDS (마이크로소프트 SQL 서버) Amazon Kendra 에 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.

4. 데이터 원본 추가 페이지에서 Amazon RDS (Microsoft SQL Server) 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Amazon RDS (Microsoft SQL Server) 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. 소스에서 다음 정보를 입력합니다.
   b. 호스트 - 데이터베이스 호스트 이름을 입력합니다.
   c. 포트 - 데이터베이스 포트를 입력합니다.
   d. 인스턴스 - 데이터베이스 인스턴스를 입력합니다.
   e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.
   f. 인증에서 - 다음 정보를 입력합니다.
      - AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 Secrets Manager 암호를 만들어 Amazon RDS (Microsoft SQL Server) 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.
        A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.
           I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 'AmazonKendra-Amazon RDS (Microsoft SQL Server) -'라는 접두사가 자동으로 추가됩니다.
           II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.
        B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명과 인덱스 콘텐츠에 액세스합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.

> **Note**
>
> 테이블 이름에 특수 문자 (영숫자가 아닌) 가 포함된 경우 테이블 이름 주위에 대괄호를 사용해야 합니다. 예를 들어, `[] ##*# #####`. my-database-table

- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.

- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon RDS (마이크로소프트 SQL 서버) Amazon Kendra 에 연결하려면

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 sqlserver로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.

  > **Note**
  >
  > 테이블 이름에 특수 문자 (영숫자가 아닌) 가 포함된 경우 테이블 이름 주위에 대괄호를 사용해야 합니다. 예를 들어, *[] ##*# #####*. my-database-table

- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - (Amazon RDS Microsoft SQL Server) 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공하십시오. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
```

```
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 - `CreateDataSource` Secrets Manager 암호에 액세스할 수 있는 권한을 IAM 역할에 제공하고 (Amazon RDS Microsoft SQL Server) 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Amazon RDS (Microsoft SQL Server) 데이터 원본의IAM 역할을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 - 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - (Amazon RDS Microsoft SQL Server) 데이터 원본 필드를 인덱스 필드에 매핑하도록 Amazon Kendra 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 해당 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Amazon RDS (Microsoft SQL Server) 템플릿 스키마를 참조하세요.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Amazon RDS (MySQL)

Amazon RDS (Amazon Relational Database Service) 는 클라우드에서 AWS 관계형 데이터베이스를 더 쉽게 설정, 운영 및 확장할 수 있게 해주는 웹 서비스입니다. Amazon RDS 사용자인 경우 Amazon RDS (MySQL) 데이터 소스를 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다. Amazon Kendra 데이터 소스 커넥터는 Amazon RDS MySql 5.6, 5.7 및 8.0을 지원합니다.

Amazon Kendra 콘솔과 API를 사용하여 Amazon RDS (MySQL) 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Amazon RDS (MySQL)데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화

- Virtual Private Cloud(VPC)

필수 조건

Amazon RDS (MySQL)데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Amazon RDS (MySQL) 및 AWS 계정에서 다음과 같이 변경하십시오.

Amazon RDS (MySQL)에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
>
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다. Amazon RDS 콘솔에서 이 정보를 찾을 수 있습니다.
- Amazon RDS (MySQL) 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Amazon RDS (MySQL) 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Amazon RDS (MySQL) 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Amazon RDS (MySQL)데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Amazon RDS (MySQL) 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 Amazon RDS (MySQL) 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Amazon RDS (MySQL)

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 Amazon RDS (MySQL)커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Amazon RDS (MySQL)커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. 소스에서 다음 정보를 입력합니다.

b. 호스트 - 데이터베이스 호스트 URL(예: `http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`)을 입력합니다.

c. 포트 - 데이터베이스 포트(예: `5432`)를 입력합니다.

d. 인스턴스 - 데이터베이스 인스턴스(예: `postgres`)를 입력합니다.

e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.

f. 인증에서 - 다음 정보를 입력합니다.

- AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Amazon RDS (MySQL) 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

  A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

     I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Amazon RDS (MySQL) -'가 자동으로 추가됩니다.

     II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

  B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. SQL 쿼리는 32KB 미만이어야 하며 세미콜론 (;) 을 포함하지 않아야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra Amazon RDS (MySQL)

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 mySql로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. Amazon RDS (MySQL) 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 - CreateDataSource Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 RoleArn 시기를 지정합니다.

Amazon RDS (MySQL) Amazon Kendra자세한 내용은 Amazon RDS (MySQL) 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 필드 매핑 - Amazon RDS (MySQL) 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 해당 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서의 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

구성해야 할 기타 중요한 JSON 키 목록은 Amazon RDS (MySQL) 템플릿 스키마를 참조하세요.

참고

- 업데이트된 콘텐츠를 확인할 때 Amazon Kendra 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Amazon RDS (Oracle)

Amazon RDS (Amazon Relational Database Service) 는 클라우드에서 AWS 관계형 데이터베이스를 더 쉽게 설정, 운영 및 확장할 수 있게 해주는 웹 서비스입니다. Amazon RDS (Oracle)사용자인 경우 Amazon RDS (Oracle) 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다. Amazon Kendra Amazon RDS (Oracle)데이터 소스 커넥터는 Amazon RDS 오라클 데이터베이스 21c, 오라클 데이터베이스 19c, 오라클 데이터베이스 12c를 지원합니다.

Amazon Kendra 콘솔과 API를 사용하여 Amazon RDS (Oracle) 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Amazon RDS (Oracle)데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Amazon RDS (Oracle)데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Amazon RDS (Oracle) 및 AWS 계정에서 다음과 같이 변경하십시오.

Amazon RDS (Oracle)에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
>
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다.
- Amazon RDS (Oracle) 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Amazon RDS (Oracle) 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Amazon RDS (Oracle) 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Amazon RDS (Oracle)데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Amazon RDS (Oracle) 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 Amazon RDS (Oracle) 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Amazon RDS (Oracle)

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 Amazon RDS (Oracle)커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Amazon RDS (Oracle)커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. 소스에서 다음 정보를 입력합니다.
   b. 호스트 - 데이터베이스 호스트 이름을 입력합니다.
   c. 포트 - 데이터베이스 포트를 입력합니다.

d. 인스턴스 - 데이터베이스 인스턴스를 입력합니다.

e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.

f. 인증에서 - 다음 정보를 입력합니다.

- AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Amazon RDS (Oracle) 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

  A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

     I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Amazon RDS (Oracle) -'가 자동으로 추가됩니다.

     II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

  B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.

- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

   b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

   c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra Amazon RDS (Oracle)

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 oracle로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. Amazon RDS (Oracle) 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 - `CreateDataSource` Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Amazon RDS (Oracle) Amazon Kendra자세한 내용은 Amazon RDS (Oracle) 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Amazon RDS (Oracle) 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Amazon RDS (오라클) 템플릿 스키마를 참조하세요.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Amazon RDS (PostgreSQL)

Amazon RDS AWS 클라우드에서 관계형 데이터베이스를 더 쉽게 설정, 운영 및 확장할 수 있게 해주는 웹 서비스입니다. Amazon RDS 사용자인 경우 Amazon RDS (PostgreSQL) 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다. Amazon Kendra Amazon RDS (PostgreSQL)데이터 소스 커넥터는 PostgreSQL 9.6을 지원합니다.

Amazon Kendra 콘솔과 API를 사용하여 Amazon RDS (PostgreSQL) 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Amazon RDS (PostgreSQL)데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건

- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Amazon RDS (PostgreSQL)데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Amazon RDS (PostgreSQL) 및 AWS 계정에서 다음과 같이 변경하십시오.

Amazon RDS (PostgreSQL)에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다. Amazon RDS 콘솔에서 이 정보를 찾을 수 있습니다.
- Amazon RDS (PostgreSQL) 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형 및 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Amazon RDS (PostgreSQL) 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Amazon RDS (PostgreSQL) 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Amazon RDS (PostgreSQL)데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Amazon RDS (PostgreSQL) 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 Amazon RDS (PostgreSQL) 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Amazon RDS (PostgreSQL)

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.

4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 Amazon RDS (PostgreSQL)커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Amazon RDS (PostgreSQL)커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

    a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
    b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
    c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
    d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
    e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

    a. 소스에서 다음 정보를 입력합니다.
    b. 호스트 - 데이터베이스 호스트 URL(예: `http://`*`instance URL`*`.`*`region`*`.rds.amazonaws.com`)을 입력합니다.
    c. 포트 - 데이터베이스 포트(예: `5432`)를 입력합니다.
    d. 인스턴스 - 데이터베이스 인스턴스(예: `postgres`)를 입력합니다.
    e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.
    f. 인증에서 - 다음 정보를 입력합니다.

        - AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Amazon RDS (PostgreSQL) 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

            A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

                I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Amazon RDS (PostgreSQL) -'가 자동으로 추가됩니다.
                II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

            B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. SQL 쿼리는 32KB 미만이어야 하며 세미콜론 (;) 을 포함하지 않아야 합니다. Amazon Kendra 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.

- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra Amazon RDS (PostgreSQL)

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 postgresql로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. Amazon RDS (PostgreSQL) 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 - CreateDataSource Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 RoleArn 시기를 지정합니다.

Amazon RDS (PostgreSQL) Amazon Kendra자세한 내용은 Amazon RDS (PostgreSQL) 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Amazon RDS (PostgreSQL) 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Amazon RDS (PostgreSQL) 템플릿 스키마를 참조하세요.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Amazon S3

Amazon S3 데이터를 버킷 내에 객체로 저장하는 오브젝트 스토리지 서비스입니다. 를 사용하여 문서의 Amazon S3 버킷 리포지토리를 Amazon Kendra 인덱싱할 수 있습니다.

Warning

Amazon Kendra Amazon Kendra 보안 주체에게 S3 버킷과 상호 작용할 권한을 부여하는 버킷 정책을 사용하지 않습니다. 대신 IAM 역할을 사용합니다. 실수로 임의의 주체에게 권한을 부여하여 데이터 보안 문제가 발생하지 않도록 해당 Amazon Kendra 구성원을 신뢰할 수 있는 구성원으로 버킷 정책에 포함하지 않도록 하세요. 하지만 여러 계정에서 Amazon S3 버킷을 사용하도록 버킷 정책을 추가할 수 있습니다. 자세한 내용은 여러 계정에서 Amazon S3 를 사용할 수 있는 정책(S3 IAM 역할 탭의 데이터 소스의IAM 역할 아래)을 참조하세요. S3 데이터 원본의 IAM 역할에 대한 자세한 내용은 역할을 참조하십시오.IAM

Note

Amazon Kendra 이제 업그레이드된 Amazon S3 커넥터를 지원합니다.
콘솔이 자동으로 업그레이드되었습니다. 콘솔에서 새로 만드는 모든 커넥터는 업그레이드된 아키텍처를 사용합니다. API를 사용하는 경우 이제 TemplateConfiguration개체 대신 개체를 사용하여 커넥터를 구성해야 합니다. `S3DataSourceConfiguration`
이전 콘솔과 API 아키텍처를 사용하여 구성된 커넥터는 구성된 대로 계속 작동합니다. 하지만 편집하거나 업데이트할 수는 없습니다. 커넥터 구성을 편집하거나 업데이트하려면 새 커넥터를 만들어야 합니다.
커넥터 워크플로를 업그레이드된 버전으로 마이그레이션하는 것이 좋습니다. 이전 아키텍처를 사용하여 구성된 커넥터에 대한 지원은 2024년 6월에 종료될 예정입니다.

Amazon Kendra 콘솔 또는 TemplateConfigurationAPI를 사용하여 Amazon S3 데이터 소스에 연결할 수 있습니다.

Note

데이터 소스에 대한 동기화 상태 보고서를 생성하려면 Amazon S3 데이터 소스 문제 해결을 참조하십시오.

Amazon Kendra S3 데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- Amazon S3 데이터 소스 생성
- Amazon S3 문서 메타데이터
- Amazon S3 데이터 소스에 대한 액세스 제어
- 데이터 Amazon VPC 소스와 Amazon S3 함께 사용

지원되는 기능

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

S3 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하기 전에 먼저 S3와 AWS 계정에서 다음과 같이 변경하십시오.

S3에서 다음 사항을 갖추었는지 확인하세요.

- Amazon S3 버킷 이름을 복사했습니다.

> **Note**
>
> 버킷은 인덱스와 같은 리전에 있어야 하고 Amazon Kendra 인덱스에는 문서가 포함된 버킷에 액세스할 수 있는 권한이 있어야 합니다.

- S3 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

AWS 계정에 다음이 있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

기존 IAM 역할이 없는 경우 S3 데이터 소스를 연결할 때 콘솔을 사용하여 새 IAM 역할을 생성할 수 있습니다. Amazon Kendra API를 사용하는 경우 기존 IAM 역할의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

S3 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 S3 데이터 원본의 필수 세부 정보를 제공해야 합니다. 에 대해 아직 S3를 구성하지 않은 경우 을 Amazon Kendra참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Amazon S3

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 S3 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 S3 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 선택적 정보를 입력합니다.

a. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

b. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

c. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 데이터 원본 위치의 경우 - 데이터가 저장되는 Amazon S3 버킷의 경로를 지정합니다. S3 찾아보기를 선택하여 S3 버킷을 선택합니다.

b. 최대 파일 크기 - 이 한도 미만의 파일만 크롤링하려면 제한을 MB 단위로 지정합니다. Amazon Kendra 허용할 수 있는 최대 파일 크기는 50MB입니다.

c. (선택 사항) 메타데이터 파일 접두사 폴더 위치의 경우 - 필드/속성 및 기타 문서 메타데이터가 저장되는 폴더의 경로를 지정합니다. S3 찾아보기를 선택하여 메타데이터 폴더를 찾습니다.

d. (선택 사항) 액세스 제어 목록 구성 파일 위치의 경우 - 사용자의 JSON 구조가 포함된 파일의 경로와 문서에 대한 액세스 권한을 지정합니다. S3 찾아보기를 선택하여 ACL 파일을 찾습니다.

e. (선택 사항) 해독 키 선택 - 암호 해독 키를 사용하려면 선택합니다. 기존 AWS KMS 키를 사용하도록 선택할 수 있습니다.

f. (선택 사항) 추가 구성의 경우 - 패턴을 추가하여 특정 파일을 포함하거나 제외합니다. 모든 경로는 데이터 소스 위치 S3 버킷과 관련이 있습니다.

g. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

h. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

i. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 선택적 정보를 입력합니다.

a. 기본 필드 매핑 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만들려면 선택합니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra Amazon S3

TemplateConfigurationAPI를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 S3 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- BucketName—문서가 들어 있는 버킷의 이름.
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.

  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- IAM 역할 - CreateDataSource Secrets Manager 암호에 액세스할 수 있는 권한을 IAM 역할을 제공하고 S3 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 RoleArn 시기를 지정합니다. Amazon Kendra자세한 내용은 S3 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - CreateDataSource를 호출할 때 VpcConfiguration을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 특정 파일 이름, 파일 유형, 파일 경로를 포함할지 제외할지 여부를 지정합니다. 글로브 패턴 (와일드카드 패턴을 지정된 패턴과 일치하는 경로 이름 목록으로 확장할 수 있는 패턴) 을 사용합니다. 예제는 AWS CLI 명령 참조의 제외 및 포함 필터 사용을 참조하십시오.
- 문서 메타데이터 및 액세스 제어 구성 - 소스 URI, 문서 작성자 또는 사용자 정의 문서 속성/필드, 사용자 및 액세스할 수 있는 문서 등의 정보가 포함된 문서 메타데이터와 액세스 제어 파일을 추가합니다. 각 메타데이터 파일에는 단일 문서에 대한 메타데이터가 들어 있습니다.
- 필드 매핑 - S3 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 해당 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 _document_body 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 S3 템플릿 스키마를 참조하세요.

자세히 알아보기

S3 데이터 Amazon Kendra 소스와의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- VPC를 지원하는 Amazon Kendra S3 커넥터를 사용하여 정확한 답변 검색

Amazon S3 데이터 소스 생성

다음 예제는 Amazon S3 데이터 소스 생성을 보여줍니다. 이 예제에서는 인덱스에서 데이터를 읽을 수 있는 권한이 있는 인덱스와 IAM 역할을 이미 만들었다고 가정합니다. IAM 역할에 대한 자세한 내용은 IAM 액세스 역할을 참조하십시오. 인덱스 생성에 대한 자세한 내용은 인덱스 생성을 참조하세요.

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

다음 Python 코드 스니펫은 Amazon S3 데이터 소스를 만듭니다. 전체 예제는 시작하기(AWS SDK for Python (Boto3)) 섹션을 참조하세요.

```
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
    )
```

데이터 소스를 생성하는 데 시간이 조금 걸릴 수 있습니다. API를 사용하여 진행 상황을 모니터링할 수 있습니다. DescribeDataSource 데이터 소스 상태가 ACTIVE가 되면 데이터 소스를 사용할 준비가 된 것입니다.

다음 예제는 데이터 소스의 상태를 가져오는 방법을 보여줍니다.

CLI

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

Python

다음 Python 코드 스니펫은 S3 데이터 소스에 대한 정보를 가져옵니다. 전체 예제는 시작하기 (AWS SDK for Python (Boto3)) 섹션을 참조하세요.

```
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
```

이 데이터 소스에는 일정이 없으므로 자동으로 실행되지 않습니다. 데이터 원본을 StartDataSourceSyncJob인덱싱하려면 를 호출하여 색인을 데이터 원본과 동기화합니다.

다음 예는 데이터 소스 동기화를 보여줍니다.

CLI

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

Python

다음 Python 코드 스니펫은 Amazon S3 데이터 소스를 동기화합니다. 전체 예제는 시작하기(AWS SDK for Python (Boto3)) 섹션을 참조하세요.

```
print("Synchronize the data source.")

    sync_response = kendra.start_data_source_sync_job(
        Id = "data-source-id",
        IndexId = "index-id"
    )
```

Amazon S3 문서 메타데이터

메타데이터 파일을 사용하여 문서에 대한 추가 정보인 메타데이터를 Amazon S3 버킷의 문서에 추가할 수 있습니다. 각 메타데이터 파일은 인덱싱된 문서와 연결됩니다.

메타데이터 파일은 인덱싱된 파일과 동일한 버킷에 저장해야 합니다. 콘솔이나 데이터 소스를 만들 때 `DocumentsMetadataConfiguration` 파라미터의 `S3Prefix` 필드를 사용하여 버킷 내 Amazon S3 메타데이터 파일 위치를 지정할 수 있습니다. Amazon S3 접두사를 지정하지 않는 경우 메타데이터 파일은 인덱싱된 문서와 동일한 위치에 저장되어야 합니다.

메타데이터 파일의 Amazon S3 접두사를 지정하면 해당 파일은 색인화된 문서와 평행한 디렉토리 구조를 갖게 됩니다. Amazon Kendra 지정된 디렉터리에서만 메타데이터를 찾습니다. 메타데이터를 읽을 수 없는 경우 디렉터리 위치가 메타데이터의 위치와 일치하는지 확인하세요.

다음 예시에서는 인덱싱된 문서 위치가 메타데이터 파일 위치에 매핑되는 방식을 보여줍니다. 참고로 문서 Amazon S3 키는 메타데이터의 접두사에 추가된 다음 Amazon S3 접미사를 사용하여 메타데이터 파일의 `.metadata.json` 경로를 구성합니다. Amazon S3 메타데이터의 Amazon S3 접두사 및 `.metadata.json` 접미사를 포함한 결합된 Amazon S3 키는 총 1024자를 넘지 않아야 합니다. Amazon S3 키를 접두사 및 접미사와 결합할 때 추가 문자를 고려하려면 키를 1000자 미만으로 유지하는 것이 좋습니다.

```
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
```

문서 메타데이터는 JSON 파일에 정의되어 있습니다. 파일은 BOM 마커가 없는 UTF-8 텍스트 파일이어야 합니다. JSON 파일의 이름은 `<document>.<extension>.metadata.json`이어야 합니다. 이 예제에서 “document”는 메타데이터가 적용되는 문서의 이름이고 “extension”은 문서의 파일 확장자입니다. 문서 ID는 `<document>.<extension>.metadata.json`에서 고유해야 합니다.

JSON 파일의 내용은 이 템플릿을 따릅니다. 모든 속성/필드는 선택 사항이므로 모든 속성을 포함할 필요는 없습니다. 포함하려는 각 속성에 값을 입력해야 합니다. 값은 비워 둘 수 없습니다. 를 지정하지 않으면 검색 Amazon Kendra 결과에서 반환되는 링크가 문서가 `_source_uri` 들어 있는 Amazon S3 버킷을 가리킵니다. `DocumentId`필드에 `s3_document_id` 매핑되며 S3에 있는 문서의 절대 경로입니다.

```
{
    "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."
}
```

_created_at 및 _last_updated_at 메타데이터 필드는 ISO 8601로 인코딩된 날짜입니다. 예를 들어, 2012-03-25T12:30:10+01:00은 중부 유럽 시간대로 2012년 3월 25일 오후 12시 30분 10초에 대한 ISO 8601 날짜/시간 형식입니다.

쿼리를 필터링하거나 쿼리 응답을 그룹화하는 데 사용하는 문서에 대한 추가 정보를 Attributes 필드에 추가할 수 있습니다. 자세한 정보는 사용자 지정 문서 필드 만들기을 참조하세요.

AccessControlList 필드를 사용하여 쿼리의 응답을 필터링할 수 있습니다. 이렇게 하면 특정 사용자와 그룹만 문서에 액세스할 수 있습니다. 자세한 정보는 사용자 컨텍스트 필터링을 참조하세요.

Amazon S3 데이터 소스에 대한 액세스 제어

구성 파일을 사용하여 Amazon S3 데이터 소스의 문서에 대한 액세스를 제어할 수 있습니다. 콘솔에서 파일을 지정하거나 CreateDataSource또는 UpdateDataSourceAPI를 호출할 때 AccessControlListConfiguration 매개변수로 지정합니다.

구성 파일에는 S3 접두사를 식별하고 접두사에 대한 액세스 설정을 나열하는 JSON 구조가 포함되어 있습니다. 접두사는 경로일 수도 있고 개별 파일일 수도 있습니다. 접두사가 경로인 경우 액세스 설정은 해당 경로에 있는 모든 파일에 적용됩니다. JSON 구성 파일에는 최대 S3 접두사 수와 기본 최대 파일 크기가 있습니다. 자세한 내용은 에 대한 할당량 Amazon Kendra단원을 참조하세요.

액세스 설정에서 사용자와 그룹을 모두 지정할 수 있습니다. 인덱스를 쿼리할 때 사용자 및 그룹 정보를 지정합니다. 자세한 정보는 사용 속성으로 필터링을 참조하세요.

구성 파일의 JSON 구조는 다음 형식이어야 합니다.

```
[
```

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

데이터 Amazon VPC 소스와 Amazon S3 함께 사용

이 주제에서는 Amazon VPC를 통해 Amazon S3 커넥터를 사용하여 Amazon S3 버킷에 연결하는 방법을 보여주는 step-by-step 예제를 제공합니다. 이 예제에서는 기존 S3 버킷으로 시작한다고 가정합니다. 예제를 테스트하려면 S3 버킷에 몇 개의 문서만 업로드하는 것이 좋습니다.

를 통해 Amazon Kendra Amazon S3 버킷에 연결할 수 Amazon VPC있습니다. 이렇게 하려면 Amazon S3 데이터 원본 커넥터를 만들 때 Amazon VPC 서브넷과 Amazon VPC 보안 그룹을 지정해야 합니다.

> **Important**
>
> Amazon Kendra Amazon S3 커넥터가 Amazon S3 버킷에 액세스할 수 있도록 하려면 가상 사설 클라우드 (VPC) 에 Amazon S3 엔드포인트를 할당했는지 확인하십시오.

Amazon S3 버킷의 문서를 Amazon Kendra 동기화하려면 다음 단계를 완료해야 합니다. Amazon VPC

- 에 대한 Amazon S3 엔드포인트를 설정합니다 Amazon VPC. 엔드포인트 설정 방법에 대한 자세한 내용은 Amazon S3AWS PrivateLink 가이드의 게이트웨이 Amazon S3 엔드포인트를 참조하십시오.
- (선택 사항) 할당한 가상 사설 클라우드 (VPC) 에서 Amazon S3 버킷에 액세스할 수 있는지 확인하기 위해 버킷 정책을 확인했습니다. Amazon S3 Amazon Kendra자세한 내용은 Amazon S3 사용 설명서의 버킷 정책을 사용한 VPC 엔드포인트의 액세스 제어를 참조하십시오.

단계

- 1단계: 구성 Amazon VPC
- (선택 사항) 2단계: Amazon S3 버킷 정책 구성
- 3단계: 테스트 Amazon S3 데이터 소스 커넥터 생성

1단계: 구성 Amazon VPC

Amazon S3 게이트웨이 엔드포인트가 있는 프라이빗 서브넷과 나중에 사용할 Amazon Kendra 보안 그룹을 포함하는 VPC 네트워크를 생성합니다.

프라이빗 서브넷, S3 엔드포인트 및 보안 그룹으로 VPC를 구성하려면

1. 에 AWS Management Console 로그인하고 에서 Amazon VPC 콘솔을 엽니다. https://console.aws.amazon.com/vpc/
2. 다음과 같이 사용할 프라이빗 서브넷과 S3 엔드포인트가 있는 VPC를 생성합니다. Amazon Kendra

   탐색 창에서 내 VPC를 선택한 다음 VPC 생성을 선택합니다.

a. 생성할 리소스에서 VPC 등을 선택합니다.

b. 이름 태그의 경우 자동 생성을 활성화한 다음 입력하십시오. **kendra-s3-example**

c. IPv4/IPv6 CIDR 블록의 경우 기본값을 유지하십시오.

d. 가용 영역 (AZ) 수에서는 1을 선택합니다.

e. 사용자 지정 AZ를 선택한 다음 첫 번째 가용 영역 목록에서 가용 영역을 선택합니다.

Amazon Kendra 특정 가용 영역 집합만 지원합니다.

f. 퍼블릭 서브넷 수에서는 숫자 0을 선택합니다.

g. 프라이빗 서브넷 수에서는 숫자 1을 선택합니다.

h. NAT 게이트웨이는 없음을 선택합니다.

i. VPC 엔드포인트의 경우 게이트웨이를 선택합니다.Amazon S3 .

j. 나머지 값은 기본 설정으로 유지합니다.

k. VPC 생성을 선택합니다.

VPC 생성 워크플로가 완료될 때까지 기다리십시오. 그런 다음 View VPC를 선택하여 방금 만든 VPC를 확인합니다.

이제 퍼블릭 인터넷에 액세스할 수 없는 프라이빗 서브넷이 있는 VPC 네트워크를 생성했습니다.

3. Amazon S3 엔드포인트의 VPC 엔드포인트 ID를 복사합니다.

a. 탐색 창에서 엔드포인트를 선택합니다.

b. 엔드포인트 목록에서 VPC와 함께 방금 `kendra-s3-example-vpce-s3` 생성한 Amazon S3 엔드포인트를 찾습니다.

c. VPC 엔드포인트 ID를 기록해 둡니다.

이제 서브넷을 통해 Amazon S3 버킷에 액세스할 수 있는 Amazon S3 게이트웨이 엔드포인트를 생성했습니다.

4. 사용할 보안 그룹을 생성합니다. Amazon Kendra

a. 탐색 창에서 보안 그룹을 선택한 다음 보안 그룹 생성을 선택합니다.

b. 보안 그룹 이름에 **s3-data-source-security-group**를 입력합니다.

c. 목록에서 VPC를 선택합니다. Amazon VPC

d. 인바운드 규칙과 아웃바운드 규칙을 기본값으로 두십시오.

e. 보안 그룹 생성을 선택합니다.

이제 VPC 보안 그룹을 생성했습니다.

커넥터 구성 프로세스 중에 Amazon Kendra Amazon S3 데이터 소스 커넥터에 생성한 서브넷 및 보안 그룹을 할당합니다.

(선택 사항) 2단계: Amazon S3 버킷 정책 구성

이 선택적 단계에서는 할당된 VPC에서만 Amazon S3 버킷에 액세스할 수 있도록 Amazon S3 버킷 정책을 구성하는 방법을 알아봅니다. Amazon Kendra

Amazon Kendra 는 IAM 역할을 사용하여 Amazon S3 버킷에 액세스하므로 Amazon S3 버킷 정책을 구성할 필요가 없습니다. 하지만 퍼블릭 인터넷에서의 액세스를 제한하는 기존 정책이 있는 Amazon S3 버킷을 사용하여 Amazon S3 커넥터를 구성하려는 경우 버킷 정책을 생성하는 것이 유용할 수 있습니다.

Amazon S3 버킷 정책을 구성하려면

1. https://console.aws.amazon.com/s3/에서 Amazon S3 콘솔을 엽니다.
2. 탐색 창에서 Buckets를 선택합니다.
3. 동기화할 Amazon S3 버킷의 이름을 선택합니다 Amazon Kendra.
4. 권한 탭을 선택하고 버킷 정책까지 아래로 스크롤한 다음 편집을 클릭합니다.
5. 생성한 VPC 엔드포인트에서만 액세스를 허용하도록 버킷 정책을 추가하거나 수정합니다.

   다음은 버킷 정책의 예입니다. *bucket-name*와 를 이전에 *vpce-id*적어둔 Amazon S3 버킷 이름 및 Amazon S3 엔드포인트 ID로 바꾸십시오.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Deny",
      "Principal": "*",
      "Action": "s3:*",
      "Resource": "arn:aws:s3:::bucket-name/*",
      "Condition": {
        "StringNotEquals": {
          "aws:SourceVpce": "vpce-id"
```

```
        }
      }
    }
  ]
}
```

6. 변경 사항 저장(Save changes)을 선택합니다.

이제 S3 버킷은 생성한 특정 VPC에서만 액세스할 수 있습니다.

3단계: 테스트 Amazon S3 데이터 소스 커넥터 생성

Amazon VPC 구성을 테스트하려면 Amazon S3 커넥터를 만드세요. 그런 다음 에 설명된 단계에 따라 생성한 VPC로 구성합니다. Amazon S3

Amazon VPC 구성 값의 경우 이 예제에서 생성한 값을 선택하십시오.

- Amazon VPC(VPC) — `kendra-s3-example-vpc`
- 서브넷 — `kendra-s3-example-subnet-private1-[availability zone]`
- 보안 그룹 — `s3-data-source-security-group`

커넥터 생성이 완료될 때까지 기다리십시오. Amazon S3 커넥터를 만든 후 지금 동기화를 선택하여 동기화를 시작합니다.

Amazon S3 버킷에 있는 문서 수에 따라 동기화를 완료하는 데 몇 분에서 몇 시간이 걸릴 수 있습니다. 예제를 테스트하려면 S3 버킷에 몇 개의 문서만 업로드하는 것이 좋습니다. 구성이 올바르면 결국 동기화 상태가 완료됨으로 표시될 것입니다.

오류가 발생하는 경우 Amazon VPC 연결 문제 해결을 참조하십시오.

Amazon Kendra 웹 크롤러

Amazon Kendra Web Crawler를 사용하여 웹 페이지를 크롤링하고 인덱싱할 수 있습니다.

보안 통신 프로토콜인 HTTPS(Hypertext Transfer Protocol Secure)를 사용하는 공개 웹 사이트 또는 회사 내부 웹사이트만 크롤링할 수 있습니다. 웹 사이트를 크롤링할 때 오류가 발생하면 웹 사이트가 크롤링되지 않도록 차단되었을 수 있습니다. 내부 웹 사이트를 크롤링하기 위해 웹 프록시를 설정할 수 있습니다. 이 웹 프록시는 공개용이어야 합니다. 인증을 사용하여 웹 사이트에 액세스하고 크롤링할 수도 있습니다.

인덱싱할 웹 사이트를 선택할 때 Amazon 이용 정책과 기타 모든 Amazon 약관을 준수해야 합니다. Amazon Kendra 웹 크롤러는 자신의 웹 페이지 또는 인덱싱할 권한이 있는 웹 페이지를 인덱싱할 때만 사용해야 한다는 점을 기억하십시오. Amazon Kendra Web Crawler가 웹 사이트를 인덱싱하지 못하게 하는 방법을 알아보려면 을 참조하십시오. Amazon Kendra 웹 크롤러용 *robots.txt* 파일 구성

> **Note**
>
> Amazon Kendra 웹 크롤러를 남용하여 소유하지 않은 웹 사이트 또는 웹 페이지를 공격적으로 크롤링하는 행위는 허용되는 사용으로 간주되지 않습니다.

Amazon Kendra 커넥터에는 두 가지 버전이 있습니다. web crawler 각 버전에 지원되는 기능은 다음과 같습니다.

Amazon Kendra 웹 크롤러 커넥터 v1.0/ API WebCrawlerConfiguration

- 웹 프록시
- 포함/제외 필터

Amazon Kendra 웹 크롤러 커넥터 v2.0/ API TemplateConfiguration

- 필드 매핑
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- 웹 프록시
- 웹 사이트를 위한 기본, NTLM/Kerberos, SAML 및 양식 인증
- Virtual Private Cloud(VPC)

> **Important**
>
> 웹 크롤러 v2.0 커넥터 생성은 에서 지원되지 않습니다. AWS CloudFormation지원이 필요한 경우 웹 크롤러 v1.0 커넥터를 사용하십시오. AWS CloudFormation

Amazon Kendra 웹 크롤러 데이터 원본 커넥터의 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

Amazon Kendra 웹 크롤러 커넥터 v1.0

웹 크롤러를 사용하여 Amazon Kendra 웹 페이지를 크롤링하고 인덱싱할 수 있습니다.

보안 통신 프로토콜인 HTTPS(Hypertext Transfer Protocol Secure)를 사용하는 웹사이트와 공개 웹 사이트만 크롤링할 수 있습니다. 웹 사이트를 크롤링할 때 오류가 발생하면 웹 사이트가 크롤링되지 않도록 차단되었을 수 있습니다. 내부 웹 사이트를 크롤링하기 위해 웹 프록시를 설정할 수 있습니다. 이 웹 프록시는 공개용이어야 합니다.

인덱싱할 웹 사이트를 선택할 때 Amazon 이용 정책과 기타 모든 Amazon 약관을 준수해야 합니다. Amazon Kendra 웹 크롤러는 자신의 웹 페이지 또는 인덱싱할 권한이 있는 웹 페이지를 인덱싱할 때만 사용해야 한다는 점을 기억하십시오. Amazon Kendra Web Crawler가 웹 사이트를 인덱싱하지 못하게 하는 방법을 알아보려면 을 참조하십시오. Amazon Kendra 웹 크롤러용 *robots.txt* 파일 구성

> **Note**
>
> Amazon Kendra 웹 크롤러를 남용하여 소유하지 않은 웹 사이트 또는 웹 페이지를 공격적으로 크롤링하는 행위는 허용되는 사용으로 간주되지 않습니다.

Amazon Kendra 웹 크롤러 데이터 원본 커넥터의 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

지원되는 기능

- 웹 프록시

- 포함/제외 필터

필수 조건

웹 사이트를 Amazon Kendra 인덱싱하는 데 사용하기 전에 먼저 웹 사이트 및 AWS 계정의 세부 정보를 확인하세요.

웹 사이트의 경우 다음이 있어야 합니다.

- 인덱싱하려는 웹 사이트의 시드 또는 사이트맵 URL을 복사했습니다.
- 기본 인증이 필요한 웹 사이트의 경우: 사용자 이름과 암호를 기록하고 웹 사이트의 호스트 이름과 포트 번호를 복사했습니다.
- 선택 사항: 웹 프록시를 사용하여 크롤링하려는 내부 웹 사이트에 연결하려는 경우 웹 사이트의 호스트 이름과 포트 번호를 복사했습니다. 이 웹 프록시는 공용이어야 합니다. Amazon Kendra 에서는 기본 인증으로 뒷받침되는 웹 프록시 서버 또는 인증 없이 연결할 수 있는 웹 프록시 서버에 연결할 수 있습니다.
- 인덱싱하려는 각 웹 페이지 문서가 고유한지, 동일한 인덱스에 사용할 다른 데이터 소스 전체를 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

AWS 계정에 다음이 있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- 인증이 필요한 웹 사이트 또는 인증과 함께 웹 프록시를 사용하는 경우 인증 자격 증명을 AWS Secrets Manager 비밀에 저장하고 API를 사용하는 경우 비밀의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 web crawler 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

web crawler데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 web crawler 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 web crawler 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 web crawler

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 웹 크롤러 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 웹 크롤러 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. 소스의 경우 사용 사례에 따라 소스 URL과 소스 사이트맵 중에서 선택하고 각 값을 입력합니다.

최대 10개의 소스 URL과 3개의 사이트맵을 추가할 수 있습니다.

> **Note**
>
> 사이트맵을 크롤링하려면 기본 또는 루트 URL이 사이트맵 페이지에 나열된 URL과 동일한지 확인하세요. 예를 들어 사이트맵 URL이 https://example.com/sitemap-page.html인 경우 이 사이트맵 페이지에 나열된 URL도 기본 URL “https://example.com/”을 사용해야 합니다.

b. (선택 사항) 웹 프록시의 경우 다음 정보를 입력합니다.

i. 호스트 이름 - 웹 프록시가 필요한 호스트 이름입니다.

ii. 포트 번호 - 호스트 URL 전송 프로토콜에서 사용하는 포트입니다. 포트 번호는 0에서 65535 사이의 숫자 값이어야 합니다.

iii. 웹 프록시 보안 인증의 경우 - 웹 프록시 연결에 인증이 필요한 경우 기존 보안 암호를 선택하거나 보안 인증 정보를 저장할 새 보안 암호를 생성합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

iv. AWS Secrets Manager Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름. 'AmazonKendra-WebCrawler-'라는 접두사가 보안 암호 이름에 자동으로 추가됩니다.

B. 사용자 이름 및 암호의 경우 - 웹 사이트의 기본 보안 인증 정보를 입력합니다.

C. 저장을 선택합니다.

c. (선택 사항) 인증 받은 호스트 - 인증 받은 호스트를 더 추가하려면 선택합니다.

d. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

e. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 크롤링 범위 - 크롤링하려는 웹 페이지의 종류를 선택합니다.

b. 크롤링 깊이 - 시드 URL에서 크롤링해야 하는 Amazon Kendra 수준의 수를 선택합니다.

c. 고급 크롤링 설정 및 추가 구성은 다음 정보를 입력합니다.

i. 최대 파일 크기 - 크롤링할 최대 웹 페이지 또는 첨부 파일 크기입니다. 최소 0.000001MB(1바이트). 최대 50MB.

ii. 페이지당 최대 링크 수 - 페이지당 크롤링된 최대 링크 수입니다. 링크는 표시되는 순서대로 크롤링됩니다. 페이지당 최소 1개 링크. 페이지당 최대 1000개의 링크.

iii. 최대 제한 - 1분간 호스트 이름당 크롤링되는 최대 URL 수입니다. 분당 호스트 이름당 최소 1개 URL. 분당 호스트 이름당 최대 300개 URL.

iv. 정규식 패턴 - 특정 URL을 포함하거나 제외하기 위한 정규 표현식 패턴을 추가합니다. 최대 100개의 패턴을 추가할 수 있습니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본과 Amazon Kendra 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra 연결하려면: web crawler

WebCrawlerConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- URL - SeedUrlConfiguration 및 SiteMapsConfiguration를 사용하여 크롤링할 웹 사이트의 시드나 시작 지점 URL 또는 웹 사이트의 사이트맵 URL을 지정합니다.

  > **Note**
  >
  > 사이트맵을 크롤링하려면 기본 또는 루트 URL이 사이트맵 페이지에 나열된 URL과 동일한지 확인하세요. 예를 들어 사이트맵 URL이 https://example.com/sitemap-page.html인 경우 이 사이트맵 페이지에 나열된 URL도 기본 URL “https://example.com/”을 사용해야 합니다.

- 보안 암호 Amazon 리소스 이름(ARN) - 웹사이트에서 기본 인증을 요구하는 경우 호스트 이름, 포트 번호, 사용자 이름 및 암호의 기본 보안 인증 정보를 저장하는 보안 암호를 제공합니다. AuthenticationConfiguration API를 사용하여 보안 암호 ARN을 제공합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  AWS Secrets Manager 보안 암호를 사용하여 웹 프록시 보안 인증 정보를 제공할 수도 있습니다. ProxyConfiguration API를 사용하여 웹 사이트 호스트 이름과 포트 번호를 제공하고, 선택적으로 웹 프록시 보안 인증 정보를 저장하는 보안 암호를 제공합니다.

- IAM 역할 - 역할에 `CreateDataSource` Secrets Manager 비밀번호에 액세스할 수 있는 권한을 제공하고 웹 크롤러 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. IAM Amazon Kendra자세한 내용은 웹 크롤러 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- 크롤링 모드 - 웹 사이트 호스트 이름만 크롤링할지, 하위 도메인이 포함된 호스트 이름을 크롤링할지, 아니면 웹 페이지가 연결된 다른 도메인도 크롤링할지 선택합니다.
- 시드 수준에서부터 크롤링할 '깊이' 또는 수준 수. 예를 들어 시드 URL 페이지는 깊이 1이고 이 페이지에서 크롤링되는 모든 하이퍼링크는 깊이 2입니다.
- 크롤링할 단일 웹 페이지의 최대 URL 수입니다.
- 크롤링할 웹 페이지의 최대 크기(MB 단위)입니다.

- 1분간 웹 사이트 호스트당 크롤링되는 최대 URL 수입니다.
- 내부 웹 사이트에 연결하고 크롤링하기 위한 웹 프록시 호스트 및 포트 번호입니다. 예를 들어, https://a.example.com/page1.html의 호스트 이름은 “a.example.com”이고 포트 번호는 HTTPS의 표준 포트인 443입니다. 웹 사이트 호스트에 연결하는 데 웹 프록시 보안 인증이 필요한 경우, 보안 인증을 저장하는 AWS Secrets Manager 보안 암호를 만들 수 있습니다.
- 사용자 인증이 필요한 웹 사이트에 액세스하고 크롤링하기 위한 인증 정보입니다.
- 사용자 지정 문서 보강 도구를 사용하여 HTML 메타 태그를 필드로 추출할 수 있습니다. 자세한 내용을 알아보려면 수집 프로세스 중 문서 메타데이터 사용자 지정을 참조하세요. HTML 메타태그 추출의 예는 CDE 예제를 참조하세요.
- 포함 및 제외 필터 - 특정 URL을 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

자세히 알아보기

데이터 Amazon Kendra 원본과의 통합에 대한 자세한 내용은 다음을 참조하십시오. web crawler

- Amazon Kendra의 웹 크롤러를 사용하여 지식 발견을 재구상하십시오.

Amazon Kendra 웹 크롤러 커넥터 v2.0

웹 크롤러를 사용하여 Amazon Kendra 웹 페이지를 크롤링하고 인덱싱할 수 있습니다.

보안 통신 프로토콜인 HTTPS(Hypertext Transfer Protocol Secure)를 사용하는 공개 웹 사이트 또는 회사 내부 웹사이트만 크롤링할 수 있습니다. 웹 사이트를 크롤링할 때 오류가 발생하면 웹 사이트가 크롤링되지 않도록 차단되었을 수 있습니다. 내부 웹 사이트를 크롤링하기 위해 웹 프록시를 설정할 수 있습니다. 이 웹 프록시는 공개용이어야 합니다. 인증을 사용하여 웹 사이트에 액세스하고 크롤링할 수도 있습니다.

Amazon Kendra 웹 크롤러 v2.0은 Selenium 웹 크롤러 패키지와 Chromium 드라이버를 사용합니다. Amazon Kendra 지속적 통합 (CI) 을 사용하여 Selenium 버전과 Chromium 드라이버를 자동으로 업데이트합니다.

인덱싱할 웹 사이트를 선택할 때 Amazon 이용 정책과 기타 모든 Amazon 약관을 준수해야 합니다. Amazon Kendra 웹 크롤러는 자신의 웹 페이지 또는 인덱싱 권한이 있는 웹 페이지를 인덱싱할 때만 사용해야 한다는 점을 기억하십시오. Amazon Kendra Web Crawler가 웹 사이트를 인덱싱하지 못하게 하는 방법을 알아보려면 을 참조하십시오. Amazon Kendra 웹 크롤러용 *robots.txt* 파일 구성 . Amazon Kendra 웹 크롤러를 남용하여 소유하지 않은 웹 사이트 또는 웹 페이지를 공격적으로 크롤링하는 행위는 허용되는 사용으로 간주되지 않습니다.

Amazon Kendra 웹 크롤러 데이터 원본 커넥터의 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

> **Note**
>
> 웹 크롤러 커넥터 v2.0은 암호화된 버킷에서 웹 사이트 목록을 크롤링하는 것을 지원하지 않습니다. AWS KMS Amazon S3 관리 키를 사용한 서버 측 암호화만 지원합니다. Amazon S3

> **Important**
>
> Web Crawler v2.0 커넥터 생성은 에서 지원되지 않습니다. AWS CloudFormation지원이 필요한 경우 웹 크롤러 v1.0 커넥터를 사용하십시오. AWS CloudFormation

주제

- 지원되는 기능
- 필수 조건
- 연결 지침

지원되는 기능

- 필드 매핑
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- 웹 프록시
- 웹 사이트를 위한 기본, NTLM/Kerberos, SAML 및 양식 인증
- Virtual Private Cloud(VPC)

필수 조건

웹 사이트를 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 웹 사이트 및 AWS 계정의 세부 정보를 확인하세요.

웹 사이트의 경우 다음이 있어야 합니다.

- 인덱싱하려는 웹 사이트의 시드 또는 사이트맵 URL을 복사했습니다. URL을 텍스트 파일에 저장하고 이를 Amazon S3 버킷에 업로드할 수 있습니다. 텍스트 파일의 각 URL은 별도의 줄에 형식을 지정해야 합니다. 사이트맵을 Amazon S3 버킷에 저장하려면 사이트맵 XML을 복사하여 XML 파일에 저장해야 합니다. 여러 개의 사이트맵 XML 파일을 ZIP 파일로 묶을 수도 있습니다.

  > **Note**
  >
  > (온프레미스/서버) AWS Secrets Manager 는 포함된 엔드포인트 정보가 데이터 소스 구성 세부 정보에 지정된 엔드포인트 정보와 동일한지 Amazon Kendra 확인합니다. 이렇게 하면 사용자가 작업을 수행할 권한이 없지만 구성된 보안 암호에 액세스하여 작업을 수행하는 데 Amazon Kendra 를 프록시로 사용하는 보안 문제인 혼동된 대리자 문제를 방지하는 데 도움이 됩니다. 나중에 엔드포인트 정보를 변경하는 경우 새 보안 암호를 생성하여 이 정보를 동기화해야 합니다.

- 기본, NTLM 또는 Kerberos 인증이 필요한 웹 사이트의 경우:
  - 사용자 이름과 암호가 포함된 웹사이트 보안 인증 정보를 기록해 두었습니다.

    > **Note**
    >
    > Amazon Kendra Web Crawler v2.0은 암호 해싱을 포함하는 NTLM 인증 프로토콜과 암호 암호화를 포함하는 Kerberos 인증 프로토콜을 지원합니다.

- SAML 또는 로그인 양식 인증이 필요한 웹 사이트의 경우:
  - 사용자 이름과 암호가 포함된 웹사이트 보안 인증 정보를 기록해 두었습니다.
  - 사용자 이름 필드(SAML을 사용하는 경우 사용자 이름 버튼 포함), 암호 필드 및 버튼의 XPaths(XML Path Language)를 복사하고 로그인 페이지 URL을 복사했습니다. 웹 브라우저의 개발자 도구를 사용하여 요소의 XPaths를 찾을 수 있습니다. XPaths는 일반적으로 다음 형식을 따릅니다. `//tagname[@Attribute='Value']`

> **Note**
>
> Amazon Kendra Web Crawler v2.0은 헤드리스 Chrome 브라우저와 양식의 정보를 사용하여 OAuth 2.0으로 보호된 URL을 통해 액세스를 인증하고 권한을 부여합니다.

- 선택 사항: 웹 프록시를 사용하여 크롤링하려는 내부 웹 사이트에 연결하려는 경우 웹 프록시 서버의 호스트 이름과 포트 번호를 복사했습니다. 웹 프록시는 공개되어야 합니다. Amazon Kendra 기본 인증으로 뒷받침되는 웹 프록시 서버 또는 인증 없이 연결할 수 있는 웹 프록시 서버에 연결할 수 있습니다.
- 선택 사항: VPC를 사용하여 크롤링하려는 내부 웹 사이트에 연결하려는 경우 Virtual Private Cloud(VPC) 서브넷 ID를 복사했습니다. 자세한 내용은 구성을 참조하십시오 Amazon VPC.
- 인덱싱하려는 각 웹 페이지 문서가 고유한지, 동일한 인덱스에 사용할 다른 데이터 소스 전체를 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

AWS 계정에 다음이 있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 생성하고 API를 사용하는 경우 IAM 역할의 Amazon 리소스 이름을 기록했습니다.

> **Note**
>
> 인증 유형 및 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 보안 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- 인증이 필요한 웹 사이트 또는 인증과 함께 웹 프록시를 사용하는 경우 인증 자격 증명을 AWS Secrets Manager 비밀에 저장하고 API를 사용하는 경우 비밀의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 web crawler 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

web crawler데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 web crawler 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 web crawler 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 web crawler

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 웹 크롤러 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 웹 크롤러 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. 소스— 소스 URL, 소스 사이트맵, 소스 URL 파일, 소스 사이트맵 파일 중에서 선택합니다. 최대 100개의 시드 URL 목록이 포함된 텍스트 파일을 사용하기로 선택한 경우 파일이 저장되는 Amazon S3 버킷의 경로를 지정합니다. 사이트맵 XML 파일을 사용하기로 선택한 경우 파일이 저장되는 Amazon S3 버킷의 경로를 지정합니다. 여러 개의 사이트맵 XML 파일을 ZIP 파일로 묶을 수도 있습니다. 그렇지 않으면 최대 10개의 시드 또는 시작점 URL과 최대 3개의 사이트맵 URL을 수동으로 입력할 수 있습니다.

> **Note**
>
> 사이트맵을 크롤링하려면 기본 또는 루트 URL이 사이트맵 페이지에 나열된 URL과 동일한지 확인하세요. 예를 들어 사이트맵 URL이 https://example.com/sitemap-page.html인 경우 이 사이트맵 페이지에 나열된 URL도 기본 URL "https://example.com/"을 사용해야 합니다.

웹 사이트에 액세스하는 데 인증이 필요한 웹 사이트의 경우 기본, NTLM/Kerberos, SAML 또는 양식 인증을 선택할 수 있습니다. 그렇지 않으면 인증 없음 옵션을 선택하세요.

> **Note**
>
> 나중에 데이터 소스를 편집하여 사이트맵에 대한 인증을 통해 시드 URL을 변경하려면 새 데이터 소스를 만들어야 합니다. Amazon Kendra 는 인증을 위해 Secrets Manager 보안 암호의 시드 URL 엔드포인트 정보를 사용하여 데이터 소스를 구성하므로 사이트맵으로 변경할 때 데이터 소스를 재구성할 수 없습니다.

- AWS Secrets Manager 비밀 - 웹 사이트에 액세스하는 데 동일한 인증이 필요한 웹 사이트인 경우 기존 암호를 선택하거나 새 암호를 만들어 웹 사이트 자격 Secrets Manager 증명을 저장하십시오. 새 암호를 만들기로 선택하면 AWS Secrets Manager 비밀 창이 열립니다.

  기본 또는 NTML/Kerberos 인증을 선택한 경우 보안 암호와 사용자 이름 및 암호를 입력합니다. NTLM 인증 프로토콜에는 암호 해싱이 포함되고 Kerberos 인증 프로토콜에는 암호 암호화가 포함됩니다.

  SAML 또는 양식 인증을 선택한 경우 보안 암호와 사용자 이름 및 암호를 입력합니다. 사용자 이름 필드에는 XPath를 사용하고, SAML을 사용하는 경우 사용자

이름 버튼에는 XPath를 사용합니다. 암호 필드와 버튼, 로그인 페이지 URL에는 XPaths를 사용하세요. 웹 브라우저의 개발자 도구를 사용하여 요소의 XPaths(XML 경로 언어)를 찾을 수 있습니다. XPaths는 일반적으로 다음 형식을 따릅니다. `//tagname[@Attribute='Value']`

b. (선택 사항) 웹 프록시 - 내부 웹 사이트에 연결하는 데 사용할 프록시 서버의 호스트 이름과 포트 번호를 입력합니다. 예를 들어, https://a.example.com/page1.html의 호스트 이름은 “a.example.com”이고 포트 번호는 HTTPS의 표준 포트인 443입니다. 웹 사이트 호스트에 연결하는 데 웹 프록시 자격 증명이 필요한 경우 자격 증명을 저장하는 웹 프록시 자격 증명을 만들 수 있습니다. AWS Secrets Manager

c. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

d. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

e. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위 - 도메인, 파일 크기 및 링크를 비롯한 웹 페이지 크롤링에 대한 제한을 설정하고 정규식 패턴을 사용하여 URL을 필터링합니다.

i. (선택 사항) 도메인 크롤링 범위 - 웹 사이트 도메인만 크롤링할지, 하위 도메인이 있는 도메인만 크롤링할지, 아니면 웹 페이지가 연결된 다른 도메인도 크롤링할지 선택합니다. 기본적으로 크롤링하려는 웹 사이트의 Amazon Kendra 도메인만 크롤링합니다.

ii. (선택 사항) 추가 구성 - 다음 설정을 구성합니다.

- 크롤링 깊이 - 시드 수준에서부터 크롤링할 '깊이' 또는 수준 수. 예를 들어 시드 URL 페이지는 깊이 1이고 이 페이지에서 크롤링되는 모든 하이퍼링크는 깊이 2입니다.
- 최대 파일 크기 - 크롤링할 최대 웹 페이지 또는 첨부 파일 크기(MB)입니다.
- 페이지당 최대 링크 - 크롤링할 단일 웹 페이지의 최대 URL 수입니다.

- 최대 크롤링 속도 제한 - 1분간 웹사이트 호스트당 크롤링되는 최대 URL 수입니다.
- 파일 - 웹 페이지가 링크된 파일을 크롤링하려면 선택합니다.
- URL 크롤링 및 인덱싱 - 특정 URL을 크롤링하고 해당 URL 웹 페이지의 하이퍼링크를 인덱싱하는 것을 포함하거나 제외하는 정규 표현식 패턴 목록을 추가합니다.

b. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

c. 동기화 실행 일정 - 빈도에서 Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

d. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. Amazon Kendra 생성된 웹 페이지 및 파일의 기본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra 연결하려면 web crawler

TemplateConfiguration API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 `WEBCRAWLERV2` 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 `TEMPLATE` 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource

- URL - 크롤링할 웹 사이트의 시드나 시작 지점 URL 또는 웹 사이트의 사이트맵 URL을 지정합니다. 시드 URL 목록을 저장하는 Amazon S3 버킷의 경로를 지정할 수 있습니다. 시드 URL의 텍스트 파일의 각 URL은 별도의 줄에 형식을 지정해야 합니다. 사이트맵 XML 파일을 저장하는 Amazon S3 버킷의 경로를 지정할 수도 있습니다. 여러 사이트맵 파일을 ZIP 파일로 묶고 Amazon S3 버킷에 ZIP 파일을 저장할 수 있습니다.

  > **Note**
  >
  > 사이트맵을 크롤링하려면 기본 또는 루트 URL이 사이트맵 페이지에 나열된 URL과 동일한지 확인하세요. 예를 들어 사이트맵 URL이 https://example.com/sitemap-page.html인 경우 이 사이트맵 페이지에 나열된 URL도 기본 URL “https://example.com/”을 사용해야 합니다.

- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - `FORCED_FULL_CRAWL`모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - `FULL_CRAWL`데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 인증 - 웹 사이트에 동일한 인증이 필요한 경우, `BasicAuth`, `NTLM_Kerberos`, `SAML` 또는 `Form` 인증을 지정하세요. 웹 사이트에 인증이 필요하지 않은 경우 `NoAuthentication`을 지정하세요.
- 보안 암호 Amazon 리소스 이름(ARN) - 웹사이트에서 기본, NTLM 또는 Kerberos 인증을 사용하려는 경우 사용자 이름 및 암호의 보안 인증 정보를 저장하는 보안 암호를 제공합니다. AWS Secrets Manager 보안 암호의 Amazon 리소스 이름(ARN)을 제공합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

  ```
  {
      "seedUrlsHash": "Hash representation of all seed URLs",
      "userName": "user name",
      "password": "password"
  }
  ```

  웹 사이트에 SAML 인증이 필요한 경우 보안 암호는 다음 키가 있는 JSON 구조에 저장됩니다.

```
{
    "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"
}
```

웹 사이트에 양식 인증이 필요한 경우 보안 암호는 다음 키가 있는 JSON 구조에 저장됩니다.

```
{
    "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"
}
```

웹 브라우저의 개발자 도구를 사용하여 요소의 XPaths(XML 경로 언어)를 찾을 수 있습니다. XPaths는 일반적으로 다음 형식을 따릅니다. `//tagname[@Attribute='Value']`

AWS Secrets Manager 보안 암호를 사용하여 웹 프록시 보안 인증 정보를 제공할 수도 있습니다.

- IAM 역할 - `CreateDataSource` Secrets Manager 비밀번호에 액세스할 수 있는 권한을 IAM 역할을 제공하고 웹 크롤러 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 웹 크롤러 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.

- 도메인 범위 - 하위 도메인만 포함하여 웹 사이트 도메인을 크롤링할지, 아니면 웹 페이지가 연결된 다른 도메인도 크롤링할지 선택합니다. 기본적으로 크롤링하려는 웹 사이트의 Amazon Kendra 도메인만 크롤링합니다.
- 시드 수준에서부터 크롤링할 '깊이' 또는 수준 수. 예를 들어 시드 URL 페이지는 깊이 1이고 이 페이지에서 크롤링되는 모든 하이퍼링크는 깊이 2입니다.
- 크롤링할 단일 웹 페이지의 최대 URL 수입니다.
- 크롤링할 웹 페이지나 첨부 파일의 최대 크기(MB 단위)입니다.
- 1분간 웹 사이트 호스트당 크롤링되는 최대 URL 수입니다.
- 내부 웹 사이트에 연결하고 크롤링하기 위한 웹 프록시 호스트 및 포트 번호입니다. 예를 들어, https://a.example.com/page1.html의 호스트 이름은 “a.example.com”이고 포트 번호는 HTTPS의 표준 포트인 443입니다. 웹 사이트 호스트에 연결하는 데 웹 프록시 보안 인증이 필요한 경우, 보안 인증을 저장하는 AWS Secrets Manager 보안 암호를 만들 수 있습니다.
- 포함 및 제외 필터 - 특정 URL을 크롤링하고 해당 URL 웹 페이지의 하이퍼링크를 인덱싱하는 작업을 포함할지 제외할지 여부를 지정합니다.

  > **Note**
  >
  > 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 필드 매핑 - 웹 페이지 및 웹 페이지 파일의 필드를 인덱스 필드에 매핑하도록 선택합니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

구성해야 할 기타 중요한 JSON 키 목록은 Amazon Kendra 웹 크롤러 템플릿 스키마를 참조하세요.

Amazon Kendra 웹 크롤러용 `robots.txt` 파일 구성

Amazon Kendra AWS 고객이 원하는 문서를 인덱싱하고 검색하는 데 사용하는 지능형 검색 서비스입니다. 웹에서 문서를 인덱싱하기 위해 고객은 인덱싱해야 하는 URL과 기타 운영 매개 변수를 나타내는 Amazon Kendra Web Crawler를 사용할 수 있습니다. Amazon Kendra 고객은 특정 웹 사이트를 인덱싱하기 전에 승인을 받아야 합니다.

Amazon Kendra 웹 크롤러는 및 와 같은 표준 robots.txt 지침을 준수합니다. `Allow` `Disallow` 웹 사이트 `robots.txt` 파일을 수정하여 웹 크롤러가 웹 사이트를 크롤링하는 방식을 Amazon Kendra 제어할 수 있습니다.

웹 크롤러가 Amazon Kendra 웹 사이트에 액세스하는 방법 구성

Amazon Kendra 웹 크롤러가 및 지시문을 사용하여 웹 사이트를 인덱싱하는 방식을 제어할 수 있습니다. `Allow` `Disallow` 또한 인덱싱되는 웹 페이지와 크롤링되지 않는 웹 페이지를 제어할 수 있습니다.

Amazon Kendra Web Crawler가 허용되지 않은 웹 페이지를 제외한 모든 웹 페이지를 크롤링하도록 허용하려면 다음 지침을 사용하십시오.

```
User-agent: amazon-kendra    # Amazon Kendra Web Crawler
Disallow: /credential-pages/ # disallow access to specific pages
```

Amazon Kendra Web Crawler가 특정 웹 페이지만 크롤링하도록 허용하려면 다음 지침을 사용하십시오.

```
User-agent: amazon-kendra    # Amazon Kendra Web Crawler
Allow: /pages/ # allow access to specific pages
```

Amazon Kendra Web Crawler가 모든 웹 사이트 콘텐츠를 크롤링하도록 허용하고 다른 로봇의 크롤링은 허용하지 않으려면 다음 지침을 사용하십시오.

```
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
```

웹 크롤러가 웹 사이트를 크롤링하지 못하도록 중지 Amazon Kendra

지침을 사용하여 Amazon Kendra 웹 크롤러가 웹 사이트의 색인을 생성하지 못하게 할 수 있습니다. `Disallow` 또한 인덱싱되는 웹 페이지와 크롤링되지 않는 웹 페이지를 제어할 수 있습니다.

Amazon Kendra Web Crawler가 웹 사이트를 크롤링하지 못하게 하려면 다음 지침을 사용하십시오.

```
User-agent: amazon-kendra # Amazon Kendra Web Crawler
Disallow: / # disallow access to any pages
```

Amazon Kendra 또한 웹 크롤러는 HTML 페이지의 메타 태그에 있는 noindex 로봇과 nofollow 지시문을 지원합니다. 이 지시문은 웹 크롤러가 웹 페이지를 인덱싱하지 못하게 하고 웹 페이지의 모든 링크를 따라가지 않도록 합니다. 문서 섹션에 메타 태그를 넣어 로봇 규칙의 규칙을 지정합니다.

예를 들어, 아래 웹 페이지에는 지시문 로봇 noindex 및 nofollow가 포함됩니다.

```
<html>
<head>
    <meta name="robots" content="noindex, nofollow"/>
    ...
</head>
<body>...</body>
</html>
```

Amazon Kendra 웹 크롤러와 관련하여 질문이나 우려 사항이 있는 경우 지원 팀에 문의할 수 있습니다.AWS

Amazon WorkDocs

Amazon WorkDocs 콘텐츠 생성, 편집, 저장 및 공유를 위한 안전한 콘텐츠 협업 서비스입니다. 를 사용하여 Amazon WorkDocs 데이터 원본을 Amazon Kendra 인덱싱할 수 있습니다.

Amazon Kendra 콘솔과 WorkDocsConfigurationAPI를 Amazon Kendra 사용하여 Amazon WorkDocs 데이터 소스에 연결할 수 있습니다.

Amazon WorkDocs 오레곤, 노스버지니아, 시드니, 싱가포르, 아일랜드 지역에서 사용할 수 있습니다.

Amazon Kendra WorkDocs 데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra WorkDocs 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 변경 로그

필수 조건

WorkDocs 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 WorkDocs 및 AWS 계정에서 다음과 같이 변경하십시오.

WorkDocs에서 다음이 있는지 확인하세요.

- Amazon WorkDocs 저장소의 Amazon WorkDocs 디렉터리 ID (조직 ID) 를 기록해 두었습니다.
- 동일한 색인에 사용하려는 다른 데이터 소스에서 각 문서가 고유한지 확인했습니다. WorkDocs 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

AWS 계정에 다음이 있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

기존 IAM 역할이 없는 경우 WorkDocs 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할을 만들 수 있습니다. Amazon Kendra API를 사용하는 경우 기존 IAM 역할의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

WorkDocs 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 WorkDocs 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 WorkDocs 을 Amazon Kendra참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Amazon WorkDocs

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.

2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 WorkDocs 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 WorkDocs 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. Amazon WorkDocs 사이트별 조직 ID - 색인을 생성하려는 Amazon WorkDocs 사이트의 ID를 선택합니다. 사이트가 이미 생성되어 있어야 합니다.
   b. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

   > **Note**
   >
   > IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

   c. 다음을 선택합니다.
7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

   a. 문서 설명 크롤링 - 크롤링하려는 Amazon WorkDocs 개체 또는 콘텐츠 유형.

b. 변경 로그 사용 - 모든 파일을 동기화하는 대신 새 콘텐츠나 수정된 콘텐츠만 사용하여 색인을 업데이트하려면 이 옵션을 선택합니다.

c. 정규식 패턴 - 특정 페이지 및 자산을 포함하거나 제외하기 위한 정규 표현식 패턴입니다.

d. 동기화 시 빈도 실행 일정 —데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra 연결하려면 Amazon WorkDocs

WorkDocsConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- Amazon WorkDocs 디렉터리 ID - 디렉터리의 조직 ID를 지정합니다. Amazon WorkDocs Active Directory로 이동한 다음 디렉터리로 이동하여 AWS Directory Service에서 조직 ID를 찾을 수 있습니다.
- IAM 역할 - WorkDocs 디렉터리에 액세스할 수 있는 권한을 IAM 역할에 제공하고 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. `CreateDataSource` WorkDocs Amazon Kendra자세한 내용은 데이터 소스의 IAM 역할을 참조하십시오. WorkDocs

다음 선택적 기능도 추가할 수 있습니다.

- 변경 로그 - 인덱스에서 문서를 Amazon Kendra 업데이트해야 하는지 여부를 결정하기 위해 WorkDocs 데이터 소스 변경 로그 메커니즘을 사용해야 하는지 여부.

> **Note**
>
> Amazon Kendra 가 모든 문서를 스캔하지 않도록 하려면 변경 로그를 사용하세요. 변경 로그가 크면 WorkDocs 데이터 원본의 문서를 스캔하는 데 걸리는 시간이 변경 로그를 처리하는 시간보다 Amazon Kendra 적을 수 있습니다. WorkDocs데이터 원본을 색인과 처음으로 동기화하는 경우 모든 문서가 스캔됩니다.

- 포함 및 제외 필터 - 특정 문서와 문서 설명을 포함할지 또는 제외할지 여부를 지정합니다. 각 설명은 별도의 문서로 인덱싱됩니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 — WorkDocs 데이터 원본 필드를 인덱스 필드에 매핑하도록 선택합니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

자세히 알아보기

WorkDocs 데이터 Amazon Kendra 소스와의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- Amazon Kendra Amazon WorkDocs 커넥터로 시작하기

Box

Box는 파일 호스팅 기능을 제공하는 클라우드 스토리지 서비스입니다. 댓글, 작업, 웹 링크 등 Box 콘텐츠의 콘텐츠를 인덱싱하는 Amazon Kendra 데 사용할 수 있습니다.

Amazon Kendra 콘솔과 BoxConfigurationAPI를 Amazon Kendra 사용하여 Box 데이터 소스에 연결할 수 있습니다.

Amazon Kendra Box 데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Box 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 변경 로그, 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Box 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하기 전에 먼저 Box와 AWS 계정에서 다음과 같이 변경하세요.

Box에서 다음 사항을 갖추었는지 확인하세요.

- A Box Enterprise 또는 Box Enterprise Plus 계정.
- Box 개발자 콘솔에서 JSON 웹 토큰 (JWT) 을 사용한 서버 측 인증을 사용하여 Box 맞춤형 앱을 구성했습니다. 자세한 내용은 사용자 지정 앱 만들기에 대한 Box 설명서 및 JWT 인증 구성에 대한 Box 설명서를 참조하세요.

- 앱 액세스 수준을 App + Enterprise Access로 설정하고 as-user 헤더를 사용하여 API 호출을 할 수 있도록 허용했습니다.
- 관리자를 사용하여 Box 앱에 다음 애플리케이션 범위를 추가했습니다.
  - Box에 저장된 모든 파일 및 폴더를 작성합니다.
  - 사용자 관리
  - 그룹 관리
  - 엔터프라이즈 속성 관리
- 인증 자격 증명으로 사용할 클라이언트 ID, 클라이언트 암호, 공개 키 ID, 개인 키 ID, 암호 문구 및 엔터프라이즈 ID를 포함하는 공개/개인 키 쌍을 구성했습니다. 자세한 내용은 공개 및 개인 키 쌍을 참조하십시오.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- Box 개발자 콘솔 설정 또는 Box 앱에서 Box 엔터프라이즈 ID를 복사했습니다. 예: *801234567*.
- Box 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

귀하의 AWS 계정란에 다음이 있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Box 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Box 데이터 소스를 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Box 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Box 데이터 소스의 필수 세부 정보를 제공해야 합니다. 아직 Box를 구성하지 않은 경우 을 Amazon Kendra참조하십시오필수 조건.

Console

Amazon Kendra Box에 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Box 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Box 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 색인을 기준으로 문서를 필터링할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. Box 엔터프라이즈 ID - Box 엔터프라이즈 ID를 입력합니다. 예: *801234567*.

b. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

c. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 Secrets Manager 암호를 만들어 Box 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

i. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Box-'가 자동으로 추가됩니다.

ii. 클라이언트 ID, 클라이언트 암호, 공개 키 ID, 개인 키 ID 및 암호문의 경우 - Box에서 구성한 공개/개인 키의 값을 입력합니다.

iii. 비밀번호를 추가하고 저장합니다.

d. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

e. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

f. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

g. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

   a. Box 폴더 ID - 크롤링하려는 특정 Box 폴더 ID를 입력합니다. 그렇지 않으면 모든 폴더의 콘텐츠가 크롤링됩니다.

   b. Box 파일 - 웹 링크, 댓글, 작업을 크롤링할지 여부를 선택합니다.

   c. 추가 구성의 경우 - 정규 표현식 패턴을 추가하여 특정 콘텐츠를 포함하거나 제외합니다.

   d. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

      - 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
      - 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
      - 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

   e. 빈도에 대한 동기화 실행 일정 —데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

   f. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Box에 Amazon Kendra 연결하려면

BoxConfigurationAPI를 사용하여 다음을 지정해야 합니다.

Box Enterprise ID - Box Enterprise ID를 입력합니다. 엔터프라이즈 ID는 Box 개발자 콘솔 설정에서 또는 Box에서 앱을 구성할 때 찾을 수 있습니다.

- 비밀 Amazon 리소스 이름 (ARN) - Box 계정의 인증 자격 증명이 포함된 Secrets Manager 비밀의 Amazon 리소스 이름 (ARN) 을 입력합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "clientID": "client-id",
    "clientSecret": "client-secret",
    "publicKeyID": "public-key-id",
    "privateKey": "private-key",
    "passphrase": "pass-phrase"
}
```

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 Box 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Box 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - 데이터 소스 구성의 일부로 `VpcConfiguration`를 지정합니다. VPC 사용을 위한 Amazon Kendra 구성을 참조하세요.
- 변경 로그 - 색인에서 문서를 Amazon Kendra 업데이트해야 하는지 여부를 결정하기 위해 Box 데이터 소스 변경 로그 메커니즘을 사용해야 하는지 여부.

> **Note**
>
> Amazon Kendra 가 모든 문서를 스캔하지 않도록 하려면 변경 로그를 사용하세요. 변경 로그가 크면 변경 로그를 처리하는 것보다 Box 데이터 원본의 문서를 스캔하는 데 걸리는 시간이 더 Amazon Kendra 적을 수 있습니다. Box 데이터 소스를 인덱스와 처음으로 동기화하는 경우 모든 문서가 스캔됩니다.

- 댓글, 작업, 웹 링크 - 이러한 유형의 콘텐츠를 크롤링할지 여부를 지정하세요.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 포함 및 제외 필터 - 특정 Box 파일 및 폴더를 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Box 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

자세히 알아보기

Box 데이터 Amazon Kendra 소스와의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- Amazon Kendra Box 커넥터 시작하기

Confluence

Confluence는 프로젝트 계획, 소프트웨어 개발 및 제품 관리를 공유, 저장 및 작업하도록 설계된 협업 작업 관리 도구입니다. Confluence 스페이스, 페이지 (중첩된 페이지 포함), 블로그, 색인된 페이지 및 블로그의 댓글과 첨부 파일을 인덱싱하는 데 사용할 Amazon Kendra 수 있습니다.

Amazon Kendra 컨플루언스 서버/데이터 센터와 컨플루언스 클라우드를 모두 지원합니다.

> **Note**
>
> 기본적으로 Confluence 아카이브와 개인 Amazon Kendra 공간을 인덱싱하지 않습니다. 데이터 소스를 만들 때 인덱싱하도록 선택할 수 있습니다. 스페이스를 인덱싱하지 Amazon Kendra 않으려면 Confluence에서 해당 스페이스를 비공개로 표시하세요.

Amazon Kendra 콘솔, TemplateConfigurationAPI 또는 API를 Amazon Kendra 사용하여 Confluence 데이터 소스에 연결할 수 있습니다. ConfluenceConfiguration

Amazon Kendra 컨플루언스 커넥터에는 두 가지 버전이 있습니다. 각 버전에 지원되는 기능은 다음과 같습니다.

컨플루언스 커넥터 V1.0/ API ConfluenceConfiguration

- 필드 매핑
- 사용자 액세스 제어

- 포함/제외 필터
- (Confluence 서버에만 해당) Virtual Private Cloud(VPC)

컨플루언스 커넥터 V2.0/ API TemplateConfiguration

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 패턴
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

> **Note**
>
> 컨플루언스 커넥터 ConfluenceConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. 컨플루언스 커넥터 V2.0/ API로 마이그레이션하거나 이를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra Confluence 데이터 소스 커넥터 문제 해결은 을 참조하십시오. 데이터 소스 문제 해결

주제

- Confluence 커넥터 V1.0
- Confluence 커넥터 V2.0

Confluence 커넥터 V1.0

Confluence는 프로젝트 계획, 소프트웨어 개발 및 제품 관리를 공유, 저장 및 작업하도록 설계된 협업 작업 관리 도구입니다. Confluence 스페이스, 페이지(중첩된 페이지 포함), 블로그, 인덱싱된 페이지 및 블로그에 설명과 첨부 파일을 인덱싱하는 데 Amazon Kendra 를 사용할 수 있습니다.

> **Note**
>
> 컨플루언스 커넥터 ConfluenceConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. 컨플루언스 커넥터 V2.0/ API로 마이그레이션하거나 이를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra Confluence 데이터 소스 커넥터 문제 해결은 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Confluence 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- (Confluence 서버에만 해당) Virtual Private Cloud(VPC)

필수 조건

Confluence 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Confluence 및 계정에서 다음과 같이 변경하십시오. AWS

Confluence에서 다음을 충족해야 합니다.

- 다음과 같은 방법으로 Confluence 인스턴스 내의 모든 콘텐츠를 볼 수 있는 Amazon Kendra 권한을 부여했습니다.
  - 그룹 구성원 Amazon Kendra `confluence-administrators` 만들기.
  - 모든 기존 스페이스, 블로그 및 페이지에 대한 사이트 관리자 권한 부여.
- Confluence 인스턴스의 URL을 복사했습니다.
- SSO(Single Sign-On) 사용자의 경우: Confluence 데이터 센터에서 Confluence 인증 방법을 구성할 때 사용자 이름 및 암호에 대해 로그인 시 표시 페이지를 활성화했습니다.
- Confluence 서버의 경우
  - Amazon Kendra에 연결할 Confluence 관리 계정 사용자 이름과 암호가 들어 있는 기본보안 인증 정보를 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

  - 선택 사항: Confluence 계정에서 Amazon Kendra에 연결할 개인용 액세스 토큰을 생성했습니다. 자세한 내용은 개인용 액세스 토큰 생성에 대한 Confluence 설명서를 참조하세요.
- Confluence 클라우드의 경우
  - Amazon Kendra에 연결할 Confluence 관리 계정 사용자 이름과 암호가 들어 있는 기본보안 인증 정보를 기록했습니다.
- Confluence 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

귀하의 AWS 계정계정에는 다음이 있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Confluence 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Confluence 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. Amazon Kendra API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Confluence 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 수 있도록 Amazon Kendra Confluence 자격 증명의 세부 정보를 제공해야 합니다. 아직 Confluence를 구성하지 않은 경우 참조하십시오. Amazon Kendra 필수 조건

Console

컨플루언스에 Amazon Kendra 연결하려면

1. AWS 관리 콘솔에 로그인하고 콘솔을 엽니다.Amazon Kendra
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 소스 추가 페이지에서 Confluence 커넥터 V1.0을 선택한 다음 데이터 소스 추가를 선택합니다.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. 컨플루언스 클라우드와 컨플루언스 서버 중에서 선택하세요.

b. Confluence 클라우드를 선택하는 경우 다음 정보를 입력합니다.

   i. Confluence URL - 사용자의 Confluence URL.

   ii. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Confluence Secrets Manager 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

   - AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

     I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Confluence-'가 자동으로 추가됩니다.

     II. 사용자 이름 및 암호의 경우 —Confluence 사용자 이름과 암호를 입력합니다.

     III. 인증 저장을 선택합니다.

c. Confluence 서버를 선택하는 경우 다음 정보를 입력합니다.

   i. Confluence URL - Confluence 사용자 이름 및 암호.

   ii. (선택 사항) 웹 프록시의 경우 다음 정보를 입력합니다.

   A. 호스트 이름 - Confluence 계정의 호스트 이름.

   B. 포트 번호 - 호스트 URL 전송 프로토콜에서 사용하는 포트 번호.

   iii. 인증의 경우 기본 인증 또는 (Confluence 서버만 해당) 개인용 액세스 토큰을 선택합니다.

   iv. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Confluence Secrets Manager 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

   - AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

     I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Confluence-'가 자동으로 추가됩니다.

     II. 사용자 이름 및 암호의 경우 - Confluence에서 구성한 인증 자격 증명 값을 입력합니다. 기본 인증을 사용하는 경우 Confluence 사용자 이름 (이메일 ID) 과 암호 (API 토큰) 를 사용하십시오. 개인용 액세스 토큰을 사용하는 경우 Confluence 계정에서 구성한 개인용 액세스 토큰의 세부 정보를 입력합니다.

     III. 비밀번호를 저장하고 추가하세요.

d. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

e. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 개인 스페이스 포함 및 보관된 스페이스 포함 - 이 데이터 소스에 포함할 선택적 스페이스 유형을 선택합니다.

b. 추가 구성의 경우 - 정규 표현식 패턴을 지정하여 특정 콘텐츠를 포함하거나 제외합니다. 최대 100개의 패턴을 추가할 수 있습니다.

c. 선택한 스페이스 내에서 첨부 파일을 크롤링하도록 선택할 수도 있습니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본과 Amazon Kendra 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 스페이스, 페이지, 블로그의 경우 - Amazon Kendra 생성된 기본 데이터 원본 필드 또는 추가 제안 필드 매핑 중에서 선택하여 인덱스 필드를 추가합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

컨플루언스에 연결하려면 Amazon Kendra

API를 사용하여 ConfluenceConfiguration다음을 지정해야 합니다.

- 컨플루언스 버전 - 사용 중인 Confluence 인스턴스의 버전을 CLOUD 또는 SERVER로 지정합니다.
- 비밀 Amazon 리소스 이름 (ARN) - Confluence 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다.

  Confluence 서버를 사용하는 경우 Confluence 사용자 이름과 암호 또는 개인 액세스 토큰을 인증 자격 증명으로 사용할 수 있습니다.

  Confluence 사용자 이름과 암호를 인증 자격 증명으로 사용하는 경우 다음 자격 증명을 JSON 구조로 시크릿에 저장합니다. Secrets Manager

  ```
  {
      "username": "user name",
      "password": "password"
  }
  ```

  개인용 액세스 토큰을 사용하여 Confluence 서버를 연결하는 Amazon Kendra경우 다음 자격 증명을 JSON 구조로 시크릿에 저장합니다. Secrets Manager

  ```
  {
      "patToken": "personal access token"
  }
  ```

  컨플루언스 클라우드를 사용하는 경우 컨플루언스 사용자 이름과 컨플루언스에 구성된 API 토큰을 비밀번호로 사용합니다. 다음 자격 증명을 JSON 구조로 시크릿에 저장합니다. Secrets Manager

  ```
  {
      "username": "user name",
      "password": "API token"
  }
  ```

- IAM 역할 - 역할에 CreateDataSource Secrets Manager 암호에 액세스할 수 있는 권한을 제공하고 Confluence 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 RoleArn 때 지정합니다. IAM Amazon Kendra자세한 내용은 Confluence 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- 웹 프록시 - 웹 프록시를 통해 Confluence URL 인스턴스에 연결할지 여부입니다. 이 옵션을 Confluence 서버에 사용할 수 있습니다.

- (Confluence 서버만 해당) Virtual Private Cloud(VPC) - `VpcConfiguration`을 데이터 소스 구성의 일부로 지정합니다. VPC 사용을 Amazon Kendra 위한 구성을 참조하십시오.
- 포함 및 제외 필터 - 특정 스페이스, 블로그 게시물, 페이지, 스페이스 및 연결을 포함하거나 제외하는 정규식 패턴을 지정합니다. 첨부 파일을 인덱싱하도록 선택하면 인덱싱된 페이지 및 블로그에 대한 첨부 파일만 인덱싱됩니다.

  > **Note**
  >
  > 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 필드 매핑 - Confluence 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

  > **Note**
  >
  > 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서의 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

자세히 알아보기

Confluence 데이터 Amazon Kendra 원본과의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- 컨플루언스 서버 커넥터 Amazon Kendra 구성

Confluence 커넥터 V2.0

Confluence는 프로젝트 계획, 소프트웨어 개발 및 제품 관리를 공유, 저장 및 작업하도록 설계된 협업 작업 관리 도구입니다. Confluence 스페이스, 페이지(중첩된 페이지 포함), 블로그, 인덱싱된 페이지 및 블로그에 설명과 첨부 파일을 인덱싱하는 데 Amazon Kendra 를 사용할 수 있습니다.

Amazon Kendra Confluence 데이터 소스 커넥터의 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침

지원되는 기능

Amazon Kendra Confluence 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 패턴
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Confluence 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Confluence 및 계정에서 다음과 같이 변경하십시오. AWS

Confluence에서 다음을 충족해야 합니다.

- Confluence 인스턴스의 URL을 복사했습니다. 예: *https://example.confluence.com* 또는 *https://www.example.confluence.com/* 또는 *https:// atlassian.net/*. Amazon Kendra에 연결하려면 Confluence 인스턴스 URL이 필요합니다.

  ##### ##### #### ## ### URL# atlassian.net/# ### ###.

  > **Note**
  >
  > 다음 URL 형식은 지원되지 않습니다.
  >
  > - *https://example.confluence.com/xyz*
  > - *https://www.example.confluence.com//wiki/spacekey/xxx*

- *https://atlassian.net/xyz*

> **Note**
>
> (온프레미스/서버) AWS Secrets Manager 는 포함된 엔드포인트 정보가 데이터 소스 구성 세부 정보에 지정된 엔드포인트 정보와 동일한지 Amazon Kendra 확인합니다. 이렇게 하면 사용자가 작업을 수행할 권한이 없지만 구성된 보안 암호에 액세스하여 작업을 수행하는 데 Amazon Kendra 를 프록시로 사용하는 보안 문제인 혼동된 대리자 문제를 방지하는 데 도움이 됩니다. 나중에 엔드포인트 정보를 변경하는 경우 새 보안 암호를 생성하여 이 정보를 동기화해야 합니다.

- 사용자 이름 (Confluence에 로그인하는 데 사용되는 이메일 ID) 및 암호 (Confluence API 토큰을 암호로 사용) 를 포함하는 기본 인증 자격 증명을 구성했습니다. Atlassian 계정의 API 토큰 관리를 참조하십시오.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 선택 사항: Confluence 앱 키, Confluence 앱 암호, Confluence 액세스 토큰, Confluence 인스턴스에 연결할 수 있는 Confluence 새로 고침 토큰을 포함하는 OAuth 2.0 자격 증명을 구성했습니다. Amazon Kendra 액세스 토큰이 만료되면 새로 고침 토큰을 사용하여 액세스 토큰과 새로 고침 토큰 쌍을 다시 생성할 수 있습니다. 또는 인증 프로세스를 반복할 수 있습니다. 액세스 토큰에 대한 자세한 내용을 알아보려면 OAuth 액세스 토큰 관리를 참조하세요.
- (Confluence 서버/데이터 센터에만 해당) 선택 사항: Confluence에서 개인 액세스 토큰 (PAT) 을 구성했습니다. 개인용 액세스 토큰 사용을 참조하십시오.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Confluence 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Confluence 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. Amazon Kendra API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Confluence 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 수 있도록 Amazon Kendra Confluence 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 Amazon Kendra 를 Confluence에 대해 구성하지 않은 경우 필수 조건을 참조하세요.

Console

컨플루언스에 연결하려면 Amazon Kendra

1. 에 AWS Management Console 로그인하고 콘솔을 엽니다.Amazon Kendra
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.

4. 데이터 원본 추가 페이지에서 Confluence 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Confluence 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. 소스에서 Confluence 클라우드 또는 Confluence 서버/데이터 센터를 선택합니다.
   b. Confluence URL - Confluence 호스트 URL을 입력합니다. 예: *https://example.confluence.com*.
   c. (Confluence 서버/데이터 센터에만 해당) SSL 인증서 위치 - 선택 사항 —Confluence 서버의 SSL 인증서 파일 경로를 입력합니다. Amazon S3
   d. (Confluence 서버/데이터 센터에만 해당) 웹 프록시 - 선택 사항 —웹 프록시 호스트 이름 (`http://`또는 `https://` 프로토콜 제외) 및 포트 번호 (호스트 URL 전송 프로토콜에 사용되는 포트) 를 입력합니다. 포트 번호는 0에서 65535 사이의 숫자 값이어야 합니다.
   e. 인증 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
   f. 인증 - 기본 인증, Oauth 2.0 인증 또는 개인 액세스 토큰 인증 (Confluence 서버/데이터 센터만 해당) 을 선택합니다.
   g. AWS Secrets Manager 보안 암호 - 기존 보안 암호를 선택하거나 Confluence 보안 인증 정보를 저장할 새 Secrets Manager 보안 암호를 생성합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다. 창에 다음 정보를 입력합니다.

i. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Confluence-'가 자동으로 추가됩니다.

ii. 기본 인증을 사용하는 경우 - Confluence에서 구성한 암호 이름, 사용자 이름 및 암호(암호는 Confluence API 토큰) 를 입력합니다.

OAuth2.0 인증을 사용하는 경우 - Confluence에서 구성한 암호 이름, 앱 키, 앱 암호, 액세스 토큰 및 새로 고침 토큰을 입력합니다.

(Confluence 서버/데이터 센터만 해당) 개인용 액세스 토큰 인증을 사용하는 경우 — Confluence에서 구성한 암호 이름과 Confluence 토큰을 입력합니다.

iii. 비밀번호를 저장하고 추가하세요.

h. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

i. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

j. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

k. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 콘텐츠 동기화 - 페이지, 페이지 댓글, 페이지 첨부 파일, 블로그, 블로그 댓글, 블로그 첨부 파일, 개인 공간, 보관된 공간 등의 콘텐츠 유형 중에서 동기화하도록 선택합니다.

> **Note**
>
> 페이지 댓글과 페이지 첨부 파일은 페이지를 동기화하도록 선택한 경우에만 삭제할 수 있습니다. 블로그 댓글과 블로그 첨부 파일은 블로그를 동기화하도록 선택한 경우에만 삭제할 수 있습니다.

> **Important**
>
> 추가 구성에서 스페이스 키 정규식 패턴을 지정하지 않으면 기본적으로 모든 페이지와 블로그가 크롤링됩니다.

b. 추가 구성에서 최대 파일 크기 —크롤링할 파일 크기 제한을 MB 단위로 지정합니다. Amazon Kendra Amazon Kendra 정의한 크기 제한 내에 있는 파일만 크롤링합니다. 기본 파일 크기는 50MB입니다. 최대 파일 크기는 0MB보다 크고 50MB보다 작거나 같아야 합니다.

스페이스 정규식 패턴의 경우 - 다음을 사용하여 색인에 특정 공백을 포함할지 또는 제외할지 여부를 지정합니다.

- *#### # (#: my-space-123)*

> **Note**
>
> 스페이스 키 정규식 패턴을 지정하지 않으면 기본적으로 모든 페이지와 블로그가 크롤링됩니다.

- *URL (#:. */ MySiteMyDocuments/*)
- 파일 유형 (예: .*\ *.pdf, .*\ .txt*)

엔티티 제목 정규식 패턴의 경우 - 특정 블로그, 페이지, 댓글 및 첨부 파일을 제목별로 포함하거나 제외하는 정규 표현식 패턴을 지정합니다.

> **Note**
>
> 특정 페이지나 하위 페이지의 크롤링을 포함하거나 제외하려는 경우 페이지 제목 정규식 패턴을 사용할 수 있습니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 Amazon Kendra 모든 콘텐츠가 크롤링되고 색인됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다. 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

b. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra 컨플루언스에 연결하려면

API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. TemplateConfiguration 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 CONFLUENCEV2 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 호스트 URL - 컨플루언스 호스트 URL 인스턴스를 지정합니다. 예: *https://example.confluence.com*.
- 동기화 모드 - 데이터 소스 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 인증 유형 - 인증 유형 (예:,) 을 지정합니다 (Confluence 서버만 해당). Basic OAuth2 Personal-token
- (선택 사항—Confluence 서버만 해당) SSL 인증서 위치 - SSL 인증서를 저장하는 데 사용한 S3bucketName 및 s3certificateName을 지정합니다.
- 비밀 Amazon 리소스 이름 (ARN) - Confluence에서 구성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. 기본 인증을 사용하는 경우 보안 암호는 다음 키가 있는 JSON 구조에 저장됩니다.

```
{
    "username": "email ID or user name",
    "password": "Confluence API token"
}
```

OAuth 2.0 인증을 사용하는 경우 보안 암호는 다음 키를 사용하여 JSON 구조에 저장됩니다.

```
{
    "confluenceAppKey": "app key",
    "confluenceAppSecret": "app secret",
    "confluenceAccessToken": "access token",
    "confluenceRefreshToken": "refresh token"
}
```

(Confluence 서버만 해당) 기본 인증을 사용하는 경우 암호에는 다음 키가 있는 JSON 구조가 저장됩니다.

```
{
    "hostUrl": "Confluence Server host URL",
    "username": "Confluence Server user name",
    "password": "Confluence Server password"
}
```

(Confluence 서버만 해당) 개인 액세스 토큰 인증을 사용하는 경우 암호에는 다음 키가 있는 JSON 구조가 저장됩니다.

```
{
    "hostUrl": "Confluence Server host URL",
    "patToken": "personal access token"
}
```

- IAM 역할 —역할에 Secrets Manager 보안 액세스 권한을 제공하고 IAM Confluence 커넥터 및 에 필요한 퍼블릭 API를 호출하도록 호출할 `RoleArn` 시기를 지정합니다. `CreateDataSource` Amazon Kendra자세한 내용은 Confluence 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 파일 크기 —크롤링할 최대 파일 크기를 지정합니다.
- 문서/콘텐츠 유형 - 페이지, 페이지 댓글, 페이지 첨부 파일, 블로그, 블로그 댓글, 블로그 첨부 파일, 스페이스 및 보관된 스페이스를 크롤링할지 여부를 지정합니다.
- 포함 및 제외 필터 - 특정 스페이스, 페이지, 블로그, 해당 의견 및 첨부 파일을 포함할지 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 웹 프록시 - 웹 프록시를 통해 Confluence URL 인스턴스에 연결하려는 경우 웹 프록시 정보를 지정합니다. 이 옵션을 Confluence 서버에 사용할 수 있습니다.
- 액세스 제어 목록 (ACL) - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 ACL 정보를 크롤링할지 여부를 지정합니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 필드 매핑 - Confluence 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 해당 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Confluence 템플릿 스키마를 참조하십시오.

참고

- 개인 액세스 토큰(PAT)은 Confluence 클라우드에서 사용할 수 없습니다.

사용자 지정 데이터 소스 커넥터

아직 데이터 원본 커넥터를 제공하지 Amazon Kendra 않는 저장소가 있는 경우 사용자 지정 데이터 원본을 사용하세요. 의 데이터 원본을 사용하여 리포지토리를 동기화할 수 없는 경우에도 이를 사용하여 Amazon Kendra Amazon Kendra데이터 원본이 제공하는 것과 동일한 실행 기록 측정항목을 볼 수 있습니다. 이를 사용하여 Amazon Kendra 데이터 원본과 사용자 지정 원본 간에 일관된 동기화 모니터링

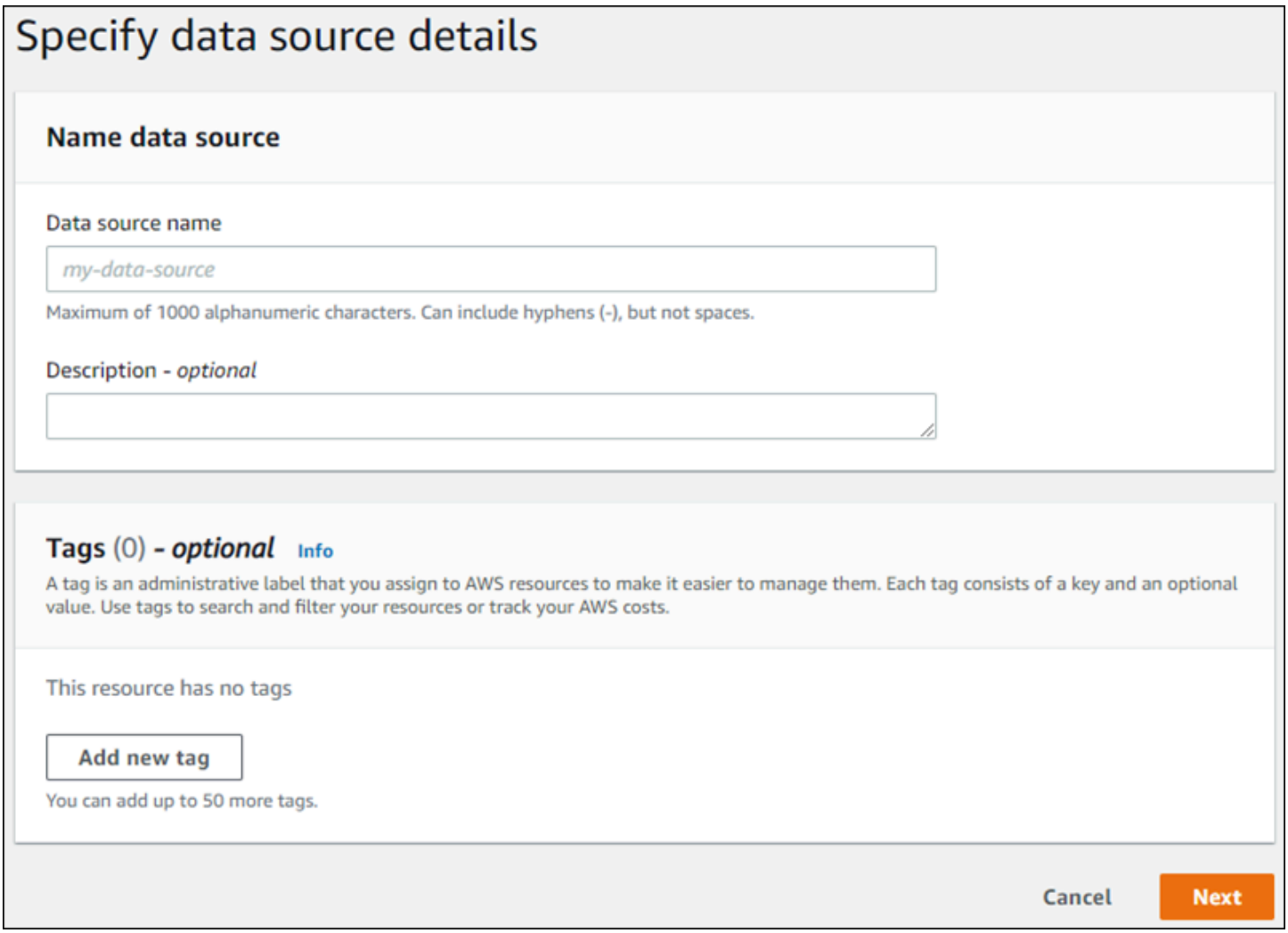
환경을 만들 수 있습니다. 특히 사용자 지정 데이터 소스를 사용하면 BatchPut문서 및 BatchDelete문서 API를 사용하여 만든 데이터 소스 커넥터의 동기화 지표를 확인할 수 있습니다.

Amazon Kendra 사용자 지정 데이터 소스 커넥터의 문제를 해결하려면 데이터 소스 문제 해결을 참조하세요.

사용자 지정 데이터 소스를 만들면 인덱싱할 문서를 선택하는 방법을 완전히 제어할 수 있습니다. Amazon Kendra 데이터 소스 동기화 작업을 모니터링하는 데 사용할 수 있는 지표 정보만 제공합니다. 데이터 소스가 인덱싱하는 문서를 결정하는 크롤러를 만들고 실행해야 합니다.

Document 객체를 사용하여 문서의 기본 제목을 지정해야 하며, `Query` 결과의 `_source_uri` DocumentAttribute응답에 `DocumentURI` 포함되도록 하려면 문서의 기본 제목을 지정해야 `DocumentTitle` 합니다.

콘솔이나 Source API를 사용하여 사용자 지정 데이터 CreateData원본의 식별자를 만듭니다. 콘솔을 사용하려면 데이터 소스에 이름을 지정하고 선택적으로 설명과 리소스 태그를 지정하세요. 데이터 소스가 생성되면 데이터 소스 ID가 표시됩니다. 데이터 소스를 인덱스와 동기화할 때 사용하려면 이 ID를 복사합니다.



`CreateDataSource` API를 사용하여 사용자 지정 데이터 소스를 생성할 수도 있습니다. API는 데이터 소스를 동기화할 때 사용할 ID를 반환합니다. `CreateDataSource` API를 사용하여 사용자 지정 데이터 소스를 만들 때는 `Configuration`, `RoleArn` 또는 `Schedule` 파라미터를 설정할 수 없습니다. 이러한 매개변수를 설정하면 `ValidationException` 예외가 Amazon Kendra 반환됩니다.

사용자 지정 데이터 소스를 사용하려면 Amazon Kendra 인덱스 업데이트를 담당하는 애플리케이션을 만드세요. 애플리케이션은 사용자가 만든 크롤러에 따라 달라집니다. 크롤러는 리포지토리의 문서를 읽고 Amazon Kendra에 무엇을 보낼지 결정합니다. 애플리케이션에서는 다음 절차를 수행합니다.

1. 리포지토리를 크롤링하고 리포지토리에서 추가, 업데이트 또는 삭제된 문서 목록을 만드세요.
2. StartDataSourceSyncJob API를 호출하여 동기화 작업이 시작되고 있음을 알립니다. 데이터 소스 ID를 제공하여 동기화 중인 데이터 소스를 식별합니다. Amazon Kendra 실행 ID를 반환하여 특정 동기화 작업을 식별합니다.

3. BatchDeleteDocument API를 호출하여 색인에서 문서를 제거합니다. 데이터 소스 ID와 실행 ID를 제공하여 동기화 중인 데이터 소스 및 이 업데이트와 관련된 작업을 식별할 수 있습니다.
4. StopDataSourceSyncJob API를 호출하여 동기화 작업 종료를 알립니다. StopDataSourceSyncJob API를 호출한 후에는 연결된 실행 ID가 더 이상 유효하지 않습니다.
5. 인덱스 및 데이터 소스 식별자와 함께 ListDataSourceSyncJobs API를 호출하여 데이터 원본의 동기화 작업을 나열하고 동기화 작업에 대한 지표를 확인합니다.

동기화 작업을 종료한 후 새 동기화 작업을 시작할 수 있습니다. 제출된 모든 문서가 인덱스에 추가되기까지 일정 시간이 걸릴 수 있습니다. ListDataSourceSyncJobs API를 사용하여 동기화 작업의 상태를 확인합니다. Status가 동기화 작업에 대해 반환한 값이 SYNCING_INDEXING인 경우 일부 문서는 여전히 인덱싱되고 있습니다. 이전 작업의 상태가 FAILED 또는 SUCCEEDED 일 때 새 동기화 작업을 시작할 수 있습니다.

StopDataSourceSyncJob API를 호출한 후에는 BatchPutDocument 또는 BatchDeleteDocument API를 호출할 때 동기화 작업 식별자를 사용할 수 없습니다. 이렇게 하면 제출된 모든 문서가 API의 FailedDocuments 응답 메시지로 반환됩니다.

필수 속성

BatchPutDocumentAPI를 Amazon Kendra 사용하여 문서를 제출할 때 각 문서에는 해당 문서가 속한 데이터 소스 및 동기화 실행을 식별하는 두 가지 속성이 필요합니다. 사용자 지정 데이터 소스의 문서를 Amazon Kendra 인덱스에 올바르게 매핑하려면 다음 두 속성을 제공해야 합니다.

- _data_source_id - 데이터 소스의 식별자입니다. 콘솔 또는 CreateDataSource API로 데이터 소스를 만들면 이 값이 반환됩니다.
- _data_source_sync_job_execution_id - 동기화 실행의 식별자. StartDataSourceSyncJob API와 인덱스 동기화를 시작하면 이 값이 반환됩니다.

다음은 사용자 지정 데이터 소스를 사용하여 문서를 인덱싱하는 데 필요한 JSON입니다.

```
{
    "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"
}
```

BatchDeleteDocument API를 사용하여 인덱스에서 문서를 제거하는 경우 DataSourceSyncJobMetricTarget 파라미터에 다음 두 필드를 지정해야 합니다.

- DataSourceId - 데이터 소스의 식별자입니다. 콘솔 또는 CreateDataSource API로 데이터 소스를 만들면 이 값이 반환됩니다.
- DataSourceSyncJobId - 동기화 실행의 식별자. StartDataSourceSyncJob API와 인덱스 동기화를 시작하면 이 값이 반환됩니다.

다음은 BatchDeleteDocument API를 사용하여 인덱스에서 문서를 삭제하는 데 필요한 JSON입니다.

```
{
    "DataSourceSyncJobMetricTarget": {
        "DataSourceId": "data source identifier",
        "DataSourceSyncJobId": "sync job identifier"
    },
    "DocumentIdList": [
        "document identifier"
    ],
    "IndexId": "index identifier"
}
```

지표 보기

동기화 작업이 완료된 후 DataSourceSyncJobMetrics API를 사용하여 동기화 작업과 관련된 지표를 가져올 수 있습니다. 이를 사용하여 사용자 지정 데이터 소스 동기화를 모니터링할 수 있습니다.

BatchPutDocument API나 BatchDeleteDocument API의 일부로 동일한 문서를 여러 번 제출하거나 추가 및 삭제를 위해 문서를 제출한 경우, 지표에서 문서가 한 번만 계산됩니다.

- DocumentsAdded - 이 동기화 작업과 관련된 BatchPutDocument API를 사용하여 제출된 문서 수가 처음으로 인덱스에 추가되었습니다. 동기화된 상태에서 문서를 두 번 이상 추가하도록 제출하는 경우 해당 문서는 지표에서 한 번만 계산됩니다.
- DocumentsDeleted - 인덱스에서 삭제된 이 동기화 작업과 연결된 BatchDeleteDocument API를 사용하여 제출된 문서 수입니다. 동기화된 상태에서 문서를 두 번 이상 삭제하도록 제출하는 경우 해당 문서는 지표에서 한 번만 계산됩니다.
- DocumentsFailed - 이 동기화 작업과 관련된 문서 중 인덱싱에 실패한 문서 수입니다. 이러한 문서는 Amazon Kendra 에 의해 인덱싱이 승인되었지만 인덱스되거나 삭제할 수 없었습니다. 에서 문서를 수락하지 않으면 BatchPutDocument 및 BatchDeleteDocument API의 FailedDocuments 응답 속성에 문서 식별자가 반환됩니다. Amazon Kendra
- DocumentsModified—이 동기화 작업과 관련된 BatchPutDocument API를 사용하여 제출된 수정된 문서 중 색인에서 수정된 문서 수. Amazon Kendra

Amazon Kendra 또한 문서를 인덱싱하는 동안 Amazon CloudWatch 메트릭을 내보냅니다. 자세한 내용은 Amazon Kendra 모니터링을 참조하십시오. Amazon CloudWatch

Amazon Kendra 사용자 지정 데이터 소스의 DocumentsScanned 메트릭을 반환하지 않습니다. 또한 Amazon Kendra 데이터 소스의 지표 문서에 나열된 CloudWatch 지표를 내보냅니다.

자세히 알아보기

사용자 지정 데이터 Amazon Kendra 소스와의 통합에 대해 자세히 알아보려면 다음을 참조하십시오.

- 사용자 지정 데이터 소스 추가 Amazon Kendra

사용자 지정 데이터 소스(Java)

다음 코드는 Java를 사용한 사용자 지정 데이터 소스의 샘플 구현을 제공합니다. 프로그램은 먼저 사용자 지정 데이터 소스를 만든 다음 인덱스에 새로 추가된 문서를 사용자 지정 데이터 소스와 동기화합니다.

다음 코드는 사용자 지정 데이터 소스를 만들고 사용하는 방법을 보여줍니다. 애플리케이션에서 사용자 지정 데이터 소스를 사용하는 경우 인덱스를 데이터 소스와 동기화할 때마다 새 데이터 소스(일회성 프로세스)를 만들 필요가 없습니다. 인덱스 ID와 데이터 소스 ID를 사용하여 데이터를 동기화합니다.

```
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

Dropbox는 클라우드 스토리지, 문서 정리, 문서 템플릿 작성 서비스를 제공하는 파일 호스팅 서비스입니다. Dropbox 사용자는 Dropbox 파일, Dropbox Paper, Dropbox Paper 템플릿, 저장된 웹 페이지 바로가기를 인덱싱하는 데 사용할 Amazon Kendra 수 있습니다. 특정 Dropbox 파일, Dropbox Paper, Dropbox Paper 템플릿, 저장된 웹 페이지 바로가기를 Amazon Kendra 인덱싱하도록 구성할 수도 있습니다.

Amazon Kendra 드롭박스 비즈니스용 드롭박스와 드롭박스 어드밴스드를 모두 지원합니다.

Amazon Kendra 콘솔과 API를 사용하여 드롭박스 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Dropbox 데이터 소스 커넥터 문제를 해결하는 방법은 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Dropbox 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Dropbox 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Dropbox와 계정에서 다음과 같이 변경하세요. AWS

Dropbox에서 다음 사항을 갖추었는지 확인하세요.

- Dropbox Advanced 계정을 만들고 관리자 사용자를 설정했습니다.
- 고유한 앱 이름을 사용하여 Dropbox 앱을 구성하고 범위 지정 액세스를 활성화했습니다. 앱 생성에 대한 Dropbox 설명서를 참조하세요.
- Dropbox 콘솔에서 전체 Dropbox 권한을 활성화하고 다음 권한을 추가했습니다.
  - files.content.read
  - files.metadata.read
  - sharing.read
  - file_requests.read
  - groups.read
  - team_info.read
  - team_data.content.read
- 기본 보안 인증을 위해 Dropbox 앱 키, Dropbox 앱 보안 암호, Dropbox 액세스 토큰을 기록해 두었습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- Dropbox 앱을 위한 임시 Oauth 2.0 액세스 토큰을 구성하고 복사했습니다. 이 토큰은 일시적이며 4시간 후에 만료됩니다. OAuth 인증에 대한 Dropbox 설명서를 참조하세요.

> **Note**
>
> 4시간 후에 만료되는 일회용 액세스 토큰을 사용하는 것보다 만료되지 않는 Dropbox 새로 고침 액세스 토큰을 만드는 것이 좋습니다. 새로 고침 액세스 토큰은 영구적이며 만료되지 않으므로 향후에도 데이터 소스를 계속 동기화할 수 있습니다.

- 권장: 중단없이 데이터 소스를 계속 동기화할 수 Amazon Kendra 있도록 만료되지 않는 Dropbox 영구 새로 고침 토큰을 구성하세요. 새로 고침 토큰에 대한 Dropbox 설명서를 참조하세요.
- Dropbox 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 AWS 계정있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Dropbox 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 비밀번호가 없는 경우 Dropbox 데이터 소스를 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 비밀번호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Dropbox 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Dropbox 데이터 소스의 필수 세부 정보를 제공해야 합니다. 아직 Dropbox를 구성하지 않았다면 을 Amazon Kendra참조하십시오. 필수 조건

Console

드롭박스에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Dropbox 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 'V2.0' 태그가 있는 Dropbox 커넥터를 선택하세요.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 사용 - 문서를 필터링하여 색인에 사용할 언어를 선택하세요. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

b. 인증 토큰 유형 - 영구 토큰 (권장) 또는 임시 액세스 토큰 중 하나를 선택합니다.

c. AWS Secrets Manager 비밀: 기존 비밀번호를 선택하거나 Dropbox 인증 자격 Secrets Manager 증명을 저장할 새 비밀번호를 만드세요. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름. 비밀 이름에는 접두사 AmazonKendra '-Dropbox-'가 자동으로 추가됩니다.

B. 앱 키, 앱 암호 및 토큰 정보 (영구 또는 임시) 의 경우 - Dropbox에 구성된 인증 자격 증명 값을 입력합니다.

ii. 비밀번호를 저장하고 추가하세요.

d. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

e. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

f. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

g. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 항목 또는 콘텐츠 유형 선택 —크롤링하려는 Dropbox 항목 또는 콘텐츠 형식을 선택합니다.

b. 정규식 패턴의 추가 구성 - 특정 파일을 포함하거나 제외하기 위한 정규 표현식 패턴을 추가합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하면 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 파일, Dropbox Paper, Dropbox Paper 템플릿 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 선택하여 색인에 매핑할 수 있습니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

드롭박스에 연결하려면 Amazon Kendra

API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다 TemplateConfiguration. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 DROPBOX 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 액세스 토큰 유형 - 인증 자격 증명을 저장하는 AWS Secrets Manager 암호에 영구 액세스 토큰을 사용할지 아니면 임시 액세스 토큰을 사용할지를 지정합니다.

> **Note**
>
> 4시간 후에 만료되는 일회용 액세스 토큰을 사용하는 것보다 Dropbox에서 만료되지 않는 새로 고침 액세스 토큰을 만드는 것이 좋습니다. Dropbox 개발자 콘솔에서 앱과 새로 고침 액세스 토큰을 만들고 암호에 액세스 토큰을 입력합니다.

- 비밀 Amazon 리소스 이름 (ARN): Dropbox 계정의 인증 자격 증명이 포함된 Secrets Manager 비밀번호의 Amazon 리소스 이름 (ARN) 을 입력합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "appKey": "Dropbox app key",
    "appSecret": "Dropbox app secret",
    "accesstoken": "temporary access token or refresh access token"
}
```

- ID 크롤러 —ID 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- IAM 역할 `CreateDataSource` Secrets Manager —비밀번호에 액세스할 수 있는 권한을 가진 IAM 역할을 제공하고 Dropbox 커넥터 및 에 필요한 공개 API를 호출하기 위해 어떤 역할을 호출할지 `RoleArn` 지정하십시오. Amazon Kendra자세한 내용은 Dropbox 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 문서/콘텐츠 유형: Dropbox에 저장된 파일, Dropbox Paper 문서, Dropbox Paper 템플릿, 웹 페이지 바로 가기를 크롤링할지 여부를 지정합니다.
- 포함 및 제외 필터 - 특정 파일을 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 액세스 제어 목록 (ACL) - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 ACL 정보를 크롤링할지 여부를 지정합니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Dropbox 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Dropbox 템플릿 스키마를 참조하십시오.

자세히 알아보기

Dropbox 데이터 소스와 Amazon Kendra의 통합에 대한 자세한 내용은 다음을 참조하세요.

- Amazon Kendra용 Dropbox 커넥터를 사용하여 Dropbox 콘텐츠 인덱싱

Drupal

Drupal은 웹 사이트 및 웹 애플리케이션을 만드는 데 사용할 수 있는 오픈 소스 콘텐츠 관리 시스템(CMS)입니다. Drupal에서 다음을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

- 콘텐츠 - 기사, 기본 페이지, 기본 블록, 사용자 정의 콘텐츠 유형, 사용자 정의 블록 유형, 사용자 정의 콘텐츠 유형, 사용자 정의 블록 유형

- 설명 - 모든 콘텐츠 유형 및 블록 유형에 해당
- 첨부 파일 - 모든 콘텐츠 유형 및 블록 유형에 해당

Amazon Kendra 콘솔 또는 API를 Amazon Kendra 사용하여 Drupal 데이터 소스에 연결할 수 있습니다. TemplateConfiguration

Amazon Kendra Drupal 데이터 소스 커넥터의 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

Amazon Kendra Drupal 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Drupal 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Drupal 및 계정에서 다음과 같이 변경하십시오. AWS

Drupal에서 다음 사항을 갖추었는지 확인하세요.

- Drupal (Standard) Suite 계정과 관리자 역할을 가진 사용자를 만들었습니다.
- Drupal 사이트 이름을 복사하고 호스트 URL을 구성했습니다. 예: *https://<hostname>/<drupalsitename>*.

- 사용자 이름(Drupal 웹 사이트 로그인 사용자 이름) 및 암호(Drupal 웹 사이트 암호)를 포함하는 기본 보안 인증을 구성했습니다.
- 권장: OAuth 2.0 보안 인증 토큰을 구성했습니다. 이 토큰을 Drupal 암호 부여, 클라이언트 ID, 클라이언트 보안 암호, 사용자 이름(Drupal 웹 사이트 로그인 사용자 이름) 및 암호(Drupal 웹 사이트 암호)와 함께 사용하여 Amazon Kendra에 연결합니다.
- 관리자 역할을 사용하여 Drupal 계정에 다음 권한을 추가했습니다.
  - 블록 관리
  - 블록_콘텐츠 디스플레이 관리
  - 블록_콘텐츠 필드 관리
  - 블록_콘텐츠 양식 디스플레이 관리
  - 뷰 관리
  - 사용자 이메일 주소 보기
  - 게시되지 않은 자체 콘텐츠 보기
  - 페이지 수정본 보기
  - 기사 수정본 보기
  - 모든 수정본 보기
  - 관리 테마 보기
  - 콘텐츠 액세스
  - 콘텐츠 개요 액세스
  - 설명 액세스
  - 콘텐츠 검색
  - 파일 개요 액세스
  - 상황별 링크에 액세스

> **Note**
>
> 사용자 정의 콘텐츠 유형 또는 사용자 정의 블록 유형이 있거나 Drupal 웹 사이트에 보기 및 블록이 추가된 경우 관리자 액세스 권한을 제공해야 합니다.

에 다음이 AWS 계정있는지 확인하십시오.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.

- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Drupal 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Drupal 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Drupal 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Drupal 자격 증명의 세부 정보를 제공해야 합니다. 아직 Drupal을 구성하지 않은 경우 참조하십시오. Amazon Kendra 필수 조건

Console

드루팔에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.

4. 데이터 원본 추가 페이지에서 Drupal 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Drupal 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
    a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
    b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
    c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
    d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
    e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
    a. 소스 내, 호스트 URL의 경우 - Drupal 사이트의 호스트 URL. 예: *https://<hostname>/<drupalsitename>*.
    b. SSL 인증서 위치 - Amazon S3 버킷에 저장된 SSL 인증서의 경로를 입력합니다.
    c. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
    d. 인증의 경우 - 사용 사례에 따라 기본 인증, OAuth 2.0 인증 중에서 선택합니다.
    e. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 Secrets Manager 암호를 생성하여 Drupal 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.
        i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.
            A. 기본 인증을 선택한 경우 복사한 보안 암호 이름, 사용자 이름(Drupal 사이트 사용자 이름), 암호(Drupal 사이트 암호)를 입력하고 저장 및 보안 암호 추가를 선택합니다.
            B. OAuth 2.0 인증을 선택한 경우 Drupal 계정에서 생성한 보안 암호 이름, 사용자 이름(Drupal 사이트 사용자 이름), 암호(Drupal 사이트 암호), 클라이언트 ID, 클라이언트 보안 암호를 입력하고 보안 암호 저장 및 추가를 선택합니다.

ii. 저장을 선택합니다.

f. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

g. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

> **Note**
>
> 기사, 기본 페이지, 기본 블록을 크롤링하도록 선택하면 해당 기본 필드가 자동으로 동기화됩니다. 설명, 첨부 파일, 사용자 지정 필드 및 기타 사용자 지정 개체를 동기화하도록 선택할 수도 있습니다.

- 개체 선택의 경우:
  - 기사 - 기사, 설명, 첨부 파일을 크롤링할지 여부를 선택합니다.
  - 기본 페이지 - 기본 페이지, 설명, 첨부 파일을 크롤링할지 여부를 선택합니다.
  - 기본 블록 - 기본 블록, 설명, 첨부 파일을 크롤링할지 여부를 선택합니다.

- 사용자 지정 콘텐츠 유형 및 사용자 지정 블록을 추가하도록 선택할 수도 있습니다.

b. 추가 구성 - 선택 사항의 경우:

- 정규식 패턴의 경우 - 정규 표현식 패턴을 추가하여 특정 개체 제목 및 파일 이름을 포함하거나 제외합니다. 최대 100개의 패턴을 추가할 수 있습니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 콘텐츠, 댓글 및 첨부 파일의 경우 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Drupal에 Amazon Kendra 연결하려면

TemplateConfiguration API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 DRUPAL 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - Drupal 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다.

  기본 인증을 사용하는 경우 보안 암호는 다음 키가 있는 JSON 구조에 저장됩니다.

```
{
    "username": "user name",
    "password": "password"
}
```

  OAuth 2.0 인증을 사용하는 경우 보안 암호는 다음 키를 사용하여 JSON 구조에 저장됩니다.

```
{
    "username": "user name",
    "password": "password",
    "clientId": "client id",
    "clientSecret": "client secret"
}
```

> **Note**
>
> > **Note**
> >
> > 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 권한을 역할을 제공하고 IAM Drupal 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Drupal 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 콘텐츠, 설명 및 첨부 파일을 포함할지 여부를 지정할 수 있습니다. 또한 정규 표현식 패턴을 지정하여 콘텐츠, 설명 및 첨부 파일을 포함하거나 제외할 수 있습니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

- 필드 매핑 - Drupal 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 해당 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Drupal 템플릿 스키마를 참조하세요.

참고

- Drupal API에는 공식적인 제한 한도가 없습니다.
- Drupal에는 Java SDK를 사용할 수 없습니다.
- Drupal 데이터는 네이티브 JSON API를 사용해야만 가져올 수 있습니다.
- Drupal 뷰와 연결되지 않은 콘텐츠 유형은 크롤링할 수 없습니다.
- Drupal 블록에서 데이터를 크롤링하려면 관리자 액세스 권한이 필요합니다.
- HTTP 동사를 사용하여 사용자 정의 콘텐츠 유형을 생성하는 데 사용할 수 있는 JSON API는 없습니다.
- 기사, 기본 페이지, 기본 블록, 사용자 정의 콘텐츠 유형 및 사용자 정의 블록 유형의 문서 본문과 설명은 HTML 형식으로 표시됩니다. HTML 콘텐츠의 형식이 올바르지 않으면 HTML 관련 태그가 문서 본문과 설명에 나타나고 Amazon Kendra 검색 결과에 표시됩니다.
- 설명이나 본문이 없는 콘텐츠 유형 및 블록 유형은 인제스트되지 않습니다. Amazon Kendra해당 콘텐츠 또는 블록 유형의 댓글과 첨부 파일만 색인에 통합됩니다. Amazon Kendra

GitHub

GitHub 버전 제어와 함께 코드 저장 및 관리 서비스를 제공하는 소프트웨어 개발을 위한 웹 기반 호스팅 서비스입니다. GitHub 엔터프라이즈 클라우드 (SaaS) 및 GitHub 엔터프라이즈 서버 (온프레미스) 리포지토리 파일을 인덱싱하고, 요청 및 풀 리퀘스트를 발행하고, 풀 리퀘스트 코멘트를 발행하고, 풀 리퀘스트 코멘트 첨부 파일을 발행하는 데 사용할 Amazon Kendra 수 있습니다. 또한 특정 파일을 포함하거나 제외하도록 선택할 수 있습니다.

> **Note**
>
> Amazon Kendra 이제 업그레이드된 GitHub 커넥터를 지원합니다.
> 콘솔이 자동으로 업그레이드되었습니다. 콘솔에서 새로 만드는 모든 커넥터는 업그레이드된 아키텍처를 사용합니다. API를 사용하는 경우 이제 TemplateConfiguration개체 대신 개체를 사용하여 커넥터를 구성해야 합니다. `GitHubConfiguration`
> 이전 콘솔과 API 아키텍처를 사용하여 구성된 커넥터는 구성된 대로 계속 작동합니다. 하지만 편집하거나 업데이트할 수는 없습니다. 커넥터 구성을 편집하거나 업데이트하려면 새 커넥터를 만들어야 합니다.
> 커넥터 워크플로를 업그레이드된 버전으로 마이그레이션하는 것이 좋습니다. 이전 아키텍처를 사용하여 구성된 커넥터에 대한 지원은 2024년 6월에 종료될 예정입니다.

Amazon Kendra 콘솔과 TemplateConfigurationAPI를 Amazon Kendra 사용하여 GitHub 데이터 소스에 연결할 수 있습니다.

Amazon Kendra GitHub 데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra GitHub 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

GitHub 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 GitHub 및 AWS 계정에서 다음과 같이 변경하십시오.

에서 GitHub 다음 사항을 갖추었는지 확인하세요.

- GitHub 조직에 대한 관리자 권한을 가진 GitHub 사용자를 생성했습니다.
- Git Hub에서 개인 액세스 토큰을 인증 자격 증명으로 사용하도록 구성했습니다. 개인용 액세스 토큰 생성에 대한 GitHub 설명서를 참조하십시오.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 권장: 인증 자격 증명을 위한 OAuth 토큰을 구성했습니다. API 제한 한도 및 커넥터 성능을 높이려면 OAuth 토큰을 사용하세요. OAuth GitHub 인증에 대한 설명서를 참조하십시오.
- 사용하는 GitHub 서비스 유형의 GitHub 호스트 URL을 기록해 두었습니다. 예를 들어 GitHub 클라우드의 호스트 URL은 *https://api.github.com* 일 수 있고 GitHub 서버의 호스트 URL은 *on-prem-host-urlhttps:///api/v3/#* 수 있습니다.
- 연결하려는 GitHub 엔터프라이즈 클라우드 (SaaS) 계정 또는 GitHub 엔터프라이즈 서버 (온-프레미스) 계정의 조직 이름을 기록해 두었습니다. GitHub GitHub 데스크톱에 로그인하고 프로필 사진 드롭다운에서 내 조직을 선택하여 조직 이름을 찾을 수 있습니다.
- 선택 사항 (서버만 해당): SSL 인증서를 생성하고 Amazon S3 버킷에 저장된 인증서의 경로를 복사했습니다. 보안 SSL 연결이 필요한 GitHub 경우 이를 사용하여 연결할 수 있습니다. OpenSSL을 사용하여 컴퓨터에 자체 서명된 X509 인증서를 생성하기만 하면 됩니다. OpenSSL을 사용하여 X509 인증서를 만드는 예제는 X509 인증서 생성 및 서명을 참조하세요.
- 다음 권한을 추가했습니다.

  GitHub 엔터프라이즈 클라우드 (SaaS) 용

  - `repo:status`— 퍼블릭 및 프라이빗 리포지토리의 커밋 상태에 대한 읽기/쓰기 액세스 권한을 부여합니다. 이 범위는 코드에 대한 액세스 권한을 부여하지 않고 다른 사용자나 서비스에 비공개 리포지토리 커밋 상태에 대한 액세스 권한을 부여하는 데에만 필요합니다.

- repo_deployment— 퍼블릭 및 프라이빗 리포지토리의 배포 상태에 대한 액세스 권한을 부여합니다. 이 범위는 코드에 대한 액세스 권한을 부여하지 않고 다른 사용자나 서비스에 배포 상태에 대한 액세스 권한을 부여하는 데에만 필요합니다.
- public_repo— 공용 리포지토리에 대한 액세스를 제한합니다. 여기에는 코드에 대한 읽기/쓰기 권한, 커밋 상태, 리포지토리 프로젝트, 공동 작업자, 공용 리포지토리 및 조직의 배포 상태가 포함됩니다. 공용 리포지토리를 시작하는 데도 필요합니다.
- repo:invite— 저장소에서 공동 작업하라는 초대에 대한 수락/거부 기능을 부여합니다. 이 범위는 코드에 대한 액세스 권한을 부여하지 않고 다른 사용자나 서비스에 초대에 대한 액세스 권한을 부여하는 데에만 필요합니다.
- security_events— 부여: 코드 스캔 API의 보안 이벤트에 대한 읽기 및 쓰기 액세스 권한. 이 범위는 코드에 대한 액세스 권한을 부여하지 않고 다른 사용자나 서비스에 보안 이벤트에 대한 액세스 권한을 부여하는 데에만 필요합니다.
- read:org— 조직 멤버십, 조직 프로젝트 및 팀 멤버십에 대한 읽기 전용 액세스.
- user:email— 사용자의 이메일 주소에 대한 읽기 권한을 부여합니다. Amazon Kendra에서 ACL을 크롤링하는 데 필요합니다.
- user:follow— 다른 사용자를 팔로우하거나 언팔로우할 수 있는 액세스 권한을 부여합니다. Amazon Kendra에서 ACL을 크롤링하는 데 필요합니다.
- read:user— 사용자의 프로필 데이터를 읽을 수 있는 액세스 권한을 부여합니다. Amazon Kendra에서 ACL을 크롤링하는 데 필요합니다.
- workflow— 작업 워크플로 파일을 추가하고 업데이트할 GitHub 수 있는 권한을 부여합니다. 경로와 컨텐트가 같은 같은 파일이 같은 저장소의 다른 분기에 있는 경우 이 범위 없이 워크플로 파일을 커밋할 수 있습니다.

자세한 내용은 문서의 OAuth 앱 범위를 참조하십시오. GitHub

GitHub 엔터프라이즈 서버용 (온프레미스)

- repo:status— 퍼블릭 및 프라이빗 리포지토리의 커밋 상태에 대한 읽기/쓰기 액세스 권한을 부여합니다. 이 범위는 코드에 대한 액세스 권한을 부여하지 않고 다른 사용자나 서비스에 비공개 리포지토리 커밋 상태에 대한 액세스 권한을 부여하는 데에만 필요합니다.
- repo_deployment— 퍼블릭 및 프라이빗 리포지토리의 배포 상태에 대한 액세스 권한을 부여합니다. 이 범위는 코드에 대한 액세스 권한을 부여하지 않고 다른 사용자나 서비스에 배포 상태에 대한 액세스 권한을 부여하는 데에만 필요합니다.
- public_repo— 공용 리포지토리에 대한 액세스를 제한합니다. 여기에는 코드에 대한 읽기/쓰기 권한, 커밋 상태, 리포지토리 프로젝트, 공동 작업자, 공용 리포지토리 및 조직의 배포 상태가 포함됩니다. 공용 리포지토리를 시작하는 데도 필요합니다.

- `repo:invite`— 저장소에서 공동 작업하라는 초대에 대한 수락/거부 기능을 부여합니다. 이 범위는 코드에 대한 액세스 권한을 부여하지 않고 다른 사용자나 서비스에 초대에 대한 액세스 권한을 부여하는 데에만 필요합니다.
- `security_events`— 부여: 코드 스캔 API의 보안 이벤트에 대한 읽기 및 쓰기 액세스 권한. 이 범위는 코드에 대한 액세스 권한을 부여하지 않고 다른 사용자나 서비스에 보안 이벤트에 대한 액세스 권한을 부여하는 데에만 필요합니다.
- `read:user`— 사용자의 프로필 데이터를 읽을 수 있는 액세스 권한을 부여합니다. Amazon Q 비즈니스에서 ACL을 크롤링하는 데 필요합니다.
- `user:email`— 사용자의 이메일 주소에 대한 읽기 액세스 권한을 부여합니다. Amazon Q 비즈니스에서 ACL을 크롤링하는 데 필요합니다.
- `user:follow`— 다른 사용자를 팔로우하거나 언팔로우할 수 있는 액세스 권한을 부여합니다. Amazon Q 비즈니스에서 ACL을 크롤링하는 데 필요합니다.
- `site_admin`— 사이트 관리자에게 GitHub 엔터프라이즈 서버 관리 API 엔드포인트에 대한 액세스 권한을 부여합니다.
- `workflow`— GitHub 작업 워크플로 파일을 추가하고 업데이트할 수 있는 권한을 부여합니다. 경로와 컨텐트가 같은 같은 파일이 같은 저장소의 다른 분기에 있는 경우 이 범위 없이 워크플로 파일을 커밋할 수 있습니다.

  자세한 내용은 GitHub 문서의 OAuth 앱 범위 및 개발자의 OAuth 앱 범위 이해를 참조하십시오. GitHub

- 동일한 색인에 사용하려는 다른 데이터 소스에서 각 문서가 고유한지 확인했습니다. GitHub 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하세요. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- GitHub 인증 자격 증명을 AWS Secrets Manager 비밀에 저장하고 API를 사용하는 경우 비밀의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 GitHub 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

GitHub 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 GitHub 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 GitHub 을 Amazon Kendra참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 GitHub

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 GitHub 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 GitHub 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. GitHub출처 GitHub —엔터프라이즈 클라우드와 GitHub엔터프라이즈 서버 중에서 선택하세요.

b. GitHub 호스트 URL —예를 들어 GitHub 클라우드의 호스트 URL은 *https://api.github.com*, GitHub 서버의 호스트 URL은 *https:///api/v3/#* 수 있습니다. on-prem-host-url

c. GitHub 조직 이름 - 조직 이름을 입력합니다. GitHub GitHub 계정에서 조직 정보를 찾을 수 있습니다.

d. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

e. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 GitHub 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '- GitHub -'가 자동으로 추가됩니다.

B. GitHub토큰의 경우 - 에 구성된 인증 자격 증명 값을 입력합니다. GitHub

ii. 암호를 저장하고 추가합니다.

f. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

g. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon

Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 리포지토리 선택 - 모든 리포지토리를 크롤링하도록 선택하거나 선택합니다.

선택한 리포지토리를 크롤링하기로 선택한 경우 리포지토리의 이름을 추가하고 선택적으로 특정 분기의 이름을 추가합니다.

b. 콘텐츠 유형 - 파일, 이슈, 풀 리퀘스트 등에서 크롤링하려는 콘텐츠 유형을 선택합니다.

c. 정규식 패턴 - 특정 페이지 및 자산을 포함하거나 제외하기 위한 정규 표현식 패턴을 추가합니다.

d. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

e. 빈도에 대한 동기화 실행 일정 —데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

f. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra 연결하려면 GitHub

TemplateConfiguration API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 `GITHUB` 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 `TEMPLATE` 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- GitHub유형 - 유형을 `SAAS` 또는 `ON_PREMISE` 로 지정합니다.
- 호스트 URL - GitHub 호스트 URL 또는 API 엔드포인트 URL을 지정합니다. 예를 들어 GitHub SaaS/엔터프라이즈 클라우드를 사용하는 경우 호스트 URL은 일 수 있고`https://api.github.com`, GitHub 온프레미스/엔터프라이즈 서버의 경우 호스트 URL은 다음과 같을 수 있습니다. `https://on-prem-host-url/api/v3/`
- 조직 이름 - 계정의 조직 이름을 지정합니다. GitHub GitHub 데스크톱에 로그인하고 프로필 사진 드롭다운에서 내 조직을 선택하면 조직 이름을 찾을 수 있습니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.

  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- ID 크롤러 Amazon Kendra—의 ID 크롤러를 활성화할지 여부를 지정합니다. ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. GitHub 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "personalToken": "token"
}
```

- IAM 역할 - CreateDataSource Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 RoleArn 시기를 지정합니다. GitHub Amazon Kendra자세한 내용은 GitHub 데이터 원본의IAM 역할을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - CreateDataSource를 호출할 때 VpcConfiguration을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.

> **Note**
>
> GitHub 서버를 사용하는 경우 서버에 Amazon VPC 연결하려면 를 GitHub 사용해야 합니다.

- 리포지토리 필터 - 이름 및 브랜치 이름을 기준으로 리포지토리를 필터링합니다.
- 문서/콘텐츠 유형 - 리포지토리 문서, 이슈, 이슈 코멘트, 코멘트 첨부 파일 발행, 풀 리퀘스트, 풀 리퀘스트 코멘트, 풀 리퀘스트 코멘트 첨부 파일을 크롤링할지 여부를 지정합니다.
- 포함 및 제외 필터 - 특정 파일 및 폴더를 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 액세스 제어 목록 (ACL) - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 ACL 정보를 크롤링할지 여부를 지정합니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - GitHub 데이터 원본 필드를 인덱스 필드에 매핑하도록 선택합니다. Amazon Kendra 문서, 커밋, 이슈, 이슈 첨부 파일, 이슈 코멘트, 풀 리퀘스트, 풀 리퀘스트 첨부, 풀 리퀘스트 코멘트 등의 필드를 포함할 수 있습니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> Amazon Kendra에서 문서를 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 소스의 문서 본문 필드 이름을 인덱스 필드 이름에 매핑해야 합니다. `_document_body` 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 GitHub 템플릿 스키마를 참조하세요.

자세히 알아보기

GitHub 데이터 Amazon Kendra 소스와의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- 커넥터의 힘으로 GitHub 리포지토리 검색을 재구상해 보세요. Amazon Kendra GitHub

Gmail

Gmail은 Google에서 개발한 이메일 클라이언트로, 이를 통해 첨부 파일이 포함된 이메일 메시지를 보낼 수 있습니다. 폴더와 라벨을 사용하여 이메일 수신함 내에서 Gmail 메시지를 정렬하고 저장할 수 있습니다. 이메일 메시지 및 메시지 첨부파일을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다. 인덱싱할 특정 이메일 메시지, 메시지 첨부 파일 및 레이블을 포함하거나 Amazon Kendra 제외하도록 구성할 수도 있습니다.

Amazon Kendra 콘솔과 API를 사용하여 Gmail 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Gmail 데이터 소스 커넥터 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

지원되는 기능

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Gmail 데이터 소스의 색인을 생성하는 Amazon Kendra 데 사용하려면 먼저 Gmail과 계정에서 다음과 같이 변경하세요. AWS

Gmail에서 다음 사항을 갖추었는지 확인하세요.

- Google Cloud Platform 관리자 계정을 만들고 Google Cloud 프로젝트를 생성했습니다.
- 관리자 계정에서 Gmail API 및 관리자 SDK API를 활성화했습니다.
- 서비스 계정을 만들고 Gmail용 JSON 프라이빗 키를 다운로드했습니다. 프라이빗 키를 만들고 액세스하는 방법에 대한 자세한 내용은 서비스 계정 키 생성 및 서비스 계정 보안 인증 생성 방법에 대한 Google Cloud 설명서를 참조하세요.
- 인증 자격 증명으로 사용할 관리자 계정 이메일, 서비스 계정 이메일, 비공개 키를 복사했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 인덱싱하려는 사용자 및 공유 디렉터리에 대해 다음과 같은 Oauth 범위(관리자 역할 사용)를 추가했습니다.
  - https://www.googleapis.com/auth/admin.directory.user.readonly
  - https://www.googleapis.com/auth/gmail.readonly
- Gmail 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Gmail 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 비밀번호가 없는 경우 Gmail 데이터 소스를 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 비밀번호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Gmail 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Gmail 자격 증명의 세부정보를 제공해야 합니다. 아직 Gmail을 구성하지 않았다면 을 Amazon Kendra 참조하십시오. 필수 조건

Console

Gmail에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 소스 추가 페이지에서 Gmail 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2(해당하는 경우) 를 사용하는 경우 'V2.0' 태그가 있는 Gmail 커넥터를 선택하세요.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 색인을 기준으로 문서를 필터링할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

b. 비밀번호 인증에서 - 기존 비밀번호를 선택하거나 새 Secrets Manager 비밀번호를 만들어 Gmail 인증 자격 증명을 저장합니다.AWS Secrets Manager 새 비밀번호를 만들기로 선택하면 AWS Secrets Manager 비밀 창이 열립니다.

- AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름.

B. 클라이언트 이메일 - Google 서비스 계정에서 복사한 클라이언트 이메일.

C. 관리자 계정 이메일 - 사용하려는 관리자 계정 이메일.

D. 프라이빗 키 - Google 서비스 계정에서 복사한 프라이빗 키.

E. 비밀번호를 저장하고 추가하세요.

c. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

d. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

e. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 엔티티 유형의 경우 - 메시지 첨부 파일을 동기화하도록 선택합니다.

b. (선택 사항) 추가 구성에서 다음 정보를 입력합니다.

i. 날짜 범위 - 날짜 범위를 입력하여 크롤링하려는 이메일의 시작 날짜와 종료 날짜를 지정합니다.

ii. 이메일 도메인 - “받는 사람”, “보낸 사람”, “cc” 및 “bcc” 이메일 도메인을 기반으로 특정 이메일을 포함하거나 제외합니다.

iii. 제목의 키워드 - 이메일 제목의 키워드를 기준으로 이메일을 포함하거나 제외합니다.

> **Note**
>
> 입력한 모든 주제 키워드와 일치하는 문서를 포함하도록 선택할 수도 있습니다.

iv. 레이블 - 정규 표현식 패턴을 추가하여 특정 이메일 레이블을 포함하거나 제외합니다.

v. 첨부 파일 - 정규 표현식 패턴을 추가하여 특정 이메일 첨부 파일을 포함하거나 제외합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

> **Important**
>
> 영구 삭제된 Gmail 메시지, 신규, 수정 또는 삭제된 콘텐츠 동기화를 업데이트하는 API가 없기 때문입니다.
>
> - Gmail에서 영구 삭제된 메일은 색인에서 삭제되지 않습니다. Amazon Kendra
> - Gmail 이메일 라벨의 변경사항을 동기화하지 않습니다.

Gmail 데이터 소스 라벨 변경사항과 영구 삭제된 이메일 메시지를 Amazon Kendra 인덱스에 동기화하려면 정기적으로 전체 크롤링을 실행해야 합니다.

d. 동기화 실행 일정에서 빈도 - 데이터 소스 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

> **Note**
>
> Amazon Kendra Gmail 데이터 소스 커넥터는 API 제한으로 인해 맞춤 색인 필드 생성을 지원하지 않습니다.

b. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Gmail에 Amazon Kendra 연결하려면

TemplateConfiguration API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 소스 - TemplateConfigurationJSON 스키마를 사용할 `GMAIL` 때와 같이 데이터 소스 유형을 지정합니다. 또한 API를 `TEMPLATE` 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.

- FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
- FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

> **Important**
>
> 영구 삭제된 Gmail 메시지, 신규, 수정 또는 삭제된 콘텐츠 동기화를 업데이트하는 API가 없기 때문입니다.
>
> - Gmail에서 영구 삭제된 메일은 색인에서 삭제되지 않습니다. Amazon Kendra
> - Gmail 이메일 라벨의 변경사항을 동기화하지 않습니다.
>
> Gmail 데이터 소스 라벨 변경사항과 영구 삭제된 이메일 메시지를 Amazon Kendra 색인에 동기화하려면 정기적으로 전체 크롤링을 실행해야 합니다.

- 비밀 Amazon 리소스 이름 (ARN) - Gmail 계정의 인증 자격 증명이 포함된 Secrets Manager 비밀의 Amazon 리소스 이름 (ARN) 을 입력합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "adminAccountEmailId": "service account email",
    "clientEmailId": "user account email",
    "privateKey": "private key"
}
```

- IAM 역할 - CreateDataSource Secrets Manager 비밀에 액세스할 권한을 가진 IAM 역할을 제공하고 Gmail 커넥터 및 Gmail 커넥터에 필요한 공개 API를 호출하기 위해 호출할 RoleArn 시기를 지정합니다. Amazon Kendra자세한 내용은 Gmail 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - CreateDataSource를 호출할 때 VpcConfiguration을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 특정 “받는 사람”, “보낸 사람”, “cc”, “bcc” 이메일을 포함할지 아니면 제외할지를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Gmail 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

> **Note**
>
> Amazon Kendra Gmail 데이터 소스 커넥터는 API 제한으로 인해 맞춤 색인 필드 생성을 지원하지 않습니다.

구성해야 할 기타 중요한 JSON 키 목록은 Gmail 템플릿 스키마를 참조하세요.

자세히 알아보기

Gmail 데이터 Amazon Kendra 소스와의 통합에 대해 자세히 알아보려면 다음을 참조하세요.

- Amazon Kendra용 Gmail 커넥터를 사용하여 Google Workspace의 이메일을 지능적으로 검색하세요.

참고

- 영구 삭제된 Gmail 메시지를 업데이트하는 API가 없으므로 FULL_CRAWL/새 콘텐츠, 수정 콘텐츠, 삭제된 콘텐츠가 동기화됩니다.
  - Gmail에서 영구 삭제된 메일은 색인에서 삭제되지 않습니다. Amazon Kendra
  - Gmail 이메일 라벨의 변경사항을 동기화하지 않습니다.

  Gmail 데이터 소스 라벨 변경사항과 영구 삭제된 이메일 메시지를 Amazon Kendra 색인에 동기화하려면 정기적으로 전체 크롤링을 실행해야 합니다.
- Amazon Kendra Gmail 데이터 소스 커넥터는 API 제한으로 인해 맞춤 색인 필드 생성을 지원하지 않습니다.

Google Drive

Google Drive는 클라우드 기반 파일 스토리지 서비스입니다. Google Drive 데이터 소스의 공유 드라이브, 내 드라이브, 나와 공유 폴더에 저장된 문서를 인덱싱하는 데 Amazon Kendra 를 사용할 수 있습니다. Google Workspace 문서와 문서 유형에 나열된 문서를 모두 인덱싱할 수 있습니다. 포함 및 제외 필터를 사용하여 파일 이름, 파일 유형, 파일 경로별로 콘텐츠를 인덱싱할 수도 있습니다.

Amazon Kendra 콘솔, TemplateConfigurationAPI 또는 API를 Amazon Kendra 사용하여 Google 드라이브 데이터 소스에 연결할 수 있습니다. GoogleDriveConfiguration

Amazon Kendra Google 드라이브 커넥터에는 두 가지 버전이 있습니다. 각 버전에 지원되는 기능은 다음과 같습니다.

구글 드라이브 커넥터 V1.0/ API GoogleDriveConfiguration

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터

구글 드라이브 커넥터 V2.0/ API TemplateConfiguration

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화

- Virtual Private Cloud(VPC)

> **Note**
>
> 구글 드라이브 커넥터 V1.0/구글 DriveConfiguration API에 대한 지원은 2023년에 종료될 예정입니다. Google 드라이브 커넥터 V2.0/ API로 마이그레이션하거나 이를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra Google 드라이브 데이터 소스 커넥터 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- Google Drive 커넥터 V1.0
- Google Drive 커넥터 V2.0

Google Drive 커넥터 V1.0

Google Drive는 클라우드 기반 파일 스토리지 서비스입니다. Google 드라이브 데이터 원본의 공유 드라이브, 내 드라이브, 나와 공유한 항목 폴더에 저장된 문서 및 댓글을 색인화하는 데 사용할 Amazon Kendra 수 있습니다. Google Workspace 문서와 문서 유형에 나열된 문서를 인덱싱할 수 있습니다. 포함 및 제외 필터를 사용하여 파일 이름, 파일 유형, 파일 경로별로 콘텐츠를 인덱싱할 수도 있습니다.

> **Note**
>
> 구글 드라이브 커넥터 V1.0/구글 DriveConfiguration API에 대한 지원은 2023년에 종료될 예정입니다. Google 드라이브 커넥터 V2.0/ API로 마이그레이션하거나 이를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra Google 드라이브 데이터 소스 커넥터 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침

- 자세히 알아보기

지원되는 기능

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터

필수 조건

Google 드라이브 데이터 소스의 색인을 생성하는 Amazon Kendra 데 사용하려면 먼저 Google 드라이브와 AWS 계정에서 다음과 같이 변경하세요.

Google Drive에서 다음 사항을 갖추었는지 확인하세요.

- 수퍼 관리자 역할에 의해 액세스 권한을 부여 받은 경우 또는 관리자 권한이 있는 사용자일 수 있습니다. 수퍼 관리자 역할을 통해 액세스 권한을 부여 받은 경우 본인에게 수퍼 관리자 역할이 필요하지 않습니다.
- G Suite 도메인 전체 위임 활성화를 활성화하고 해당 계정을 사용하여 JSON 키를 프라이빗 키로 사용하여 서비스 계정을 생성했습니다.
- 사용자 계정 이메일과 서비스 계정 이메일을 복사했습니다. 연결할 때 사용자 계정 Amazon Kendra 이메일을 관리자 계정 이메일로 입력하고 서비스 계정 이메일을 고객 이메일로 AWS Secrets Manager 시크릿에 입력합니다.

  > **Note**
  >
  > 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 계정에 관리자 SDK API 및 Google Drive API를 추가했습니다.
- 수퍼 관리자 역할을 사용하여 서비스 계정에 다음 권한을 추가(또는 수퍼 관리자 역할을 가진 사용자에게 추가를 요청)했습니다.
  - 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
- Google Drive 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

귀하의 AWS 계정란에 다음이 있는지 확인하십시오.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Google Drive 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 비밀번호가 없는 경우 Google 드라이브 데이터 소스를 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 비밀번호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Google 드라이브 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Google 드라이브 데이터 소스의 필수 세부정보를 제공해야 합니다. Google 드라이브를 아직 구성하지 않은 경우 Amazon Kendra 참조하십시오필수 조건.

Console

Google 드라이브에 Amazon Kendra 연결하려면

1. AWS 관리 콘솔에 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 소스 추가 페이지에서 Google Drive 커넥터 V1.0을 선택한 다음 커넥터 추가를 선택합니다.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

   e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. 인증 유형의 경우 - 기존 인증과 신규 인증 중에서 선택합니다. 기존 암호를 사용하기로 선택한 경우 암호 선택을 사용하여 암호를 선택합니다.

   b. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 옵션이 열립니다.

      - AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

        A. 보안 암호 이름 - 보안 암호의 이름. 비밀번호에는 접두사 AmazonKendra '-Google Drive-'가 자동으로 추가됩니다.

        B. 관리자 계정 이메일, 클라이언트 이메일, 프라이빗 키의 경우 - Google Drive 계정에서 생성하고 다운로드한 보안 인증 값을 입력합니다.

        C. 인증 저장을 선택합니다.

   c. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 만들어 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

d. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

   a. 사용자 계정 제외 - 인덱스에서 제외하려는 Google Drive 사용자. 최대 100개의 사용자 계정을 추가할 수 있습니다.

   b. 공유 드라이브 제외 - 인덱스에서 제외하려는 Google Drive 공유 드라이브. 최대 100개의 공유 드라이브를 추가할 수 있습니다.

   c. 파일 유형 드라이브 제외 - 인덱스에서 제외하려는 Google Drive 파일 유형. 선택한 MIME 유형을 편집하도록 선택할 수도 있습니다.

   d. 추가 구성 - 정규 표현식 패턴을 지정하여 특정 콘텐츠를 포함하거나 제외합니다. 최대 100개의 패턴을 추가할 수 있습니다.

   e. 빈도 - Amazon Kendra 가 데이터 소스와 동기화하는 빈도.

   f. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. GoogleDrive 필드 이름 및 추가 제안 필드 매핑의 경우 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

   b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

   c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Google 드라이브에 Amazon Kendra 연결하려면

GoogleDriveConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 비밀 Amazon 리소스 이름 (ARN) —Google 드라이브 계정의 인증 자격 증명이 포함된 Secrets Manager 비밀의 Amazon 리소스 이름 (ARN) 을 입력합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "clientAccount": "service account email",
    "adminAccount": "user account email"",
    "privateKey": "private key"
}
```

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 Google 드라이브 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Google Drive 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- 포함 및 제외 필터 - Amazon Kendra 는 기본적으로 Google Drive의 모든 문서를 인덱싱합니다. 공유 드라이브, 사용자 계정, 문서 MIME 유형, 파일에 특정 콘텐츠를 포함할지 제외할지 여부를 지정할 수 있습니다. 사용자 계정을 제외하도록 선택하면 해당 계정이 소유한 내 드라이브의 어떤 파일도 인덱싱되지 않습니다. 사용자와 공유한 파일은 파일 소유자도 제외되지 않는 한 인덱싱됩니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 필드 매핑 - Google Drive 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서의 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

자세히 알아보기

Google 드라이브 데이터 Amazon Kendra 소스와의 통합에 대해 자세히 알아보려면 다음을 참조하세요.

- Amazon Kendra Google 드라이브 커넥터 시작하기

Google Drive 커넥터 V2.0

Google Drive는 클라우드 기반 파일 스토리지 서비스입니다. Google 드라이브 데이터 원본의 공유 드라이브, 내 드라이브, 나와 공유한 항목 폴더에 저장된 문서 및 댓글을 인덱싱하는 데 사용할 Amazon Kendra 수 있습니다. Google Workspace 문서와 문서 유형에 나열된 문서를 인덱싱할 수 있습니다. 포함 및 제외 필터를 사용하여 파일 이름, 파일 유형, 파일 경로별로 콘텐츠를 인덱싱할 수도 있습니다.

> **Note**
>
> 구글 드라이브 커넥터 V1.0/구글 DriveConfiguration API에 대한 지원은 2023년에 종료될 예정입니다. Google 드라이브 커넥터 V2.0/ API로 마이그레이션하거나 이를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra Google 드라이브 데이터 소스 커넥터 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Google 드라이브 데이터 소스를 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Google 드라이브와 AWS 계정에서 다음과 같이 변경하세요.

Google Drive에서 다음 사항을 갖추었는지 확인하세요.

- 수퍼 관리자 역할에 의해 액세스 권한을 부여 받은 경우 또는 관리자 권한이 있는 사용자일 수 있습니다. 수퍼 관리자 역할을 통해 액세스 권한을 부여 받은 경우 본인에게 수퍼 관리자 역할이 필요하지 않습니다.
- 관리자 계정 이메일, 클라이언트 이메일(서비스 계정 이메일), 프라이빗 키를 포함하는 Google Drive 서비스 계정 연결 보안 인증을 구성했습니다. 서비스 계정 키 생성 및 삭제에 관한 Google Cloud 문서를 참조하세요.

  > **Note**
  >
  > 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- server-to-server인증을 위해 G Suite 도메인 전체 위임 활성화를 활성화하여 Google 클라우드 서비스 계정 (사용자 ID를 수임할 권한이 위임된 계정) 을 만든 다음 해당 계정을 사용하여 JSON 비공개 키를 생성했습니다.

  > **Note**
  >
  > 프라이빗 키는 서비스 계정을 만든 후에 생성해야 합니다.

- 사용자 계정에 관리자 SDK API 및 Google Drive API를 추가했습니다.

- 선택 사항: 클라이언트 ID, 클라이언트 보안 암호, 새로고침 토큰을 특정 사용자의 연결 보안 인증 정보로 포함하는 Google Drive OAuth 2.0 연결 보안 인증을 구성했습니다. 개별 계정 데이터를 크롤링하려면 이 정보가 필요합니다. OAuth 2.0을 사용하여 API에 액세스하는 방법에 대한 Google 설명서를 참조하세요.
- 수퍼 관리자 역할을 사용하여 서비스 계정에 다음 OAuth 범위를 추가(또는 수퍼 관리자 역할을 가진 사용자에게 추가를 요청)했습니다. Google Workspace 도메인의 모든 사용자에 대한 모든 문서 및 액세스 제어(ACL) 정보를 크롤링하려면 다음과 같은 API 범위가 필요합니다.
  - https://www.googleapis.com/auth/drive.readonly—모든 Google Drive 파일 확인 및 다운로드
  - https://www.googleapis.com/auth/drive.metadata.readonly—Google Drive에 있는 파일의 메타데이터 보기
  - https://www.googleapis.com/auth/admin.directory.group.readonly—그룹, 그룹 별칭, 구성원 정보 검색에만 해당하는 범위. 이는 ID 크롤러에 필요합니다. Amazon Kendra
  - https://www.googleapis.com/auth/admin.directory.user.readonly—사용자 또는 사용자 별칭을 검색에만 해당하는 범위. 이는 Amazon Kendra 아이덴티티 크롤러에 사용자를 나열하고 ACL을 설정하는 데 필요합니다.
  - https://www.googleapis.com/auth/cloud-platform—대용량 Google Drive 파일의 콘텐츠를 가져오기 위한 액세스 토큰의 생성 범위.
  - https://www.googleapis.com/auth/forms.body.readonly—Google Forms에서 데이터를 가져오는 범위.

  Forms API를 지원하려면 다음과 같은 추가 범위를 추가하세요.

  - https://www.googleapis.com/auth/forms.body.readonly
- Google Drive 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 AWS 계정있는지 확인하십시오.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Google Drive 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 비밀번호가 없는 경우 Google 드라이브 데이터 소스를 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 비밀번호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Google 드라이브 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Google 드라이브 데이터 소스의 필수 세부정보를 제공해야 합니다. Google 드라이브를 아직 구성하지 않은 경우 Amazon Kendra 참조하십시오필수 조건.

Console

Google 드라이브에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.

4. 데이터 원본 추가 페이지에서 Google 드라이브 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 'V2.0' 태그가 있는 Google 드라이브 커넥터를 선택하세요.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 색인을 기준으로 문서를 필터링할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

   b. 인증의 경우 - 사용 사례에 따라 Google 서비스 인증, OAuth 2.0 인증 중에서 선택합니다.

   c. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 Secrets Manager 암호를 만들어 Google 드라이브 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

      i. Google 서비스 계정을 선택한 경우 비밀번호의 이름, 서비스 계정 구성의 관리자 또는 '서비스 계정 사용자'의 이메일 ID (관리자 이메일), 서비스 계정의 이메일 ID (클라이언트 이메일), 서비스 계정에서 만든 비공개 키를 입력합니다.

         비밀번호를 저장하고 추가하세요.

      ii. OAuth 2.0 인증을 선택한 경우 OAuth 계정에서 만든 암호, 클라이언트 ID, 클라이언트 암호, 새로 고침 토큰의 이름을 입력합니다.

         비밀번호를 저장하고 추가하세요.

   d. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

e. (Google 서비스 계정 인증 사용자만 해당)

ID 크롤러 - ID 크롤러를 사용할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

f. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

g. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 콘텐츠 동기화 - 크롤링할 옵션이나 콘텐츠를 선택합니다. 내 드라이브 (개인 폴더), 공유 드라이브 (나와 공유된 폴더) 또는 둘 다를 크롤링하도록 선택할 수 있습니다. 파일 코멘트도 포함할 수 있습니다.

b. 추가 구성 - 선택 사항에서 다음과 같은 선택적 정보를 입력할 수도 있습니다.

i. 대상 사용자 - 크롤링하려는 문서에 특정 대상 고객을 추가합니다.

ii. 최대 파일 크기 —크롤링할 파일의 최대 크기 제한을 MB 단위로 설정합니다.

iii. 사용자 이메일 - 포함하거나 제외하려는 사용자 이메일을 추가합니다.

iv. 공유 드라이브 - 포함하거나 제외하려는 공유 드라이브 이름을 추가합니다.

v. MIME 유형 - 포함하거나 제외하려는 MIME 유형을 추가합니다.

vi. 엔티티 정규식 패턴 - 정규 표현식 패턴을 추가하여 지원되는 모든 엔티티의 특정 첨부 파일을 포함하거나 제외합니다. 최대 100개의 패턴을 추가할 수 있습니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가

크롤링되고 인덱싱됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

> **Important**
>
> Google Drive API는 영구 삭제된 파일에서 설명을 검색하는 것을 지원하지 않습니다. 휴지통에 저장된 파일의 설명은 복구할 수 있습니다. 파일이 삭제되면 커넥터는 색인에서 Amazon Kendra 댓글을 삭제합니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 실행 기록 동기화에서 데이터 소스를 동기화할 Amazon S3 때 자동 생성된 보고서를 저장하도록 선택합니다. 이는 데이터 원본을 동기화할 때 문제를 추적하는 데 유용합니다.

f. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 파일용 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

> **Note**
>
> Google Drive API는 사용자 지정 필드 생성을 지원하지 않습니다. Google Drive 커넥터에는 사용자 지정 필드 매핑을 사용할 수 없습니다.

b. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Google 드라이브에 Amazon Kendra 연결하려면

TemplateConfigurationAPI를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 GOOGLEDRIVEV2 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 인증 유형 - 서비스 계정 인증을 사용할지 OAuth 2.0 인증을 사용할지를 지정합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 Amazon Kendra 변경될 때 색인을 업데이트하는 방법을 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

  > **Important**
  >
  > Google Drive API는 영구 삭제된 파일에서 설명을 검색하는 것을 지원하지 않습니다. 휴지통에 저장된 파일의 설명은 복구할 수 있습니다. 파일이 삭제되면 커넥터는 색인에서 Amazon Kendra 댓글을 삭제합니다.

- 비밀 Amazon 리소스 이름 (ARN) —Google 드라이브 계정에서 만든 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 입력합니다. Google 서비스 계정 인증을 사용하는 경우 보안 암호는 다음 키가 있는 JSON 구조에 저장됩니다.

```
{
    "clientEmail": "user account email",
    "adminAccountEmail": "service account email",
    "privateKey": "private key"
}
```

OAuth 2.0 인증을 사용하는 경우 보안 암호는 다음 키를 사용하여 JSON 구조에 저장됩니다.

```
{
    "clientID": "OAuth client ID",
    "clientSecret": "client secret",
    "refreshToken": "refresh token"
}
```

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 Google 드라이브 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Google Drive 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 내 드라이브, 공유 드라이브, 댓글 —이러한 유형의 콘텐츠를 크롤링할지 여부를 지정할 수 있습니다.
- 포함 및 제외 필터 - 특정 사용자 계정, 공유 드라이브 및 MIME 유형을 포함할지 또는 제외할지 여부를 지정할 수 있습니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 액세스 제어 목록 (ACL) - ACL이 있고 액세스 제어에 사용하려는 경우 문서의 ACL 정보를 크롤링할지 여부를 지정합니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL

정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 필드 매핑 - Google Drive 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Google Drive 템플릿 스키마를 참조하세요.

참고

- Google Drive UI는 사용자 지정 필드 생성을 지원하지 않으므로 Google Drive 커넥터에는 사용자 지정 필드 매핑을 사용할 수 없습니다.
- Google Drive API는 영구 삭제된 파일에서 설명을 검색하는 것을 지원하지 않습니다. 휴지통에 저장된 파일을 제외하고 설명은 복구할 수 있습니다. 파일이 삭제되면 Amazon Kendra 커넥터는 색인에서 Amazon Kendra 주석을 삭제합니다.
- Google Drive API는 .docx 파일에 있는 설명을 반환하지 않습니다.

IBM DB2

IBM DB2는 IBM에서 개발한 관계형 데이터베이스 관리 시스템입니다. IBM DB2 사용자인 경우 Amazon Kendra 를 사용하여 IBM DB2 데이터 소스를 인덱싱할 수 있습니다. Amazon Kendra IBM DB2데이터 소스 커넥터는 DB2 11.5.7을 지원합니다.

Amazon Kendra 콘솔과 API를 사용하여 IBM DB2 데이터 원본에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra IBM DB2데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

IBM DB2데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 IBM DB2 및 AWS 계정에서 다음과 같이 변경하십시오.

IBM DB2에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
>
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다.
- IBM DB2 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- IBM DB2 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 IBM DB2 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

IBM DB2데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 IBM DB2 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 IBM DB2 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 IBM DB2

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 IBM DB2커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 IBM DB2커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. 소스에서 다음 정보를 입력합니다.

   b. 호스트 - 데이터베이스 호스트 이름을 입력합니다.

   c. 포트 - 데이터베이스 포트를 입력합니다.

   d. 인스턴스 - 데이터베이스 인스턴스를 입력합니다.

   e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.

   f. 인증에서 - 다음 정보를 입력합니다.

      - AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 IBM DB2 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

        A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-IBM DB2 -'가 자동으로 추가됩니다.

II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.

- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra IBM DB2

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 db2로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. IBM DB2 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. IBM DB2 Amazon Kendra자세한 내용은 IBM DB2 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - IBM DB2 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 IBM DB2 템플릿 스키마를 참조하세요.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Jira

Jira는 소프트웨어 개발, 제품 관리 및 버그 추적을 위한 프로젝트 관리 도구입니다. Jira 프로젝트, 이슈, 댓글, 첨부 파일, 작업 로그, 상태를 인덱싱하는 Amazon Kendra 데 사용할 수 있습니다.

Amazon Kendra 현재는 Jira Cloud만 지원합니다.

Amazon Kendra 콘솔 또는 API를 Amazon Kendra 사용하여 Jira 데이터 소스에 연결할 수 있습니다. JiraConfiguration 각각에 지원되는 기능 목록은 지원되는 기능 단원을 참조하세요.

Amazon Kendra Jira 데이터 소스 커넥터의 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Jira 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어

- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Jira 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Jira와 계정에서 다음과 같이 변경하세요. AWS

Jira에서 다음 사항을 갖추었는지 확인하세요.

- Jira ID (사용자 이름 또는 이메일) 및 Jira 자격 증명 (Jira API 토큰) 을 포함하는 API 토큰 인증 자격 증명을 구성했습니다. API 토큰 관리에 대한 Atlassian 설명서를 참조하세요.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- Jira 계정 설정에서 Jira 계정 URL을 기록했습니다. 예: *https://company.atlassian.net/*.
- Jira 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오 AWS 계정.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Jira 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Jira 데이터 소스를 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Jira 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Jira 데이터 소스의 필수 세부 정보를 제공해야 합니다. Jira를 위해 Amazon Kendra아직 구성하지 않은 경우 을 참조하십시오. 필수 조건

Console

Jira에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Jira 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Jira 커넥터를 선택합니다.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 문서에서 색인을 필터링할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. Jira 계정 URL —Jira 계정 URL을 입력합니다. 예: *https://company.atlassian.net/*.

b. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

c. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Jira Secrets Manager 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Jira-'가 자동으로 추가됩니다.

B. Jira ID의 경우 - Jira 사용자 이름 또는 이메일을 입력합니다.

C. 비밀번호/토큰의 경우 - Jira에 구성된 Jira API 토큰을 입력합니다.

ii. 비밀번호를 저장하고 추가하세요.

d. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

e. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

f. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

g. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 인덱싱할 Jira 프로젝트 선택 —모든 프로젝트 또는 특정 프로젝트를 크롤링하도록 선택합니다.

b. 추가 구성 —특정 상태와 이슈 유형을 지정합니다. 댓글, 첨부 파일, 작업 로그를 크롤링하도록 선택합니다. 정규 표현식 패턴을 사용하여 특정 콘텐츠를 포함하거나 제외할 수 있습니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Jira에 Amazon Kendra 연결하려면

JiraConfiguration API를 사용하여 다음을 지정해야 합니다.

- 데이터 소스 URL - Jira 계정 URL을 지정합니다. 예: *company.atlassian.net*.
- 비밀 Amazon 리소스 이름 (ARN) - Jira 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "jiraId": "Jira user name or email",
    "jiraCredential": "Jira API token"
}
```

- IAM 역할 - `CreateDataSource` Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 Jira 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Jira 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - 데이터 소스 구성의 일부로 `VpcConfiguration`를 지정합니다. VPC 사용을 위한 Amazon Kendra 구성을 참조하세요.
- 변경 로그 - 인덱스에서 문서를 Amazon Kendra 업데이트해야 하는지 여부를 결정하기 위해 Jira 데이터 소스 변경 로그 메커니즘을 사용해야 하는지 여부입니다.

> **Note**
>
> Amazon Kendra 가 모든 문서를 스캔하지 않도록 하려면 변경 로그를 사용하세요. 변경 로그가 크면 변경 로그를 처리하는 것보다 Jira 데이터 원본의 문서를 스캔하는 데 걸리는

시간이 더 Amazon Kendra 적을 수 있습니다. Jira 데이터 소스를 인덱스와 처음으로 동기화하는 경우 모든 문서가 스캔됩니다.

- 포함 및 제외 필터 - 특정 파일을 포함할지 또는 제외할지를 지정할 수 있습니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 댓글, 첨부 파일, 작업 로그 - 문제의 특정 댓글, 첨부 파일 및 작업 로그를 크롤링할지 여부를 지정할 수 있습니다.
- 프로젝트, 이슈, 상태 - 특정 프로젝트 ID, 이슈 유형 및 상태를 크롤링할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Jira 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

자세히 알아보기

Jira 데이터 Amazon Kendra 소스와의 통합에 대해 자세히 알아보려면 다음을 참조하십시오.

- Jira 클라우드 커넥터로 Jira 프로젝트를 지능적으로 검색하세요. Amazon Kendra

Microsoft Exchange

Microsoft Exchange는 메시징, 회의 및 파일 공유를 위한 엔터프라이즈 협업 도구입니다. Microsoft Exchange 사용자인 경우 Microsoft Exchange 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

Amazon Kendra 콘솔과 TemplateConfigurationAPI를 사용하여 Microsoft Exchange 데이터 원본에 연결할 Amazon Kendra 수 있습니다.

Amazon Kendra Microsoft Exchange 데이터 원본 커넥터 문제 해결에 대한 자세한 내용은 을 참조하십시오데이터 소스 문제 해결.

지원되는 기능

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Microsoft Exchange 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있으려면 먼저 Microsoft Exchange와 AWS 계정에서 이러한 변경을 수행하십시오.

Microsoft Exchange에서는 다음을 확인하세요.

- Office 365에서 Microsoft Exchange 계정을 만들었습니다.
- Microsoft 365 테넌트 ID를 기록했습니다. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
- Azure 포털에서 OAuth 애플리케이션을 구성하고 클라이언트 ID와 클라이언트 암호 또는 클라이언트 자격 증명을 기록했습니다. 자세한 내용은 Microsoft 자습서 및 등록된 앱 예제를 참조하십시오.

> **Note**
>
> Azure 포털에서 앱을 만들거나 등록할 때 비밀 ID는 실제 비밀 값을 나타냅니다. 암호 및 앱을 만들 때 실제 암호 값을 즉시 기록하거나 저장해야 합니다. Azure Portal에서 애플리케이

션 이름을 선택한 다음 인증서 및 비밀의 메뉴 옵션으로 이동하여 비밀번호에 액세스할 수 있습니다.
Azure Portal에서 애플리케이션 이름을 선택한 다음 개요 페이지로 이동하여 클라이언트 ID에 액세스할 수 있습니다. 애플리케이션 (클라이언트) ID는 클라이언트 ID입니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 커넥터 애플리케이션에 다음 권한을 추가했습니다.

Microsoft Graph	Office 365 Exchange Online
• Mail.Read(애플리케이션) • 메일. ReadBasic (애플리케이션) • 메일. ReadBasic. 전체 (애플리케이션) • Calendars.Read(애플리케이션) • User.Read.All(애플리케이션) • Contacts. Read(애플리케이션) • Notes.Read.All(애플리케이션) • Directory.Read.All(애플리케이션) • 뉴스. AccessAsUser4. 모두 (위임)	full_access_as_app(애플리케이션)

- Microsoft Exchange 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

귀하의 서류에 AWS 계정다음이 있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Microsoft Exchange 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Microsoft Exchange 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Microsoft Exchange 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Microsoft Exchange 데이터 원본의 필수 세부 정보를 제공해야 합니다. Microsoft Exchange를 위해 Amazon Kendra아직 구성하지 않은 경우 을 참조하십시오필수 조건.

Console

Microsoft Amazon Kendra Exchange에 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.

4. 데이터 원본 추가 페이지에서 Microsoft Exchange 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Microsoft Exchange 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. 테넌트 ID - Microsoft 365 테넌트 ID를 입력합니다. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
   b. 권한 부여 - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
   c. AWS Secrets Manager 암호 - 기존 암호를 선택하거나 새 Secrets Manager 암호를 만들어 Microsoft Exchange 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.
      i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.
         A. 보안 암호 이름 - 보안 암호의 이름. 접두사 'AmazonKendra- 마이크로소프트 익스체인지
         B. 클라이언트 ID의 경우 클라이언트 암호 - Microsoft Exchange에 구성된 인증 자격 증명을 Azure 포털에 입력합니다.
      ii. 암호를 저장하고 추가합니다.
   d. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

e. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

f. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 사용자 ID - 특정 이메일을 기준으로 콘텐츠를 필터링하려는 경우 사용자 이메일을 입력합니다.

b. 추가 구성 —크롤링하려는 콘텐츠 유형을 지정합니다.

- 엔티티 유형 - 캘린더 또는 연락처 콘텐츠를 크롤링하도록 선택할 수 있습니다. OneNotes
- 캘린더 크롤링 - 특정 날짜 사이에 콘텐츠를 크롤링하려면 시작 날짜와 종료 날짜를 입력합니다.
- 이메일 포함 - 크롤링하려는 특정 이메일을 필터링하려면 “받는 사람”, “보낸 사람”, 이메일 제목을 입력합니다.
- 공유 폴더 액세스 - Microsoft Exchange 데이터 원본의 액세스 제어를 위해 액세스 제어 목록을 크롤링할 수 있도록 선택합니다.
- 도메인용 정규식 —정규 표현식 패턴을 추가하여 특정 전자 메일 도메인을 포함하거나 제외합니다.
- 정규식 패턴 - 특정 페이지 및 자산을 포함하거나 제외하기 위한 정규 표현식 패턴을 추가합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.

- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

> **Note**
>
> Amazon Kendra Microsoft Exchange 데이터 원본 커넥터는 사용자 지정 필드 매핑을 지원하지 않습니다.

b. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Microsoft Amazon Kendra Exchange에 연결하려면

TemplateConfigurationAPI를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 `MSEXCHANGE` 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 `TEMPLATE` 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 테넌트 ID - Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.

- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - Microsoft Exchange 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공하십시오. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM 역할 - CreateDataSource Secrets Manager 암호에 액세스할 수 있는 권한을 IAM 역할에 제공하고 Microsoft Exchange 커넥터 및 에 필요한 공개 API를 호출하도록 호출할 RoleArn 시기를 지정합니다. Amazon Kendra자세한 내용은 Microsoft Exchange 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - CreateDataSource를 호출할 때 VpcConfiguration을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 특정 콘텐츠를 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함

필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- ACL (액세스 제어 목록) - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 ACL 정보를 크롤링할지 여부를 지정합니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Microsoft Exchange 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 다른 중요한 JSON 키 목록은 Microsoft Exchange 템플릿 스키마를 참조하십시오.

자세히 알아보기

Microsoft Exchange 데이터 Amazon Kendra 원본과의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- Amazon Kendra에 대한 Exchange 커넥터를 사용하여 Microsoft Exchange 콘텐츠를 인덱싱

마이크로소프트 OneDrive

OneDrive Microsoft는 콘텐츠를 저장, 공유 및 호스팅하는 데 사용할 수 있는 클라우드 기반 저장소 서비스입니다. OneDrive 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

Amazon Kendra 콘솔과 OneDriveConfigurationAPI를 Amazon Kendra 사용하여 OneDrive 데이터 소스에 연결할 수 있습니다.

Amazon Kendra OneDrive 커넥터에는 두 가지 버전이 있습니다. 각 버전에 지원되는 기능은 다음과 같습니다.

마이크로소프트 OneDrive 커넥터 V1.0/ API OneDriveConfiguration

- 필드 매핑
- 포함/제외 필터

마이크로소프트 OneDrive 커넥터 V2.0/ API TemplateConfiguration

- 사용자 컨텍스트 필터링
- 사용자 아이덴티티 크롤러
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

> **Note**
>
> OneDrive 커넥터 OneDriveConfiguration V1.0/API에 대한 지원은 2023년 6월에 종료될 예정입니다. OneDrive 커넥터 V2.0/ TemplateConfiguration API를 사용하는 것이 좋습니다.

Amazon Kendra OneDrive 데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 마이크로소프트 OneDrive 커넥터 V1.0
- 마이크로소프트 OneDrive 커넥터 V2.0
- 자세히 알아보기

마이크로소프트 OneDrive 커넥터 V1.0

OneDrive Microsoft는 콘텐츠를 저장, 공유 및 호스팅하는 데 사용할 수 있는 클라우드 기반 저장소 서비스입니다. Microsoft OneDrive 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

> **Note**
>
> OneDrive 커넥터 V1.0/마이크로소프트 OneDrive API에 대한 지원은 2023년 6월에 종료될 예정입니다. OneDrive 커넥터 V2.0/ API를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra OneDrive 데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

지원되는 기능

- 필드 매핑
- 포함/제외 필터

필수 조건

를 사용하여 OneDrive 데이터 원본을 Amazon Kendra 인덱싱하려면 먼저 OneDrive 및 AWS 계정에서 다음과 같이 변경하십시오.

Azure Active Directory(AD)에서 다음을 확인하세요.

- Azure Active Directory(AD) 애플리케이션을 만들었습니다.
- AD 애플리케이션 ID를 사용하여 AD 사이트에 애플리케이션의 보안 암호 키를 등록했습니다. 보안 암호 키에는 애플리케이션 ID와 보안 암호 키가 포함되어야 합니다.
- 조직의 AD 도메인을 복사했습니다.
- Microsoft Graph 옵션에서 AD 애플리케이션에 다음과 같은 애플리케이션 권한을 추가했습니다.
  - 모든 사이트 모음에서 파일 읽기(File.Read.All)
  - 모든 사용자의 전체 프로필 읽기(User.Read.All)
  - 디렉터리 데이터 읽기(Directory.Read.All)
  - 모든 그룹 읽기(Group.Read.All)
  - 모든 사이트 모음에서 항목 읽기(Site.Read.All)
- 문서를 인덱싱해야 하는 사용자의 목록을 복사했습니다. 사용자 이름 목록을 제공하도록 선택하거나 Amazon S3에 저장된 파일에 사용자 이름을 제공할 수 있습니다. 데이터 소스를 생성하면 다음을 수행할 수 있습니다.
  - 사용자 목록을 수정합니다.
  - 사용자 목록에서 Amazon S3 버킷에 저장된 목록으로 변경합니다.
  - 사용자 목록의 Amazon S3 버킷 위치를 변경합니다. 버킷 위치를 변경하는 경우 데이터 원본이 버킷에 액세스할 수 있도록 데이터 원본의 IAM 역할도 업데이트해야 합니다.

> **Note**
>
> 사용자 이름 목록을 버킷에 저장하는 경우 데이터 원본의 IAM 정책은 Amazon S3 버킷에 대한 액세스와 버킷이 암호화된 키 (있는 경우) 에 대한 액세스를 제공해야 합니다.

- 동일한 인덱스에 사용하려는 다른 데이터 소스에서 각 문서가 고유한지 확인했습니다. OneDrive 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하세요. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- OneDrive 인증 자격 증명을 AWS Secrets Manager 비밀에 저장하고 API를 사용하는 경우 비밀의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 OneDrive 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

OneDrive 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 OneDrive 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 OneDrive 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 OneDrive

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 OneDrive 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 OneDrive 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. OneDrive 테넌트 ID - 프로토콜 없이 OneDrive 테넌트 ID를 입력합니다.
   b. 인증 유형 - 신규 인증과 기존 인증 중에서 선택합니다.
   c. i. 기존을 선택하는 경우 보안 암호 선택에서 기존 보안 암호를 선택합니다.

ii. 신규를 선택한 경우 새 AWS Secrets Manager 보안 암호 섹션에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-OneDrive -'가 자동으로 추가됩니다.

B. 애플리케이션 ID 및 애플리케이션 비밀번호의 경우 - OneDrive 계정의 인증 자격 증명 값을 입력한 다음 인증 저장을 선택합니다.

d. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

e. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 사용 사례에 따라 목록 파일과 이름 목록 중에서 선택합니다.

i. 목록 파일을 선택하는 경우 다음 정보를 입력합니다.

- 위치 선택 - Amazon S3 버킷 경로를 입력합니다.

  사용자 목록 파일 추가 Amazon S3 —버킷에 사용자 목록 파일을 Amazon S3 추가하려면 선택합니다.

  사용자 로컬 그룹 매핑 - 로컬 그룹 매핑을 사용하여 콘텐츠를 필터링하려면 선택합니다.

ii. 이름 목록을 선택하는 경우 다음 정보를 입력합니다.

- 사용자 이름 - 인덱싱할 사용자 드라이브를 최대 10개까지 입력합니다. 10명 넘는 사용자를 추가하려면 이름이 포함된 파일을 만듭니다.

  다른 사용자 추가 - 사용자를 더 추가하려면 선택합니다.

  사용자 로컬 그룹 매핑 - 로컬 그룹 매핑을 사용하여 콘텐츠를 필터링하려면 선택합니다.

b. 추가 구성의 경우 - 정규 표현식 패턴을 추가하여 특정 파일을 포함하거나 제외합니다. 최대 100개의 패턴을 추가할 수 있습니다.

c. 동기화 실행 일정에서 빈도에 대해 —데이터 원본과 Amazon Kendra 동기화할 빈도를 선택합니다.

d. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 기본 데이터 원본 필드 및 추가 제안 필드 매핑의 경우 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면: Amazon Kendra OneDrive

OneDriveConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 테넌트 ID - 조직의 Azure Active Directory 도메인을 지정합니다.
- OneDrive 사용자 - 문서를 색인화해야 하는 사용자 계정 목록을 지정합니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. OneDrive 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "username": "OAuth client ID",
    "password": "client secret"
}
```

- IAM 역할 - `CreateDataSource` Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. OneDrive Amazon Kendra자세한 내용은 OneDrive 데이터 원본의IAM 역할을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- 포함 및 제외 필터 - 특정 문서를 포함할지 또는 제외할지 여부를 지정합니다.

  > **Note**
  >
  > 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 필드 매핑 - OneDrive 데이터 원본 필드를 인덱스 필드에 매핑하도록 Amazon Kendra 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

  > **Note**
  >
  > 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서의 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

마이크로소프트 OneDrive 커넥터 V2.0

OneDrive Microsoft는 콘텐츠를 저장, 공유 및 호스팅하는 데 사용할 수 있는 클라우드 기반 저장소 서비스입니다. OneDrive데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

Amazon Kendra 콘솔과 OneDriveConfigurationAPI를 Amazon Kendra 사용하여 OneDrive 데이터 소스에 연결할 수 있습니다.

> **Note**
>
> OneDrive 커넥터 OneDriveConfiguration V1.0/API에 대한 지원은 2023년 6월에 종료될 예정입니다. OneDrive 커넥터 V2.0/ TemplateConfiguration API를 사용하는 것이 좋습니다. 버전 2.0은 추가 ACL 및 ID 크롤러 기능을 제공합니다.

Amazon Kendra OneDrive 데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침

지원되는 기능

Amazon Kendra OneDrive 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

OneDrive 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 OneDrive 및 AWS 계정에서 다음과 같이 변경하십시오.

OneDrive에서 다음이 있는지 확인하세요.

- Office 365에서 OneDrive 계정을 만들었습니다.
- Microsoft 365 테넌트 ID를 기록했습니다. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
- Azure Portal에서 OAuth 응용 프로그램을 만들고 암호 인증에 사용되는 클라이언트 ID와 클라이언트 암호 또는 클라이언트 자격 증명을 기록했습니다. AWS Secrets Manager 자세한 내용은 Microsoft 자습서 및 등록된 앱 예제를 참조하십시오.

> **Note**
>
> Azure 포털에서 앱을 만들거나 등록할 때 비밀 ID는 실제 비밀 값을 나타냅니다. 암호 및 앱을 만들 때 실제 암호 값을 즉시 기록하거나 저장해야 합니다. Azure Portal에서 애플리케이션 이름을 선택한 다음 인증서 및 비밀의 메뉴 옵션으로 이동하여 비밀번호에 액세스할 수 있습니다.

Azure Portal에서 애플리케이션 이름을 선택한 다음 개요 페이지로 이동하여 클라이언트 ID에 액세스할 수 있습니다. 애플리케이션 (클라이언트) ID는 클라이언트 ID입니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- AD 애플리케이션 ID를 사용하여 AD 사이트에 애플리케이션의 보안 암호 키를 등록했습니다. 보안 암호 키에는 애플리케이션 ID와 보안 암호 키가 포함되어야 합니다.
- 조직의 AD 도메인을 복사했습니다.
- Microsoft Graph 옵션에서 AD 애플리케이션에 다음과 같은 권한을 추가했습니다.
  - 모든 사이트 모음에서 파일 읽기(File.Read.All)
  - 모든 사용자의 전체 프로필 읽기(User.Read.All)
  - 모든 그룹 읽기(Group.Read.All)
  - 모든 참고 읽기(Notes.Read.All)
- 문서를 인덱싱해야 하는 사용자의 목록을 복사했습니다. 사용자 이름 목록을 제공하도록 선택하거나 Amazon S3에 저장된 파일에 사용자 이름을 제공할 수 있습니다. 데이터 소스를 생성하면 다음을 수행할 수 있습니다.
  - 사용자 목록을 수정합니다.
  - 사용자 목록에서 Amazon S3 버킷에 저장된 목록으로 변경합니다.
  - 사용자 목록의 Amazon S3 버킷 위치를 변경합니다. 버킷 위치를 변경하는 경우 데이터 원본이 버킷에 액세스할 수 있도록 데이터 원본의 IAM 역할도 업데이트해야 합니다.

> **Note**
>
> 사용자 이름 목록을 버킷에 저장하는 경우 데이터 원본의 IAM 정책은 Amazon S3 버킷에 대한 액세스와 버킷이 암호화된 키 (있는 경우) 에 대한 액세스를 제공해야 합니다.
> OneDrive 커넥터는 Onedrive 사용자 속성에 있는 연락처 정보의 전자 메일을 사용합니다. 데이터를 크롤링하려는 사용자의 연락처 정보 페이지에 이메일 필드가 구성되어 있는지 확인하세요. 새 사용자의 경우 이 필드는 비어 있을 수 있습니다.

AWS 계정에 다음이 있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.
- OneDrive 인증 자격 증명을 AWS Secrets Manager 비밀에 저장하고 API를 사용하는 경우 비밀의 ARN을 기록했습니다.

기존 IAM 역할이나 암호가 없는 경우 OneDrive 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

OneDrive 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 OneDrive 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 OneDrive 을 Amazon Kendra참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 OneDrive

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 OneDrive 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 OneDrive 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. OneDrive 테넌트 ID - 프로토콜 없이 OneDrive 테넌트 ID를 입력합니다.

b. 권한 부여 - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록(ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

c. 인증 - 신규 인증과 기존 인증 중에서 선택합니다.

d. i. 기존을 선택하는 경우 보안 암호 선택에서 기존 보안 암호를 선택합니다.

ii. 신규를 선택한 경우 새 AWS Secrets Manager 보안 암호 섹션에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-OneDrive -'가 자동으로 추가됩니다.

B. 클라이언트 ID 및 클라이언트 암호의 경우 - 클라이언트 ID와 클라이언트 암호를 입력합니다.

e. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

f. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

g. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

h. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

8. a. 동기화 범위의 경우 - 인덱싱할 사용자 OneDrive 데이터를 선택합니다. 최대 10명의 사용자를 수동으로 추가할 수 있습니다.

   b. 추가 구성의 경우 - 정규 표현식 패턴을 추가하여 특정 콘텐츠를 포함하거나 제외합니다. 최대 100개의 패턴을 추가할 수 있습니다.

   c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

      - 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
      - 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
      - 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

   d. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

   e. 다음을 선택합니다.

9. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

   b. 다음을 선택합니다.

10. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra 연결하려면 OneDrive

TemplateConfigurationAPI를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 `ONEDRIVEV2` 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 `TEMPLATE` 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 테넌트 ID - Microsoft 365 테넌트 ID를 지정합니다. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - `FORCED_FULL_CRAWL`모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - `FULL_CRAWL`데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - `CHANGE_LOG`데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. OneDrive

  OAuth 2.0 인증을 사용하는 경우 보안 암호는 다음 키를 사용하여 JSON 구조에 저장됩니다.

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. OneDrive Amazon Kendra자세한 내용은 OneDrive 데이터 원본의IAM 역할을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 특정 파일, OneNote 섹션 및 페이지를 포함할지 또는 제외할지 여부를 지정할 수 있습니다. OneNote

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 필드 매핑 - 커넥터의 기본 제공 또는 공통 인덱스 필드만 매핑할 수 있습니다. Amazon Kendra OneDrive API 제한으로 인해 OneDrive 커넥터에는 사용자 지정 필드 매핑을 사용할 수 없습니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

구성해야 할 기타 중요한 JSON 키 목록은 OneDrive 템플릿 스키마를 참조하십시오.

자세히 알아보기

OneDrive 데이터 Amazon Kendra 소스와의 통합에 대해 자세히 알아보려면 다음을 참조하십시오.

- 에 대한 업데이트된 Microsoft OneDrive 커넥터 (V2) 를 Amazon Kendra발표합니다.

마이크로소프트 SharePoint

SharePoint 웹 콘텐츠를 사용자 지정하고 페이지, 사이트, 문서 라이브러리 및 목록을 만드는 데 사용할 수 있는 공동 웹 사이트 구축 서비스입니다. SharePoint 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

Amazon Kendra 현재 SharePoint 온라인 및 SharePoint 서버 (버전 2013, 2016, 2019 및 서브스크립션 에디션) 를 지원합니다.

Amazon Kendra 콘솔, TemplateConfigurationAPI 또는 API를 Amazon Kendra 사용하여 SharePoint 데이터 원본에 연결할 수 있습니다. SharePointConfiguration

Amazon Kendra SharePoint 커넥터에는 두 가지 버전이 있습니다. 각 버전에 지원되는 기능은 다음과 같습니다.

SharePoint 커넥터 V1.0/ API SharePointConfiguration

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 변경 로그
- Virtual Private Cloud(VPC)

SharePoint 커넥터 V2.0/ API TemplateConfiguration

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

> **Note**
>
> SharePoint 커넥터 SharePointConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. 커넥터 V2.0/ API로 마이그레이션하거나 SharePoint 커넥터 V2.0/API를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra SharePoint 데이터 소스 커넥터 문제 해결에 대한 내용은 을 참조하십시오. 데이터 소스 문제 해결

주제

- SharePoint 커넥터 V1.0
- SharePoint 커넥터 V2.0

SharePoint 커넥터 V1.0

SharePoint 웹 콘텐츠를 사용자 지정하고 페이지, 사이트, 문서 라이브러리 및 목록을 만드는 데 사용할 수 있는 공동 웹 사이트 구축 서비스입니다. SharePoint 사용자인 경우 SharePoint 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

> **Note**
>
> SharePoint 커넥터 SharePointConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. 커넥터 V2.0/ API로 마이그레이션하거나 SharePoint 커넥터 V2.0/API를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra SharePoint 데이터 소스 커넥터 문제 해결에 대한 내용은 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 변경 로그
- Virtual Private Cloud(VPC)

필수 조건

SharePoint 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 SharePoint 및 AWS 계정에서 다음과 같이 변경하십시오.

인증 자격 증명을 제공해야 하며, 이 자격 증명은 AWS Secrets Manager 비밀에 안전하게 보관됩니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

에서 SharePoint 다음을 갖추고 있는지 확인하세요.

- 색인을 생성하려는 SharePoint 사이트의 URL을 기록해 두었습니다.
- SharePoint 온라인의 경우:
  - 사이트 관리 권한이 있는 사용자 이름과 암호가 포함된 기본 보안 인증 정보를 기록해 두었습니다.
  - 선택 사항: 사용자 이름, 암호, 클라이언트 ID 및 클라이언트 보안 암호를 포함한 OAuth 2.0 보안 인증을 생성했습니다.
  - 관리자 사용자를 사용하여 Azure 포털에서 보안 기본값을 비활성화했습니다. Azure Portal에서 보안 기본 설정을 관리하는 방법에 대한 자세한 내용은 보안 기본값을 활성화/비활성화하는 방법에 대한 Microsoft 설명서를 참조하세요.
- SharePoint 서버의 경우:
  - SharePoint 서버 도메인 이름 (액티브 디렉터리의 NetBIOS 이름) 을 기록해 두었습니다. SharePoint 기본 인증 사용자 이름 및 암호와 함께 이 이름을 사용하여 SharePoint 서버를 연결할 Amazon Kendra수 있습니다.

> **Note**
>
> SharePoint 서버를 사용하고 사용자 컨텍스트에 따른 필터링을 위해 ACL (액세스 제어 목록) 을 이메일 형식으로 변환해야 하는 경우 LDAP 서버 URL 및 LDAP 검색 기준을 제공하십시오. 또는 디렉터리 도메인 재정의를 사용할 수도 있습니다. LDAP 서버 URL은 전체 도메인 이름과 포트 번호입니다(예: ldap: //example.com:389). LDAP 검색 기반은 도메인 컨트롤러 'example'과 'com'입니다. 디렉터리 도메인 재정의를 사용하면 LDAP 서버 URL 및 LDAP 검색 기반을 사용하는 대신 이메일 도메인을 사용할 수 있습니다. 예를 들어 username@example.com의 이메일 도메인은 'example.com'입니다. 도메인 확인이 염려되지 않고 단순히 이메일 도메인만 사용하려는 경우 이 재정의를 사용할 수 있습니다.

- 계정에 다음 권한을 추가했습니다. SharePoint

  SharePoint 목록의 경우

  - 항목 열기 - 서버측 파일 처리기를 사용하여 문서 소스를 확인합니다.
  - 애플리케이션 페이지 보기 - 양식, 보기 및 애플리케이션 페이지를 볼 수 있습니다. 목록을 열거합니다.
  - 항목 보기 - 목록의 항목과 문서 라이브러리의 문서를 볼 수 있습니다.
  - 버전 보기 - 목록 항목 또는 문서의 이전 버전을 볼 수 있습니다.

  SharePoint 웹사이트용

  - 디렉터리 찾아보기 - Designer 및 Web DAV 인터페이스를 사용하여 웹 사이트의 파일 및 폴더를 열거합니다. SharePoint
  - 사용자 정보 찾아보기 - 웹 사이트 사용자에 대한 정보를 볼 수 있습니다.
  - 권한 열거 - 웹 사이트, 목록, 폴더, 문서 또는 목록 항목에 대한 권한을 열거합니다.
  - 열기 - 웹 사이트, 목록 또는 폴더를 열어 컨테이너 내의 항목에 접근합니다.
  - 클라이언트 통합 기능 사용 - SOAP, WebDAV, 클라이언트 개체 모델 또는 Designer 인터페이스를 사용하여 웹 사이트에 액세스합니다. SharePoint
  - 원격 인터페이스 사용 - 클라이언트 애플리케이션을 시작하는 기능을 사용합니다.
  - 페이지 보기 - 웹 사이트에서 페이지를 봅니다.
- 동일한 색인에 사용하려는 다른 데이터 소스에서 각 문서가 고유한지 확인했습니다. SharePoint 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하세요. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형 및 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- SharePoint 인증 자격 증명을 AWS Secrets Manager 비밀에 저장하고 API를 사용하는 경우 비밀의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 SharePoint 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

SharePoint 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 SharePoint 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 SharePoint 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 SharePoint

1. AWS 관리 콘솔에 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 SharePoint 커넥터 v1.0을 선택한 다음 데이터 원본 추가를 선택합니다.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. 호스팅 방법의 경우 SharePoint —온라인과 SharePoint 서버 중에서 선택하세요.

      i. SharePoint온라인의 경우 - 저장소에 해당하는 사이트 URL을 입력합니다. SharePoint

      ii. SharePoint서버의 경우 - SharePoint 버전을 선택하고 SharePoint 리포지토리별 사이트 URL을 입력한 다음 SSL 인증서 위치의 Amazon S3 경로를 입력합니다.

   b. (SharePoint 서버만 해당) 웹 프록시용 - 내부 인스턴스의 호스트 이름과 포트 번호를 입력합니다. SharePoint 포트 번호는 0에서 65535 사이의 숫자 값이어야 합니다.

   c. 인증의 경우 - 사용 사례에 따라 다음 옵션 중에서 선택하세요.

      i. SharePoint 온라인의 경우 - 기본 인증과 OAuth 2.0 인증 중에서 선택합니다.

      ii. SharePoint 서버의 경우 - 없음, LDAP, 수동 중에서 선택합니다.

   d. AWS Secrets Manager 비밀번호의 경우 - 기존 비밀번호를 선택하거나 인증 자격 증명을 저장할 새 비밀번호를 Secrets Manager 생성합니다. SharePoint 새 암호를 만들기로 선택

하면 AWS Secrets Manager 보안 암호 창이 열립니다. 보안 암호 이름을 입력해야 합니다. 암호 이름에 접두사 AmazonKendra '- SharePoint -'가 자동으로 추가됩니다.

e. AWS Secrets Manager 보안 암호 생성 창에 다음 기타 정보를 입력합니다.

i. 사용 사례에 따라 다음 SharePoint 클라우드 인증 옵션 중에서 선택하십시오.

A. 기본 인증 - SharePoint 계정 사용자 이름을 사용자 이름으로, SharePoint 계정 비밀번호를 비밀번호로 입력합니다.

B. OAuth 2.0 인증 - SharePoint 계정 사용자 이름을 사용자 이름으로, SharePoint 계정 암호를 암호로, 자동 생성된 고유 SharePoint ID를 클라이언트 ID로, SharePoint 그리고 Amazon Kendra 양쪽에서 사용하는 공유 암호 문자열을 클라이언트 암호로 입력합니다.

ii. 사용 사례에 따라 다음 SharePoint 서버 인증 옵션 중에서 선택하십시오.

A. 없음 - SharePoint 계정 사용자 이름을 사용자 이름으로, SharePoint 계정 암호를 암호로, 서버 도메인 이름을 입력합니다.

B. LDAP ***- SharePoint ## ### ### ### ####, SharePoint ## ### ###, LDAP ## ##### (#### # ## ## ##, #: ldap: // example.com:389), LDAP ## ## (#: dc=example, dc=com) # #### #.***

C. 수동 - SharePoint 계정 사용자 이름을 사용자 이름으로, SharePoint 계정 암호를 암호로, 이메일 도메인 재정의 (디렉터리 사용자 또는 그룹의 이메일 도메인) 를 입력합니다.

iii. 저장을 선택합니다.

f. Virtual Private Cloud(VPC) - 서브넷과 VPC 보안 그룹도 추가해야 합니다.

> **Note**
>
> 서버를 사용하는 경우 VPC를 사용해야 SharePoint 합니다. Amazon VPC 다른 SharePoint 버전의 경우 선택 사항입니다.

g. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

h. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

   a. 변경 로그 사용 - 모든 파일을 동기화하는 대신 인덱스를 업데이트하려면 선택합니다.

   b. 첨부 파일 크롤링 - 첨부 파일을 크롤링하려면 선택합니다.

   c. 로컬 그룹 매핑 사용 - 문서가 제대로 필터링되었는지 확인하려면 선택합니다.

   d. 추가 구성 - 정규 표현식 패턴을 추가하여 특정 파일을 포함하거나 제외합니다. 최대 100개의 패턴을 추가할 수 있습니다.

   e. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

   f. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. Amazon Kendra 기본 필드 매핑 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

   b. 사용자 지정 필드 매핑 - 사용자 지정 데이터 소스 필드를 추가하여 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

   c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra SharePoint

SharePointConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- SharePoint버전 - 구성할 SharePoint 때 사용하는 SharePoint 버전을 지정합니다. 이 경우는 SharePoint 서버 2013, 서버 2016, SharePoint SharePoint 서버 2019 또는 SharePoint 온라인 중 무엇을 사용하든 상관 없습니다.
- 비밀 Amazon 리소스 이름 (ARN) - SharePoint 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 입력합니다. 암호는 JSON 구조로 저장됩니다.

  SharePoint 온라인 기본 인증의 경우 시크릿에 포함되어야 하는 최소 JSON 구조는 다음과 같습니다.

```
{
    "userName": "user name",
    "password": "password"
}
```

  SharePoint 온라인 OAuth 2.0 인증의 경우 시크릿에 포함되어야 하는 최소 JSON 구조는 다음과 같습니다.

```
{
    "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"
}
```

  SharePoint 서버 기본 인증의 경우 시크릿에 포함되어야 하는 최소 JSON 구조는 다음과 같습니다.

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name"
}
```

  SharePoint 서버 LDAP 인증 (사용자 컨텍스트에 따른 필터링을 위해 액세스 제어 목록 (ACL) 을 이메일 형식으로 변환해야 하는 경우 암호에 LDAP 서버 URL 및 LDAP 검색 기준을 포함할 수 있음) 의 경우 암호에 포함해야 하는 최소 JSON 구조는 다음과 같습니다.

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name"
    "ldapServerUrl": "ldap://example.com:389",
    "ldapSearchBase": "dc=example,dc=com"
}
```

SharePoint 서버 수동 인증의 경우 시크릿에 포함되어야 하는 최소 JSON 구조는 다음과 같습니다.

```
{
    "userName": "user name",
    "password": "password",
    "domain": "server domain name",
    "emailDomainOverride": "example.com"
}
```

- IAM 역할 - `CreateDataSource` Secrets Manager 비밀번호에 액세스할 수 있는 권한을 IAM 역할에 제공하고 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. SharePoint Amazon Kendra자세한 내용은 SharePoint 데이터 원본의IAM 역할을 참조하십시오.
- Amazon VPC SharePoint —서버를 사용하는 경우 데이터 원본 구성의 `VpcConfiguration` 일부로 지정하십시오. VPC 사용을 Amazon Kendra 위한 구성을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- 웹 프록시 - 웹 프록시를 통해 SharePoint 사이트 URL에 연결할지 여부. 이 옵션은 서버에만 사용할 SharePoint 수 있습니다.
- 목록 색인 - 첨부 파일의 내용을 SharePoint 목록 항목에 Amazon Kendra 인덱싱해야 하는지 여부.
- 변경 로그 - SharePoint 데이터 소스 변경 로그 메커니즘을 사용하여 색인에서 문서를 업데이트해야 하는지 여부를 Amazon Kendra 결정할지 여부입니다.

> **Note**
>
> Amazon Kendra 가 모든 문서를 스캔하지 않도록 하려면 변경 로그를 사용하세요. 변경 로그가 크면 SharePoint 데이터 원본의 문서를 스캔하는 데 걸리는 시간이 변경 로그를

처리하는 시간보다 Amazon Kendra 적을 수 있습니다. SharePoint데이터 원본을 색인과 처음으로 동기화하는 경우 모든 문서가 스캔됩니다.

- 포함 및 제외 필터 - 특정 콘텐츠를 포함할지 또는 제외할지 여부를 지정할 수 있습니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 필드 매핑 — SharePoint 데이터 원본 필드를 인덱스 필드에 매핑하도록 선택합니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

자세히 알아보기

데이터 Amazon Kendra 원본과의 통합에 대한 자세한 내용은 다음을 참조하십시오. SharePoint

- Amazon Kendra SharePoint 온라인 커넥터 시작하기

SharePoint 커넥터 V2.0

SharePoint 웹 콘텐츠를 사용자 지정하고 페이지, 사이트, 문서 라이브러리 및 목록을 만드는 데 사용할 수 있는 공동 웹 사이트 구축 서비스입니다. SharePoint 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

Amazon Kendra 현재 SharePoint 온라인 및 SharePoint 서버 (2013년, 2016년, 2019년 및 서브스크립션 에디션) 를 지원합니다.

> **Note**
>
> SharePoint 커넥터 SharePointConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. 커넥터 V2.0/ API로 마이그레이션하거나 SharePoint 커넥터 V2.0/API를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra SharePoint 데이터 소스 커넥터 문제 해결에 대한 내용은 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

Amazon Kendra SharePoint 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

SharePoint 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 SharePoint 및 AWS 계정에서 다음과 같이 변경하십시오.

인증 자격 증명을 제공해야 하며, 이 자격 증명은 AWS Secrets Manager 비밀에 안전하게 보관됩니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

SharePoint Online에서 다음 사항을 갖추었는지 확인하세요.

- SharePoint 인스턴스 URL을 복사했습니다. 입력하는 호스트 URL의 형식은 *https://yourdomain.sharepoint.com/sites/mysite*입니다. URL은 `https`로 시작하고 `sharepoint.com`을 포함해야 합니다.
- SharePoint 인스턴스 URL의 도메인 이름을 복사했습니다.
- SharePointOnline에 연결할 수 있는 사이트 관리자 권한이 있는 사용자 이름과 암호가 들어 있는 기본 인증 자격 증명을 기록해 두었습니다.
- 관리자 사용자를 사용하여 Azure 포털에서 보안 기본값을 비활성화했습니다. Azure Portal에서 보안 기본 설정을 관리하는 방법에 대한 자세한 내용은 보안 기본값을 활성화/비활성화하는 방법에 대한 Microsoft 설명서를 참조하세요.
- 콘텐츠의 Amazon Kendra 크롤링이 차단되지 않도록 SharePoint 계정에서 다단계 인증 (MFA) 을 비활성화했습니다. SharePoint
- 기본 인증 이외의 인증 유형을 사용하는 경우: 인스턴스의 테넌트 ID를 복사했습니다. SharePoint 테넌트 ID를 찾는 방법에 대한 자세한 내용은 Microsoft 365 테넌트 ID 찾기를 참조하세요.
- Microsoft Entra를 사용하여 클라우드 사용자 인증으로 마이그레이션해야 하는 경우 클라우드 인증에 대한 Microsoft 설명서를 참조하십시오.
- OAuth 2.0 인증 및 OAuth 2.0 새로 고침 토큰 인증의 경우: SharePoint Online에 연결하는 데 사용하는 사용자 이름과 암호, Azure AD에 등록한 후 생성된 클라이언트 ID 및 클라이언트 암호가 포함된 기본 인증 자격 증명을 기록해 두었습니다. SharePoint
  - ACL을 사용하지 않는 경우 다음 권한을 추가하세요.

Microsoft Graph	SharePoint
• Note.Read.All (응용 프로그램) - 모든 노트북을 읽습니다. OneNote • Sites.Read.All (애플리케이션) - 모든 사이트 모음의 항목 읽기	• AllSites.Read (위임) - 모든 사이트 모음의 항목 읽기

> **Note**
>
> Note.Read.All 및 Sites.Read.All은 문서를 크롤링하려는 경우에만 필요합니다. OneNote 특정 사이트를 크롤링하려는 경우 도메인에서 사용 가능한 모든 사이트가 아닌 특정 사이트로 권한을 제한할 수 있습니다. 사이트. 선택된 (응용 프로그램) 권한을 구성합니다. 이 API 권한을 사용하려면 Microsoft Graph API를 통해 모든 사이트에 대한 액세스 권한을 명시적으로 설정해야 합니다. 자세한 내용은 사이트에 대한 Microsoft 블로그를 참조하십시오. 선택된 권한.

- ACL을 사용하는 경우 다음 권한을 추가하세요.

Microsoft Graph	SharePoint
• Group.Member.Read.All (애플리케이션) - 모든 그룹 멤버십 읽기 • Notes.Read.All (애플리케이션) - 모든 노트북을 읽을 수 있습니다. OneNote • 사이트. FullControl.모두 (위임) - 문서의 ACL을 검색하는 데 필요합니다. • Sites.Read.All (애플리케이션) - 모든 사이트 모음의 항목 읽기 • User.Read.All (애플리케이션) - 모든 사용자의 전체 프로필 읽기	• AllSites.Read (위임) - 모든 사이트 모음의 항목 읽기

> **Note**
>
> GroupMember.Read.All 및 User.Read.All은 ID 크롤러가 활성화된 경우에만 필요합니다. 특정 사이트를 크롤링하려는 경우 도메인에서 사용 가능한 모든 사이트가 아닌 특정 사이트로 권한을 제한할 수 있습니다. 사이트. 선택된 (응용 프로그램) 권한을 구성합니다. 이 API 권한을 사용하려면 Microsoft Graph API를 통해 모든 사이트에 대한 액세스 권한을 명시적으로 설정해야 합니다. 자세한 내용은 사이트에 대한 Microsoft 블로그를 참조하십시오. 선택된 권한.

- Azure AD 앱 전용 인증의 경우: Azure AD에 등록한 후 생성한 개인 키와 클라이언트 ID SharePoint 또한 X.509 인증서도 적어 두십시오.
  - ACL을 사용하지 않는 경우 다음 권한을 추가하세요.

SharePoint
• Sites.Read.All (응용 프로그램) - 모든 사이트 모음의 항목과 목록에 액세스하는 데 필요합니다.

> **Note**
>
> 특정 사이트를 크롤링하려는 경우 도메인에서 사용 가능한 모든 사이트가 아닌 특정 사이트로 권한을 제한할 수 있습니다. 사이트. 선택된 (응용 프로그램) 권한을 구성합니다. 이 API 권한을 사용하려면 Microsoft Graph API를 통해 모든 사이트에 대한 액세스 권한을 명시적으로 설정해야 합니다. 자세한 내용은 사이트에 대한 Microsoft 블로그를 참조하십시오. 선택된 권한.

  - ACL을 사용하는 경우 다음 권한을 추가하세요.

SharePoint
• 사이트. FullControl.All (애플리케이션) - 문서의 ACL을 검색하는 데 필요합니다.

> **Note**
>
> 특정 사이트를 크롤링하려는 경우 도메인에서 사용 가능한 모든 사이트가 아닌 특정 사이트로 권한을 제한할 수 있습니다. 사이트. 선택된 (응용 프로그램) 권한을 구성합니다. 이 API 권한을 사용하려면 Microsoft Graph API를 통해 모든 사이트에 대한 액세스 권한을 명시적으로 설정해야 합니다. 자세한 내용은 사이트에 대한 Microsoft 블로그를 참조하십시오. 선택된 권한.

- SharePoint 앱 전용 인증의 경우: 앱 전용에 권한을 부여하는 동안 생성된 SharePoint 클라이언트 ID 및 클라이언트 암호와 Azure AD에 SharePoint 앱을 등록할 때 생성된 클라이언트 ID 및 클라이언트 암호를 기록해 두었습니다. SharePoint

> **Note**
>
> SharePoint 2013 버전에서는 앱 전용 인증이 지원되지 않습니다. SharePoint

- (선택 사항) OneNote 문서를 크롤링하고 Identity Crawler를 사용하는 경우 다음 권한을 추가했습니다.

Microsoft Graph
• GroupMember.Read.All (응용 프로그램) - 모든 그룹 구성원 자격 읽기 • Notes.Read.All (애플리케이션) - 모든 노트북 읽기 OneNote • Sites.Read.All (애플리케이션) - 모든 사이트 모음의 항목 읽기 • User.Read.All (애플리케이션) - 모든 사용자의 전체 프로필 읽기

> **Note**
>
> 기본 인증 및 앱 전용 인증을 사용하여 엔티티를 크롤링하는 데는 API 권한이 필요하지 않습니다. SharePoint

SharePoint Server에 다음이 있는지 확인하십시오.

- SharePoint 인스턴스 URL과 URL의 도메인 이름을 복사했습니다. SharePoint 입력하는 호스트 URL의 형식은 *https://yourcompany/sites/mysite*입니다. URL은 https로 시작해야 합니다.

> **Note**
>
> (온프레미스/서버) AWS Secrets Manager 는 에 포함된 엔드포인트 정보가 데이터 소스 구성 세부 정보에 지정된 엔드포인트 정보와 동일한지 Amazon Kendra 확인합니다. 이렇게 하면 사용자가 작업을 수행할 권한이 없지만 구성된 보안 암호에 액세스하여 작업을 수행하는 데 Amazon Kendra 를 프록시로 사용하는 보안 문제인 혼동된 대리자 문제를 방지하는 데 도움이 됩니다. 나중에 엔드포인트 정보를 변경하는 경우 새 보안 암호를 생성하여 이 정보를 동기화해야 합니다.

- 콘텐츠의 Amazon Kendra 크롤링이 차단되지 않도록 SharePoint 계정에서 다단계 인증 (MFA) 을 비활성화했습니다. SharePoint
- 액세스 제어에 SharePoint 앱 전용 인증을 사용하는 경우:
  - 사이트 수준에서 App Only를 등록할 때 생성된 SharePoint 클라이언트 ID를 복사했습니다. 클라이언트 ID 형식은 ClientId @입니다TenantId. 예: *ffa956f3-8f89-44e7-b0e4-49670756342c@888d0b57-69f1-4fb8-957f-e1f0bedf82fe*.
  - 사이트 수준에서 App Only를 등록할 때 생성된 SharePoint 클라이언트 암호를 복사했습니다.

  참고: SharePoint Server for App Only 인증을 등록한 경우 단일 사이트에 대해서만 클라이언트 ID와 클라이언트 암호가 생성되므로 SharePoint 앱 전용 인증에는 사이트 URL 하나만 지원됩니다.

> **Note**
>
> SharePoint SharePoint 2013 버전에서는 앱 전용 인증이 지원되지 않습니다.

- 액세스 제어를 위해 사용자 지정 도메인과 함께 이메일 ID를 사용하는 경우:
  - 사용자 지정 이메일 도메인 값(예: *"amazon.com"*)을 기록해 두었습니다.
- IDP 승인을 받은 도메인과 함께 이메일 ID를 사용하는 경우 다음을 복사했습니다.
  - LDAP 서버 엔드포인트(프로토콜 및 포트 번호를 포함한 LDAP 서버의 엔드포인트). 예: *ldap://example.com:389*.
  - LDAP 검색 기반(LDAP 사용자의 검색 기반). 예: *CN=Users,DC=sharepoint,DC=com*.
  - LDAP 사용자 이름과 LDAP 암호.
- 구성된 NTLM 인증 자격 증명 또는 사용자 이름 (SharePoint 계정 사용자 이름) 및 암호 (계정 암호) 를 포함하는 구성된 Kerberos 인증 자격 증명 중 하나입니다. SharePoint

에 다음이 있는지 확인하십시오 AWS 계정.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형 및 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- SharePoint 인증 자격 증명을 AWS Secrets Manager 비밀에 저장하고 API를 사용하는 경우 비밀의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 SharePoint 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

SharePoint 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 SharePoint 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 SharePoint 을 Amazon Kendra 참조하십시오필수 조건.

Console: SharePoint Online

SharePoint 온라인에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 SharePoint 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 SharePoint 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. 호스팅 방법 —온라인을 선택합니다SharePoint .

   b. SharePoint리포지토리 전용 사이트 URL - 호스트 URL을 입력합니다. SharePoint 입력하는 호스트 URL의 형식은 *https://yourdomain.sharepoint.com/sites/mysite*입니다. URL은 `https` 프로토콜로 시작해야 합니다. URL을 새 줄로 구분합니다. 최대 100개 URL을 추가할 수 있습니다.

   c. 도메인 - 도메인을 입력합니다. SharePoint 예를 들어, *https://yourdomain.sharepoint.com/sites/mysite* URL의 도메인은 *yourdomain*입니다.

   d. 권한 부여 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

      또한 사용자 ID 유형 (사용자 계정 이름 또는 Azure Portal에서 가져온 사용자 이메일) 을 선택할 수 있습니다. 지정하지 않으면 기본적으로 이메일이 사용됩니다.

   e. 인증 - 기본, OAuth 2.0, Azure AD 앱 전용 인증, 앱 전용 인증 또는 OAuth 2.0 새로 SharePoint 고침 토큰 인증을 선택합니다. 기존 AWS Secrets Manager 암호를 선택하여 인증 자격 증명을 저장하거나 암호를 생성할 수 있습니다.

i. 기본 인증을 사용하는 경우 암호에는 암호 이름, SharePoint 사용자 이름 및 암호가 포함되어야 합니다.

ii. OAuth 2.0 인증을 사용하는 경우 암호에는 SharePoint 테넌트 ID, 비밀 이름, SharePoint 사용자 이름, 암호, Azure AD에 등록할 때 생성되는 Azure AD 클라이언트 ID 및 Azure SharePoint AD에 등록할 때 생성되는 Azure AD 클라이언트 암호가 포함되어야 합니다. SharePoint

iii. Azure AD 앱 전용 인증을 사용하는 경우 암호에는 SharePoint 테넌트 ID, Azure AD 자체 서명 X.509 인증서, 암호 이름, Azure AD에 SharePoint 등록할 때 생성되는 Azure AD 클라이언트 ID, Azure AD용 커넥터를 인증하기 위한 개인 키가 포함되어야 합니다.

iv. SharePoint앱 전용 인증을 사용하는 경우 암호에는 SharePoint 테넌트 ID, 암호 이름, 테넌트 수준에서 앱 전용을 등록할 때 생성한 SharePoint 클라이언트 ID, 테넌트 수준에서 앱 전용에 등록할 때 생성되는 SharePoint 클라이언트 암호, Azure AD에 등록할 때 생성되는 Azure AD 클라이언트 ID, Azure AD에 SharePoint 등록할 때 생성되는 Azure AD 클라이언트 암호가 포함되어야 합니다. SharePoint

SharePoint ID ### ClientID @###. TenantId 예: *ffa956f3-8f89-44e7-b0e4-49670756342c@888d0b57-69f1-4fb8-957f-e1f0bedf82fe*.

v. OAuth 2.0 새로 고침 토큰 인증을 사용하는 경우 암호에는 SharePoint 테넌트 ID, 암호 이름, Azure AD에 등록할 때 생성되는 고유한 Azure AD 클라이언트 ID, Azure SharePoint AD에 등록할 때 생성되는 Azure AD 클라이언트 암호, SharePoint 연결하기 위해 생성된 새로 고침 토큰이 포함되어야 합니다. Amazon Kendra SharePoint

f. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

g. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

로컬 그룹 매핑 또는 Azure Active Directory 그룹 매핑을 크롤링하도록 선택할 수도 있습니다.

> **Note**
>
> AD 그룹 매핑 크롤링은 OAuth 2.0, OAuth 2.0 새로 고침 토큰 및 앱 전용 인증에만 사용할 수 있습니다. SharePoint

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 IAM 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

i. 개체 선택 - 크롤링할 개체를 선택합니다. 모든 개체를 크롤링하거나 파일, 첨부 파일, 링크 페이지, 이벤트, 설명 및 목록 데이터의 조합을 크롤링하도록 선택할 수 있습니다.

ii. 추가 구성에서 개체 정규식 패턴의 경우 — 모든 문서를 동기화하는 대신 링크, 페이지 및 이벤트에 대한 정규 표현식 패턴을 추가하여 특정 개체를 포함합니다.

iii. Regex 패턴 - 모든 문서를 동기화하는 대신 정규 표현식 패턴을 추가하여 파일 경로, 파일 이름, 파일 유형, OneNote 섹션 이름, OneNote 페이지 이름을 기준으로 파일을 포함하거나 제외할 수 있습니다. 최대 100개까지 추가할 수 있습니다.

> **Note**
>
> OneNote 크롤링은 OAuth 2.0, OAuth 2.0 새로 고침 토큰 및 앱 전용 인증에만 사용할 수 있습니다. SharePoint

b. 동기화 모드의 경우, 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 소스를 처음으로 Amazon Kendra 와 동기화하는 경우 기본적으로 모든 콘텐츠가 동기화됩니다.

- 전체 동기화 - 이전 동기화 상태에 관계없이 모든 콘텐츠를 동기화합니다.
- 신규 또는 수정된 문서 동기화 - 새 문서와 수정된 문서만 동기화합니다.
- 신규, 수정되었거나 삭제된 문서 동기화 - 새 문서와 수정 및 삭제된 문서만 동기화합니다.

c. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

d. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

Console: SharePoint Server

Amazon Kendra 연결하려면 SharePoint

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.

4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 SharePoint 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 SharePoint 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. 호스팅 방법 —서버를 선택합니다SharePoint.
   b. SharePoint버전 선택 SharePoint —2013, SharePoint 2016, SharePoint 2019 및 SharePoint (구독 에디션) 중 하나를 선택합니다.
   c. SharePoint리포지토리 전용 사이트 URL —호스트 URL을 입력합니다. SharePoint 입력하는 호스트 URL의 형식은 *`https://yourcompany/sites/mysite`*입니다. URL은 `https` 프로토콜로 시작해야 합니다. URL을 새 줄로 구분합니다. 최대 100개 URL을 추가할 수 있습니다.
   d. 도메인 - 도메인을 입력합니다. SharePoint 예를 들어, *`https://yourcompany/sites/mysite`* URL의 도메인은 *`yourcompany`*입니다.
   e. SSL 인증서 위치 - SSL 인증서 파일의 Amazon S3 경로를 입력합니다.
   f. (선택 사항) 웹 프록시의 경우 - 호스트 이름(`http://` 또는 `https://` 프로토콜 제외)과 호스트 URL 전송 프로토콜에서 사용하는 포트 번호를 입력합니다. 포트 번호의 숫자 값은 0에서 65535 사이여야 합니다.
   g. 권한 부여 - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록(ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

      SharePoint 서버의 경우 다음 ACL 옵션 중에서 선택할 수 있습니다.

i. IDP의 도메인이 포함된 이메일 ID - 사용자 ID는 기본 ID 공급자 (IDP) 에서 가져온 도메인의 이메일 ID를 기반으로 합니다. 인증의 일환으로 비밀번호에 IDP 연결 세부 정보를 제공합니다. Secrets Manager

ii. 사용자 지정 도메인이 포함된 이메일 ID - 사용자 ID는 사용자 지정 이메일 도메인 값을 기반으로 합니다. 예: *"amazon.com"*. 이 이메일 도메인은 액세스 제어를 위한 이메일 ID를 구성하는 데 사용됩니다. 사용자 지정 이메일 도메인을 입력해야 합니다.

iii. 도메인\ 도메인이 있는 사용자 - 사용자 ID는 도메인/사용자 ID 형식을 사용하여 구성됩니다. 유효한 도메인 이름을 제공해야 합니다. 예를 들어 액세스 제어를 구성하려면 *"sharepoint2019"*를 사용하세요.

h. 인증의 경우 SharePoint 앱 전용 인증, NTLM 인증 또는 Kerberos 인증을 선택합니다. 기존 AWS Secrets Manager 암호를 선택하여 인증 자격 증명을 저장하거나 암호를 생성할 수 있습니다.

i. NTLM 인증 또는 Kerberos 인증을 사용하는 경우 암호에는 암호 이름, 사용자 이름 및 암호가 포함되어야 합니다.

IDP의 도메인이 포함된 이메일 ID를 사용하는 경우 다음 항목도 입력하세요.

- LDAP 서버 엔드포인트 - 프로토콜 및 포트 번호를 포함한 LDAP 서버의 엔드포인트. 예: *ldap://example.com:389*.
- LDAP 검색 기반 - LDAP 사용자의 검색 기반입니다. 예: *CN=Users,DC=sharepoint,DC=com*.
- LDAP 사용자 이름 - LDAP 사용자 이름입니다.
- LDAP 암호 - 사용자의 LDAP 암호입니다.

ii. SharePoint 앱 전용 인증을 사용하는 경우 암호에는 비밀 이름, 사이트 수준에서 앱 전용을 등록할 때 생성한 SharePoint 클라이언트 ID, 사이트 수준에서 앱 전용에 등록할 때 생성되는 SharePoint 클라이언트 암호가 포함되어야 합니다.

SharePoint *##### ID ### ClientID @###. TenantId* 예: *ffa956f3-8f89-44e7-b0e4-49670756342c@888d0b57-69f1-4fb8-957f-e1f0bedf82fe*.

참고: SharePoint Server for App Only 인증을 등록한 경우 단일 사이트에 대해서만 클라이언트 ID와 클라이언트 암호가 생성되므로 앱 전용 인증에는 사이트 URL 하나만 지원됩니다. SharePoint

IDP의 도메인이 포함된 이메일 ID를 사용하는 경우 다음 항목도 입력하세요.

- LDAP 서버 엔드포인트 - 프로토콜 및 포트 번호를 포함한 LDAP 서버의 엔드포인트. 예: *ldap://example.com:389*.
- LDAP 검색 기반 - LDAP 사용자의 검색 기반입니다. 예: *CN=Users,DC=sharepoint,DC=com*.
- LDAP 사용자 이름 - LDAP 사용자 이름입니다.
- LDAP 암호 - 사용자의 LDAP 암호입니다.

i. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

j. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

   로컬 그룹 매핑 또는 Azure Active Directory 그룹 매핑을 크롤링하도록 선택할 수도 있습니다.

> **Note**
>
> AD 그룹 매핑 크롤링은 앱 전용 인증에서만 SharePoint 사용할 수 있습니다.

k. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

l. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

   a. 동기화 범위에서 다음 옵션을 선택합니다.

      i. 개체 선택 - 크롤링할 개체를 선택합니다. 모든 개체를 크롤링하거나 파일, 첨부 파일, 링크 페이지, 이벤트 및 목록 데이터의 조합을 크롤링하도록 선택할 수 있습니다.

      ii. 추가 구성에서 개체 정규식 패턴의 경우 — 모든 문서를 동기화하는 대신 링크, 페이지 및 이벤트에 대한 정규 표현식 패턴을 추가하여 특정 개체를 포함합니다.

      iii. Regex 패턴 - 모든 문서를 동기화하는 대신 파일 경로, 파일 이름, 파일 유형, OneNote섹션 이름, OneNote페이지 이름을 기준으로 파일을 포함하거나 제외하는 정규 표현식 패턴을 추가합니다. 최대 100개까지 추가할 수 있습니다.

      > **Note**
      >
      > OneNote 크롤링은 앱 전용 인증에서만 사용할 수 있습니다. SharePoint

   b. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

      - 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
      - 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
      - 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

   c. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

   d. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra 연결하려면 SharePoint

TemplateConfigurationAPI를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 `SHAREPOINTV2` 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 `TEMPLATE` 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 리포지토리 엔드포인트 메타데이터 - `siteUrls` SharePoint 인스턴스의 종료를 지정합니다. `tenantID` `domain`
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - `FORCED_FULL_CRAWL`모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - `FULL_CRAWL`데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - `CHANGE_LOG`데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로

검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

> **Note**
>
> 아이덴티티 크롤러는 로 설정한 경우에만 사용할 수 있습니다. `crawlAcl` `true`

- 리포지토리 추가 속성 - 다음을 지정하세요.
  - (Azure AD의 `s3certificateName` 경우) `s3bucketName` Azure AD 자체 서명 X.509 인증서를 저장하는 데 사용합니다.
  - 사용하는 인증 유형 (`auth_Type`) (예:,,, `OAuth2OAuth2App`, `OAuth2Certificate` 및) `Basic` `OAuth2_RefreshToken` `NTLM` `Kerberos`
  - 사용 중인 버전 (`version`) () (`ServerOnline`여부) `Server`를 사용하는 경우 `onPremVersion`를 `2013`, `2016`, `2019`, 또는 `SubscriptionEdition`로 추가 지정할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. SharePoint

  SharePoint 온라인을 사용하는 경우 기본, OAuth 2.0, Azure AD 앱 전용 및 앱 전용 인증 중에서 선택할 수 있습니다. SharePoint 다음은 각 인증 옵션의 보안 암호에 포함되어야 하는 최소 JSON 구조입니다.

  - 기본 인증

```
{
    "userName": "SharePoint account user name",
    "password": "SharePoint account password"
}
```

  - OAuth 2.0 인증

```
{
    "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"
}
```

- Azure AD 앱 전용 인증

```
{
    "clientId": "client id generated when registering SharePoint with Azure AD",
    "privateKey": "private key to authorize connection with Azure AD"
}
```

- SharePoint 앱 전용 인증

```
{
    "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"
}
```

- OAuth 2.0 새로 고침 토큰 인증

```
{
    "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"
}
```

SharePoint 서버를 사용하는 경우 SharePoint 앱 전용 인증, NTLM 인증 및 Kerberos 인증 중에서 선택할 수 있습니다. 다음은 각 인증 옵션의 보안 암호에 포함되어야 하는 최소 JSON 구조입니다.

- SharePoint 앱 전용 인증

```
{
    "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 IDP 인증을 통한 도메인을 사용한 앱 전용 인증

```
{
    "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"
}
```

- (서버만 해당) NTLM 또는 Kerberos 인증

```
{
    "siteUrlsHash": "Hash representation of SharePoint site URLs",
    "userName": "SharePoint account user name",
    "password": "SharePoint account password"
}
```

- (서버만 해당) IDP 권한 부여를 통한 도메인이 포함된 NTLM 또는 Kerberos 인증

```
{
    "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 역할 - `CreateDataSource` Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할에 제공하고 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. SharePoint Amazon Kendra자세한 내용은 SharePoint 데이터 원본의IAM 역할을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 특정 파일 및 기타 콘텐츠를 포함할지 제외할지 여부를 지정할 수 있습니다. OneNotes

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 필드 매핑 - SharePoint 데이터 원본 필드를 인덱스 필드에 매핑하도록 선택합니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 SharePoint 템플릿 스키마를 참조하십시오.

참고

- 커넥터는 파일 개체에 대해서만 사용자 지정 필드 매핑을 지원합니다.
- 모든 SharePoint 서버 버전에서 ACL 토큰은 소문자여야 합니다. IDP의 도메인이 포함된 이메일 및 사용자 지정 도메인이 포함된 이메일 ID ACL의 경우 예를 들면 *`user@sharepoint2019.com`*. 도메인이 있는 도메인\사용자 ACL의 경우의 예를 들면 *`sharepoint2013\user`*.
- 커넥터는 2013년의 변경 로그 모드/신규 또는 수정된 콘텐츠 동기화를 지원하지 않습니다. SharePoint
- 개체 이름에 % 문자가 있는 경우 커넥터는 API 제한으로 인해 이러한 파일을 건너뛰게 됩니다.
- OneNote 테넌트 ID를 사용하고 OAuth 2.0, OAuth 2.0 새로 고침 토큰 또는 온라인에 앱 전용 인증을 활성화한 경우에만 커넥터를 통해 크롤링할 수 있습니다. SharePoint SharePoint

- 커넥터는 문서 이름이 바뀌더라도 기본 이름만 사용하여 OneNote 문서의 첫 번째 섹션을 크롤링합니다.
- 커넥터는 링크 외에도 크롤링할 엔티티로 페이지 및 파일을 선택한 경우에만 SharePoint 2019, SharePoint 온라인 및 구독 에디션에서 링크를 크롤링합니다.
- 링크를 크롤링할 엔티티로 선택한 경우 커넥터는 SharePoint 2013년과 SharePoint 2016년에 링크를 크롤링합니다.
- 목록 데이터도 크롤링될 개체로 선택된 경우에만 커넥터는 목록 첨부 파일 및 설명을 크롤링합니다.
- 이벤트도 크롤링될 개체로 선택된 경우에만 커넥터는 이벤트 첨부 파일을 크롤링합니다.
- SharePoint 온라인 버전의 경우 ACL 토큰은 소문자입니다. *## ## Azure #### **### ## ### MaryMajor @domain .com# ## SharePoint #### ACL ### marymajor@domain.com # ###.*
- SharePoint 온라인 및 서버용 ID 크롤러에서 중첩된 그룹을 크롤링하려면 AD 그룹 크롤링뿐만 아니라 로컬 크롤링도 활성화해야 합니다.
- SharePoint Online을 사용하고 Azure Portal의 사용자 계정 이름이 대문자와 소문자의 조합인 경우 SharePoint API는 내부적으로 이 이름을 소문자로 변환합니다. 이로 인해 Amazon Kendra SharePoint 커넥터는 ACL을 소문자로 설정합니다.

Microsoft SQL Server

Microsoft SQL Server는 Microsoft에서 개발한 관계형 데이터베이스 관리 시스템(RDBMS)입니다. Microsoft SQL Server사용자인 경우 Microsoft SQL Server 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다. Amazon Kendra Microsoft SQL Server데이터 소스 커넥터는 MS SQL Server 2019를 지원합니다.

Amazon Kendra 콘솔과 TemplateConfigurationAPI를 Amazon Kendra 사용하여 Microsoft SQL Server 데이터 원본에 연결할 수 있습니다.

Amazon Kendra Microsoft SQL Server데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Microsoft SQL Server데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Microsoft SQL Server 및 AWS 계정에서 다음과 같이 변경하십시오.

Microsoft SQL Server에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
>
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다.
- Microsoft SQL Server 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Microsoft SQL Server 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Microsoft SQL Server 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Microsoft SQL Server데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Microsoft SQL Server 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 Microsoft SQL Server 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Microsoft SQL Server

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 Microsoft SQL Server커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Microsoft SQL Server커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. 소스에서 다음 정보를 입력합니다.

b. 호스트 - 데이터베이스 호스트 이름을 입력합니다.

c. 포트 - 데이터베이스 포트를 입력합니다.

d. 인스턴스 - 데이터베이스 인스턴스를 입력합니다.

e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.

f. 인증에서 - 다음 정보를 입력합니다.

- AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Microsoft SQL Server 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

  A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

     I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Microsoft SQL Server -'가 자동으로 추가됩니다.

     II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

  B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.

> **Note**
>
> 테이블 이름에 특수 문자 (영숫자가 아닌) 가 포함된 경우 테이블 이름 주위에 대괄호를 사용해야 합니다. 예를 들어, `[] ##*# #####`. my-database-table

- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.

- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra Microsoft SQL Server

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 sqlserver로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.

  > **Note**
  >
  > 테이블 이름에 특수 문자 (영숫자가 아닌) 가 포함된 경우 테이블 이름 주위에 대괄호를 사용해야 합니다. 예를 들어, *[] ##*# #####*. my-database-table

- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. Microsoft SQL Server 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 - `CreateDataSource` Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Microsoft SQL Server Amazon Kendra자세한 내용은 Microsoft SQL Server 데이터 소스에 대한 IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Microsoft SQL Server 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Microsoft SQL Server 템플릿 스키마를 참조하세요.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Microsoft Teams

Microsoft Teams는 메시징, 회의 및 파일 공유를 위한 엔터프라이즈 협업 도구입니다. Microsoft Teams 사용자인 경우 Microsoft Teams 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

Amazon Kendra 콘솔과 TemplateConfigurationAPI를 Amazon Kendra 사용하여 Microsoft Teams 데이터 원본에 연결할 수 있습니다.

Amazon Kendra Microsoft Teams 데이터 원본 커넥터의 문제 해결에 대한 자세한 내용은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Microsoft Teams 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있으려면 먼저 Microsoft Teams 및 AWS 계정에서 이러한 변경을 수행하십시오.

Microsoft Teams에서는 다음을 확인하세요.

- Office 365에서 Microsoft Teams 계정을 만들었습니다.
- Microsoft 365 테넌트 ID를 기록했습니다. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
- Azure Portal에서 OAuth 애플리케이션을 구성하고 클라이언트 ID와 클라이언트 암호 또는 클라이언트 자격 증명을 기록했습니다. 자세한 내용은 Microsoft 자습서 및 등록된 앱 예제를 참조하십시오.

> **Note**
>
> Azure 포털에서 앱을 만들거나 등록할 때 비밀 ID는 실제 비밀 값을 나타냅니다. 암호 및 앱을 만들 때 실제 암호 값을 즉시 기록하거나 저장해야 합니다. Azure Portal에서 애플리케이션 이름을 선택한 다음 인증서 및 비밀의 메뉴 옵션으로 이동하여 비밀번호에 액세스할 수 있습니다.
> Azure Portal에서 애플리케이션 이름을 선택한 다음 개요 페이지로 이동하여 클라이언트 ID에 액세스할 수 있습니다. 애플리케이션 (클라이언트) ID는 클라이언트 ID입니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 필요한 권한이 추가되었습니다. 모든 권한을 추가하도록 선택하거나 크롤링하려는 엔티티에 따라 더 적은 권한을 선택하여 범위를 제한할 수 있습니다. 다음 표에는 해당 엔티티별 애플리케이션 수준 권한이 나열되어 있습니다.

개체	데이터 동기화에 필요한 권한	ID 동기화에 필요한 권한
채널 게시	• ChannelMessage.Read.All • Group.Read.All	TeamMember. 모두 읽기.

개체	데이터 동기화에 필요한 권한	ID 동기화에 필요한 권한
	• User.Read • User.Read.All	
채널 첨부	• ChannelMessage. 모두 읽기. • Group.Read.All • User.Read • User.Read.All	TeamMember. 모두 읽기.
채널 위키	• Group.Read.All • User.Read • User.Read.All	TeamMember. 모두 읽기.
채팅 메시지	• Chat.Read.All • ChatMessage. 모두 읽기. • ChatMember. 모두 읽기. • User.Read • User.Read.All • Group.Read.All	TeamMember. 모두 읽기.
회의 대화	• Chat.Read.All • ChatMessage. 읽어보기 • ChatMember. 읽기. 모두 • User.Read • User.Read.All • Group.Read.All	TeamMember. 모두 읽기.

개체	데이터 동기화에 필요한 권한	ID 동기화에 필요한 권한
채팅 첨부	• Chat.Read.All • ChatMessage. 읽어보기 • ChatMember. 읽기. 모두 • User.Read • User.Read.All • Group.Read.All	TeamMember. 모두 읽기.
회의 파일	• Chat.Read.All • ChatMessage. 모두 읽기. • ChatMember. 모두 읽기. • User.Read • User.Read.All • Group.Read.All • Files.Read.All	TeamMember. 모두 읽기.
달력 회의	• Chat.Read.All • ChatMessage. 모두 읽기. • ChatMember. 모두 읽기. • User.Read • User.Read.All • Group.Read.All • Files.Read.All	TeamMember. 모두 읽기.
회의록	• User.Read • User.Read.All • Group.Read.All • Files.Read.All	TeamMember. 모두 읽기.

- Microsoft Teams 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

귀하의 서류에 AWS 계정다음이 있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Microsoft Teams 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Microsoft Teams 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Microsoft Teams 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Microsoft Teams 데이터 원본의 필수 세부 정보를 제공해야 합니다. Microsoft Teams를 위해 Amazon Kendra아직 구성하지 않은 경우 을 참조하십시오필수 조건.

Console

마이크로소프트 Amazon Kendra 팀즈에 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Microsoft Teams 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2를 사용하는 경우 (해당하는 경우), “V2.0" 태그가 있는 Microsoft Teams 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. 테넌트 ID - Microsoft 365 테넌트 ID를 입력합니다. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
   b. 권한 부여 - ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
   c. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 Secrets Manager 암호를 만들어 Microsoft Teams 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.
      i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.
         A. 보안 암호 이름 - 보안 암호의 이름. 비밀 이름에는 접두사 AmazonKendra '-Microsoft Teams-'가 자동으로 추가됩니다.

B. 클라이언트 ID 및 클라이언트 암호의 경우 - Microsoft Teams에 구성된 인증 자격 증명을 Azure 포털에 입력합니다.

ii. 암호를 저장하고 추가하세요.

d. 결제 모델 - Microsoft Teams 계정의 라이선스 및 결제 모델을 선택할 수 있습니다. 모델 A 결제 모델은 보안 준수가 필요한 라이선싱 및 결제 모델로 제한됩니다. 모델 B 결제 모델은 보안 준수가 필요하지 않은 라이선싱 및 결제 모델에 적합합니다.

e. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

f. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

g. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

h. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 콘텐츠 동기화 - 크롤링할 콘텐츠 유형을 선택합니다. 채팅, 팀 및 캘린더 콘텐츠를 크롤링하도록 선택할 수 있습니다.

b. 추가 구성 —특정 캘린더 시작 및 종료 날짜, 사용자 이메일, 팀 이름, 채널 이름, 첨부 파일 등을 지정합니다. OneNotes

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 인덱

싱됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

마이크로소프트 Amazon Kendra 팀즈에 연결하려면

TemplateConfigurationAPI를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 MSTEAMS 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource

- 테넌트 ID - Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - `FORCED_FULL_CRAWL`모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - `FULL_CRAWL`데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - `CHANGE_LOG`데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - Microsoft Teams 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공하십시오. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM 역할 - `CreateDataSource` Secrets Manager 암호에 액세스할 수 있는 권한을 IAM 역할에 제공하고 Microsoft Teams 커넥터 및 에 필요한 공개 API를 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Microsoft Teams 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 문서/콘텐츠 유형 - 채팅 메시지 및 첨부 파일, 채널 게시물 및 첨부 파일, 채널 Wiki, 캘린더 콘텐츠, 회의 채팅, 파일 및 메모를 크롤링할지 여부를 지정합니다.
- 캘린더 콘텐츠 - 캘린더 콘텐츠를 크롤링할 시작 및 종료 날짜/시간을 지정합니다.
- 포함 및 제외 필터 - Microsoft Teams의 특정 콘텐츠를 포함할지 또는 제외할지 여부를 지정할 수 있습니다. 팀 이름, 채널 이름, 파일 이름 및 파일 형식, 사용자 이메일, OneNote 섹션, 페이지를 포함하거나 제외할 수 있습니다. OneNote

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 필드 매핑 - Microsoft Teams 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 해당 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 다른 중요한 JSON 키 목록은 Microsoft Teams 템플릿 스키마를 참조하십시오.

자세히 알아보기

Microsoft Teams 데이터 Amazon Kendra 원본과의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- Microsoft Teams용 Amazon Kendra 커넥터를 사용하여 조직의 Microsoft Teams 데이터 원본을 지능적으로 검색합니다.

Microsoft Yammer

Microsoft Yammer는 메시징, 회의 및 파일 공유를 위한 엔터프라이즈 협업 도구입니다. Microsoft Yammer 사용자인 경우 Microsoft Yammer 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

Amazon Kendra 콘솔과 TemplateConfigurationAPI를 Amazon Kendra 사용하여 Microsoft Yammer 데이터 원본에 연결할 수 있습니다.

Amazon Kendra Microsoft Yammer 데이터 원본 커넥터의 문제 해결에 대한 자세한 내용은 을 참조하십시오데이터 소스 문제 해결.

지원되는 기능

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Microsoft Yammer 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있으려면 먼저 Microsoft Yammer 및 AWS 계정에서 다음과 같이 변경하십시오.

Microsoft Yammer에서는 다음을 확인하세요.

- Office 365에서 Microsoft Yammer 관리 계정을 만들었습니다.
- Microsoft Yammer 사용자 이름과 암호를 기록했습니다.
- Microsoft 365 테넌트 ID를 기록했습니다. 테넌트 ID는 Azure Active Directory 포털의 속성 또는 OAuth 애플리케이션에서 찾을 수 있습니다.
- Azure 포털에서 OAuth 응용 프로그램을 구성하고 클라이언트 ID와 클라이언트 암호 또는 클라이언트 자격 증명을 기록했습니다. 자세한 내용은 Microsoft 자습서 및 등록된 앱 예제를 참조하십시오.

> **Note**
>
> Azure 포털에서 앱을 만들거나 등록할 때 비밀 ID는 실제 비밀 값을 나타냅니다. 암호 및 앱을 만들 때 실제 암호 값을 즉시 기록하거나 저장해야 합니다. Azure Portal에서 애플리케이

션 이름을 선택한 다음 인증서 및 비밀의 메뉴 옵션으로 이동하여 비밀번호에 액세스할 수 있습니다.
Azure Portal에서 애플리케이션 이름을 선택한 다음 개요 페이지로 이동하여 클라이언트 ID에 액세스할 수 있습니다. 애플리케이션 (클라이언트) ID는 클라이언트 ID입니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- Microsoft Yammer 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Microsoft Yammer 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Microsoft Yammer 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Microsoft Yammer 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Microsoft Yammer 데이터 원본의 필수 세부 정보를 제공해야 합니다. Microsoft Yammer를 위해 Amazon Kendra아직 구성하지 않은 경우 을 참조하십시오필수 조건.

Console

Microsoft Amazon Kendra Yammer에 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Microsoft Yammer 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Microsoft Yammer 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

b. AWS Secrets Manager 암호 - 기존 암호를 선택하거나 새 Secrets Manager 암호를 만들어 Microsoft Yammer 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

   i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

      A. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에는 접두사 AmazonKendra '-Microsoft Yammer-'가 자동으로 추가됩니다.

      B. 사용자 이름, 암호의 경우 - Microsoft Yammer 사용자 이름과 암호를 입력합니다.

      C. 클라이언트 ID의 경우 클라이언트 암호 - Microsoft Yammer에 구성된 인증 자격 증명을 Azure 포털에 입력합니다.

   ii. 암호를 저장하고 추가합니다.

c. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

d. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

e. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

f. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

   a. 시작 날짜 - Microsoft Yammer에서 데이터 크롤링을 시작할 날짜를 지정합니다.

   b. 콘텐츠 동기화 - 크롤링할 콘텐츠 유형을 선택합니다. 공개 메시지, 비공개 메시지, 첨부 파일 등을 예로 들 수 있습니다.

   c. 추가 구성 - 크롤링하려는 특정 커뮤니티 이름을 지정하고 정규 표현식 패턴을 사용하여 특정 콘텐츠를 포함하거나 제외할 수도 있습니다.

   d. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

      - 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
      - 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
      - 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

   e. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

   f. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

   b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

   c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Microsoft Amazon Kendra Yammer에 연결하려면

API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다 TemplateConfiguration. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 YAMMER 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - Microsoft Yammer 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공하십시오. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "username": "user name",
    "password": "password",
    "clientId": "client ID",
    "clientSecret": "client secret"
}
```

- IAM 역할 - CreateDataSource Secrets Manager 암호에 액세스할 수 있는 권한을 IAM 역할에 제공하고 Microsoft Yammer 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 RoleArn 시기를 지정합니다. Amazon Kendra자세한 내용은 Microsoft Yammer 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 문서/콘텐츠 유형 - 커뮤니티 콘텐츠, 메시지 및 첨부 파일, 개인 메시지를 크롤링할지 여부를 지정합니다.
- 포함 및 제외 필터 - 특정 콘텐츠를 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 필드 매핑 - Microsoft Yammer 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 다른 중요한 JSON 키 목록은 Microsoft Yammer 템플릿 스키마를 참조하십시오.

자세히 알아보기

Microsoft Yammer 데이터 Amazon Kendra 원본과의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- 에 대한 Yammer 커넥터 발표 Amazon Kendra

MySQL

MySQL는 오픈 소스 관계형 데이터베이스 관리 시스템입니다. MySQL사용자인 경우 MySQL 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다. Amazon Kendra MySQL데이터 소스 커넥터는 MySQL 8.0을 지원합니다. 21.

Amazon Kendra 콘솔과 API를 Amazon Kendra 사용하여 MySQL 데이터 소스에 연결할 수 있습니다. TemplateConfiguration

Amazon Kendra MySQL데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

MySQL데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 MySQL 및 AWS 계정에서 다음과 같이 변경하십시오.

MySQL에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
>
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다.
- MySQL 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- MySQL 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 MySQL 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

MySQL데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 MySQL 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 MySQL 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 MySQL

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 MySQL커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 MySQL커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. 소스에서 다음 정보를 입력합니다.
   b. 호스트 - 데이터베이스 호스트 이름을 입력합니다.
   c. 포트 - 데이터베이스 포트를 입력합니다.
   d. 인스턴스 - 데이터베이스 인스턴스를 입력합니다.

e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.

f. 인증에서 - 다음 정보를 입력합니다.

- AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 MySQL 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

  A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

     I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-MySQL -'가 자동으로 추가됩니다.

     II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

  B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.

- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra MySQL

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 mySql로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. MySQL 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 - `CreateDataSource` Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. MySQL Amazon Kendra자세한 내용은 MySQL 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - MySQL 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Oracle Database

Oracle Database는 데이터베이스 관리 시스템입니다. Oracle Database사용자인 경우 Oracle Database 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다. Amazon Kendra Oracle Database데이터 소스 커넥터는 오라클 데이터베이스 18c, 19c 및 21c를 지원합니다.

Amazon Kendra 콘솔과 API를 사용하여 Oracle Database 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Oracle Database데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Oracle Database데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하기 전에 먼저 Oracle Database 및 AWS 계정에서 다음과 같이 변경하십시오.

Oracle Database에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
>
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다.
- Oracle Database 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Oracle Database 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Oracle Database 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Oracle Database데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Oracle Database 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 Oracle Database 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 Oracle Database

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 Oracle Database커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Oracle Database커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. 소스에서 다음 정보를 입력합니다.

b. 호스트 - 데이터베이스 호스트 이름을 입력합니다.

c. 포트 - 데이터베이스 포트를 입력합니다.

d. 인스턴스 - 데이터베이스 인스턴스를 입력합니다.

e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.

f. 인증에서 - 다음 정보를 입력합니다.

- AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Oracle Database 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

  A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

     I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Oracle Database -'가 자동으로 추가됩니다.

     II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

  B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.
- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인

됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra Oracle Database

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 oracle로 지정해야 합니다.

- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - `FORCED_FULL_CRAWL`모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - `FULL_CRAWL`데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - `CHANGE_LOG`데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. Oracle Database 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Oracle Database Amazon Kendra자세한 내용은 Oracle Database 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Oracle Database 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Oracle Database 템플릿 스키마를 참조하세요.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

PostgreSQL

PostgreSQL는 오픈 소스 데이터베이스 관리 시스템입니다. PostgreSQL사용자인 경우 PostgreSQL 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다. Amazon Kendra PostgreSQL데이터 소스 커넥터는 PostgreSQL 9.6을 지원합니다.

Amazon Kendra 콘솔과 API를 사용하여 PostgreSQL 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra PostgreSQL데이터 소스 커넥터 문제 해결은 을 참조하십시오데이터 소스 문제 해결.

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 참고

지원되는 기능

- 필드 매핑
- 사용자 컨텍스트 필터링
- 포함/제외 필터
- 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

PostgreSQL데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하기 전에 먼저 PostgreSQL 및 AWS 계정에서 다음과 같이 변경하십시오.

PostgreSQL에서 다음 사항을 갖추었는지 확인하세요.

- 데이터베이스 사용자 이름 및 암호를 기록했습니다.

> **Important**
>
> 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.

- 데이터베이스 호스트 URL, 포트, 인스턴스를 복사했습니다.
- PostgreSQL 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- PostgreSQL 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 PostgreSQL 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 Amazon Kendra있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

PostgreSQL데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 PostgreSQL 자격 증명의 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 PostgreSQL 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 PostgreSQL

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 PostgreSQL커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 PostgreSQL커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

   d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

   a. 소스에서 다음 정보를 입력합니다.

   b. 호스트 - 데이터베이스 호스트 이름을 입력합니다.

   c. 포트 - 데이터베이스 포트를 입력합니다.

   d. 인스턴스 - 데이터베이스 인스턴스를 입력합니다.

   e. SSL 인증서 위치 활성화 - SSL 인증서 파일의 Amazon S3 경로를 입력하도록 선택합니다.

   f. 인증에서 - 다음 정보를 입력합니다.

      - AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 PostgreSQL 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

        A. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

I. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-PostgreSQL -'가 자동으로 추가됩니다.

II. 데이터베이스 사용자 이름 및 암호의 경우 - 데이터베이스에서 복사한 보안 인증 값을 입력합니다.

B. 저장을 선택합니다.

g. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 동기화 범위에서 다음 옵션을 선택합니다.

- SQL 쿼리 - SELECT 및 JOIN 연산과 같은 SQL 쿼리 문을 입력합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 프라이머리 키 열 - 데이터베이스 테이블의 프라이머리 키를 입력합니다. 이는 데이터베이스 내의 테이블을 식별합니다.
- 제목 열 - 데이터베이스 테이블에 있는 문서 제목 열의 이름을 입력합니다.
- 본문 열 - 데이터베이스 테이블에 있는 문서 본문 열의 이름을 입력합니다.

b. 추가 구성 - 선택 사항에서 모든 파일을 동기화하는 대신 다음 옵션 중 하나를 선택하여 특정 콘텐츠를 동기화합니다.

- 변경 감지 열 - 콘텐츠 변경을 감지하는 Amazon Kendra 데 사용할 열 이름을 입력합니다. Amazon Kendra 이러한 열에 변경 사항이 있을 경우 콘텐츠를 다시 인덱싱합니다.
- 사용자 ID 열 - 콘텐츠에 대한 액세스를 허용할 사용자 ID가 포함된 열의 이름을 입력합니다.

- 그룹 열 - 콘텐츠에 대한 액세스를 허용할 그룹이 포함된 열의 이름을 입력합니다.
- 소스 URL 열 - 인덱싱할 소스 URL이 포함된 열의 이름을 입력합니다.
- 타임스탬프 열 —타임스탬프가 포함된 열의 이름을 입력합니다. Amazon Kendra 타임스탬프 정보를 사용하여 콘텐츠의 변경 사항을 감지하고 변경된 콘텐츠만 동기화합니다.
- 시간대 열 - 콘텐츠를 크롤링할 시간대가 포함된 열의 이름을 입력합니다.
- 타임스탬프 형식 - 콘텐츠 변경 사항을 감지하고 콘텐츠를 다시 동기화하는 데 사용할 타임스탬프 형식이 포함된 열의 이름을 입력합니다.

c. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

d. 동기화 실행 일정의 빈도 - Amazon Kendra 가 데이터 소스와 동기화할 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 소스 필드 (문서 ID, 문서 제목, 소스 URL) 중에서 색인에 매핑하려는 Amazon Kendra 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra PostgreSQL

TemplateConfigurationAPI를 사용하여 다음을 지정해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 JDBC 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 데이터베이스 유형 - 데이터베이스 유형을 postgresql로 지정해야 합니다.
- SQL 쿼리 - SELECT 및 JOIN 작업과 같은 SQL 쿼리 문을 지정합니다. SQL 쿼리는 32KB 미만이어야 합니다. Amazon Kendra 는 쿼리와 일치하는 모든 데이터베이스 콘텐츠를 크롤링합니다.
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. PostgreSQL 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "user name": "database user name",
    "password": "password"
}
```

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. PostgreSQL Amazon Kendra자세한 내용은 PostgreSQL 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 사용자 ID, 그룹, 소스 URL, 타임스탬프, 시간대를 사용하여 특정 콘텐츠를 포함할지 여부를 지정할 수 있습니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - PostgreSQL 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 PostgreSQL 템플릿 스키마를 참조하세요.

참고

- 업데이트된 내용을 Amazon Kendra 확인할 때 삭제된 데이터베이스 행은 추적되지 않습니다.
- 데이터베이스 행의 필드 이름 및 값 크기는 400KB를 초과할 수 없습니다.
- 데이터베이스 데이터 원본에 많은 양의 데이터가 있고 첫 번째 동기화 후 모든 데이터베이스 콘텐츠를 인덱싱하고 싶지 Amazon Kendra 않은 경우 새 문서, 수정된 문서 또는 삭제된 문서만 동기화하도록 선택할 수 있습니다.
- 읽기 전용 데이터베이스 자격 증명을 Amazon Kendra 제공하는 것이 가장 좋습니다.
- 가장 좋은 방법은 민감한 데이터나 개인 식별 정보(PII)가 포함된 테이블을 추가하지 않는 것입니다.

Quip

Quip은 실시간 문서 작성 기능을 제공하는 협업 생산성 소프트웨어입니다. Quip 폴더, 파일, 파일 댓글, 채팅방 및 첨부 파일을 인덱싱하는 Amazon Kendra 데 사용할 수 있습니다.

Amazon Kendra 콘솔과 API를 Amazon Kendra 사용하여 Quip 데이터 소스에 연결할 수 있습니다. QuipConfiguration

Amazon Kendra Quip 데이터 소스 커넥터 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Quip 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- Virtual Private Cloud(VPC)

필수 조건

Quip 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Quip 및 AWS 계정에서 다음과 같이 변경하십시오.

Quip에서 다음 사항을 갖추었는지 확인하세요.

- 관리자 권한이 있는 Quip 계정.
- 개인용 액세스 토큰이 포함된 Quip 보안 인증을 생성했습니다. 토큰은 비밀에 저장된 인증 자격 증명으로 사용됩니다. AWS Secrets Manager 자세한 내용은 인증에 대한 Quip 설명서를 참조하세요.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- Quip 사이트 도메인을 복사했습니다. 예를 들어 *https://quip-company.quipdomain.com/browse*, 여기서 *quipdomain*은 도메인입니다.
- Quip 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하십시오. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Quip 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Quip 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Quip 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Quip 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 Quip을 구성하지 않은 경우 을 Amazon Kendra참조하십시오. 필수 조건

Console

Quip에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Quip 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2(해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Quip 커넥터를 선택합니다.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. Quip 도메인 이름 - Quip 계정에서 복사한 Quip을 입력합니다.

b. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Quip Secrets Manager 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Quip'이 자동으로 추가됩니다.

B. Quip 토큰 - Quip 개인 액세스로 구성된 Quip을 입력합니다.

ii. 비밀번호를 추가하고 저장합니다.

c. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

d. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

e. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 크롤링할 Quip 폴더 ID 추가 - 크롤링하려는 Quip 폴더 ID.

> **Note**
>
> 루트 폴더 (모든 하위 폴더 및 문서 포함) 를 크롤링하려면 루트 폴더 ID를 추가합니다. 특정 하위 폴더를 크롤링하려면 특정 하위 폴더 ID를 추가하세요.

b. 추가 구성 (콘텐츠 유형) - 크롤링하려는 콘텐츠 유형을 입력합니다.

c. 정규식 패턴 - 특정 페이지 및 자산을 포함하거나 제외하기 위한 정규 표현식 패턴입니다. 최대 100개의 패턴을 추가할 수 있습니다.

d. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

e. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 생성된 기본 데이터 원본 필드 중에서 Amazon Kendra 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Amazon Kendra Quip에 연결하려면

QuipConfiguration API를 사용하여 다음을 지정해야 합니다.

- Quip 사이트 도메인 - 예: *https://quip-company.quipdomain.com/browse* 여기에서 *quipdomain*은 도메인입니다.
- 비밀 Amazon 리소스 이름 (ARN) - Quip 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "accessToken": "token"
```

```
}
```

- IAM 역할 - `CreateDataSource` Secrets Manager 비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 Quip 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Quip 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - 데이터 소스 구성의 일부로 `VpcConfiguration`를 지정합니다. VPC 사용을 위한 Amazon Kendra 구성을 참조하세요.
- 포함 및 제외 필터 - 특정 파일을 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 폴더 - 인덱싱하려는 Quip 폴더 및 하위 폴더를 지정합니다.

> **Note**
>
> 루트 폴더 (모든 하위 폴더 및 문서 포함) 를 크롤링하려면 루트 폴더 ID를 입력합니다. 특정 하위 폴더를 크롤링하려면 특정 하위 폴더 ID를 추가합니다.

- 첨부 파일, 채팅방, 파일 댓글 - 첨부 파일 크롤링, 대화방 콘텐츠, 파일 댓글을 포함할지 여부를 선택합니다.
- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에Amazon Kendra ACL이 있는 경우 문서에 대한 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.
- 필드 매핑 - Quip 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

자세히 알아보기

Quip 데이터 Amazon Kendra 원본과의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- Quip 커넥터를 사용하여 지능형 검색으로 Quip 문서에서 지식을 검색하십시오. Amazon Kendra

Salesforce

Salesforce는 지원, 영업 및 마케팅 팀을 관리하기 위한 CRM(고객 관계 관리) 도구입니다. 를 Amazon Kendra 사용하여 Salesforce 표준 개체 및 사용자 지정 개체를 인덱싱할 수 있습니다.

Amazon Kendra 콘솔, TemplateConfigurationAPI 또는 API를 Amazon Kendra 사용하여 Salesforce 데이터 소스에 연결할 수 있습니다. SalesforceConfiguration

Amazon Kendra Salesforce 커넥터에는 두 가지 버전이 있습니다. 각 버전에 지원되는 기능은 다음과 같습니다.

세일즈포스 커넥터 V1.0/ API SalesforceConfiguration

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터

세일즈포스 커넥터 V2.0/ API TemplateConfiguration

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화

- Virtual Private Cloud(VPC)

> **Note**
>
> 세일즈포스 커넥터 SalesforceConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. Salesforce 커넥터 V2.0/ API로 마이그레이션하거나 이를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra Salesforce 데이터 소스 커넥터 문제 해결에 대한 자세한 내용은 을 참조하십시오. 데이터 소스 문제 해결

주제

- Salesforce 커넥터 V1.0
- Salesforce 커넥터 V2.0

Salesforce 커넥터 V1.0

Salesforce는 지원, 영업 및 마케팅 팀을 관리하기 위한 CRM(고객 관계 관리) 도구입니다. 를 Amazon Kendra 사용하여 Salesforce 표준 개체 및 사용자 지정 개체를 인덱싱할 수 있습니다.

> **Important**
>
> Amazon Kendra 세일즈포스 API 버전 48을 사용합니다. Salesforce API는 하루에 수행할 수 있는 요청의 수를 제한합니다. Salesforce는 해당 요청을 초과할 경우 계속할 수 있을 때까지 재시도합니다.

> **Note**
>
> 세일즈포스 커넥터 SalesforceConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. Salesforce 커넥터 V2.0/ API로 마이그레이션하거나 이를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra Salesforce 데이터 소스 커넥터 문제 해결에 대한 자세한 내용은 을 참조하십시오. 데이터 소스 문제 해결

주제

지원되는 기능

Amazon Kendra Salesforce 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터

필수 조건

Salesforce 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Salesforce 및 계정에서 다음과 같이 변경하십시오. AWS

Salesforce에서 다음을 확인하세요.

- Salesforce 계정을 생성하고 Salesforce에 연결하는 데 사용하는 사용자 이름 및 암호를 기록했습니다.
- OAuth가 활성화된 상태로 Salesforce Connected 앱 계정을 만들고 Salesforce Connected 앱에 할당된 소비자 키(클라이언트 ID)와 소비자 보안 암호(클라이언트 보안 암호)를 복사했습니다. 클라이언트 ID와 클라이언트 암호는 암호에 저장된 인증 자격 증명으로 사용됩니다. AWS Secrets Manager 자세한 내용은 Connected 앱에 대한 Salesforce 설명서를 참조하세요.

  > **Note**
  >
  > 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- Salesforce에 연결하는 데 사용된 계정과 연결된 Salesforce 보안 토큰을 복사했습니다.
- 인덱싱하려는 Salesforce 인스턴스의 URL을 복사했습니다. 일반적으로 이는 *https://<company>.salesforce.com/*입니다. 서버에서 Salesforce Connected 앱을 실행하고 있어야 합니다.

- ReadOnly 프로필을 복제한 다음 모든 데이터 보기 및 문서 관리 권한을 추가하여 Salesforce에 대한 읽기 전용 액세스 권한을 가진 사용자를 위해 Salesforce 서버에 자격 증명을 추가했습니다. 이러한 자격 증명은 연결하는 사용자와 연결하는 Salesforce 연결 앱을 식별합니다. Amazon Kendra
- Salesforce 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 AWS 계정있는지 확인하십시오.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Salesforce 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Salesforce 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. Amazon Kendra API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Salesforce 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 수 있도록 Amazon Kendra Salesforce 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 Salesforce를 구성하지 않은 경우 참조하십시오. Amazon Kendra 필수 조건

Console

세일즈포스에 Amazon Kendra 연결하려면

1. AWS 관리 콘솔에 로그인하고 콘솔을 엽니다.Amazon Kendra
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 소스 추가 페이지에서 Salesforce 커넥터 V1.0을 선택한 다음 커넥터 추가를 선택합니다.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.
   a. 데이터 소스 이름 - 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.
   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.
   c. 기본 언어 - 문서를 필터링하여 인덱스에 사용할 수 있는 언어입니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.
   d. 새 태그 추가 - 리소스를 검색 및 필터링하거나 공유 비용을 추적하기 위한 태그입니다.
   e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.
   a. Salesforce URL - 인덱싱하려는 Salesforce 사이트의 인스턴스 URL을 입력합니다.
   b. 인증 유형에서 기존 인증과 신규 인증 정보 중 하나를 선택하여 Salesforce 보안 인증 정보를 저장합니다. 새 비밀번호를 생성하도록 선택하면 AWS Secrets Manager 비밀 창이 열립니다.
      - AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.
        A. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Salesforce-'가 자동으로 추가됩니다.
        B. 사용자 이름, 암호, 보안 토큰, 소비자 키, 소비자 보안 암호, 인증 URL - Salesforce 계정에서 생성한 보안 인증 값을 입력합니다.

C. 인증 저장을 선택합니다.

c. IAM 역할 —기존 역할을 선택하거나 새 IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 IAM 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

d. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 첨부 파일 크롤링 - 첨부된 객체, 기사, 피드를 모두 크롤링하도록 선택합니다.

b. 표준 객체, 지식 문서 및 Chatter 피드 - 크롤링하려는 Salesforce 개체 또는 콘텐츠 유형을 선택합니다.

> **Note**
>
> 표준 객체, 지식 문서 또는 chatter 피드 중 하나 이상을 인덱싱하기 위한 구성 정보를 제공해야 합니다. 지식 문서를 크롤링하기로 선택한 경우 인덱싱할 지식 문서 유형, 문서 이름, 모든 지식 문서의 표준 필드를 인덱싱할지 또는 사용자 지정 문서 유형의 필드만 인덱싱할지 여부를 지정해야 합니다. 사용자 지정 문서를 인덱싱하도록 선택한 경우 문서 유형의 내부 이름을 지정해야 합니다. 문서 유형은 최대 10개까지 지정할 수 있습니다.

c. 빈도 Amazon Kendra - 데이터 원본과 동기화하는 빈도

d. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 표준 자료 문서, 표준 개체 첨부 파일 및 추가 제안 필드 매핑의 경우 Amazon Kendra —생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

> **Note**
>
> `_document_body`에 대한 인덱스 매핑이 필요합니다. `Salesforce ID` 필드와 Amazon Kendra `_document_id` 필드 간의 매핑은 변경할 수 없습니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Salesforce에 연결하려면 Amazon Kendra

다음과 같은 API를 지정해야 합니다. SalesforceConfiguration

- 서버 URL - 인덱싱하려는 Salesforce 사이트의 인스턴스 URL입니다.
- 비밀 Amazon 리소스 이름 (ARN) —Salesforce 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공하십시오. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "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 역할 —역할을 `CreateDataSource` 호출하여 Secrets Manager 비밀에 액세스할 권한을 제공하고 IAM Salesforce 커넥터 및 에 필요한 공개 API를 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Salesforce 데이터 소스에 대한IAM 역할을 참조하세요.
- 표준 객체, 지식 문서 또는 chatter 피드 중 하나 이상을 인덱싱하기 위한 구성 정보를 제공해야 합니다.
  - 표준 객체 - 표준 객체를 크롤링하도록 선택한 경우 표준 객체 이름과 문서 내용이 포함된 표준 객체 테이블의 필드 이름을 지정해야 합니다.

- 지식 문서 - 지식 문서를 크롤링하기로 선택한 경우 인덱싱할 지식 문서 유형, 인덱싱할 지식 문서의 상태, 모든 지식 문서의 표준 필드를 인덱싱할지 또는 사용자 지정 문서 유형의 필드만 인덱싱할지 여부를 지정해야 합니다.
- Chatter 피드 - Chatter 피드를 크롤링하기로 선택한 경우 인덱싱할 콘텐츠가 포함된 Salesforce 테이블의 열 이름을 지정해야 합니다. FeedItem

다음 선택적 기능도 추가할 수 있습니다.

- 포함 및 제외 필터 - 특정 파일 첨부 파일을 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 필드 매핑 - Salesforce 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. Amazon Kendra 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서의 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

Salesforce 커넥터 V2.0

Salesforce는 지원, 영업 및 마케팅 팀을 관리하기 위한 CRM(고객 관계 관리) 도구입니다. 를 Amazon Kendra 사용하여 Salesforce 표준 개체 및 사용자 지정 개체를 인덱싱할 수 있습니다.

Amazon Kendra Salesforce 데이터 소스 커넥터는 개발자 에디션 및 엔터프라이즈 에디션과 같은 Salesforce 에디션을 지원합니다.

> **Note**
>
> 세일즈포스 커넥터 SalesforceConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. Salesforce 커넥터 V2.0/ API로 마이그레이션하거나 이를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra Salesforce 데이터 소스 커넥터 문제 해결에 대한 자세한 내용은 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Salesforce 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Salesforce 데이터 원본을 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Salesforce 및 계정에서 다음과 같이 변경하십시오. AWS

Salesforce에서 다음을 확인하세요.

- Salesforce 관리자 계정을 생성하고 Salesforce에 연결하는 데 사용하는 사용자 이름 및 암호를 기록했습니다.
- Salesforce에 연결하는 데 사용된 계정과 연결된 Salesforce 보안 토큰을 복사했습니다.
- OAuth가 활성화된 상태로 Salesforce Connected 앱 계정을 만들고 Salesforce Connected 앱에 할당된 소비자 키(클라이언트 ID)와 소비자 보안 암호(클라이언트 보안 암호)를 복사했습니다. 클라이언트 ID와 클라이언트 암호는 암호에 저장된 인증 자격 증명으로 사용됩니다. AWS Secrets Manager 자세한 내용은 Connected 앱에 대한 Salesforce 설명서를 참조하세요.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 인덱싱하려는 Salesforce 인스턴스의 URL을 복사했습니다. 일반적으로 이는 *https://<company>.salesforce.com/*입니다. 서버에서 Salesforce Connected 앱을 실행하고 있어야 합니다.
- ReadOnly 프로필을 복제한 다음 모든 데이터 보기 및 문서 관리 권한을 추가하여 Salesforce에 대한 읽기 전용 액세스 권한을 가진 사용자를 위해 Salesforce 서버에 자격 증명을 추가했습니다. 이러한 자격 증명은 연결하는 사용자와 연결하는 Salesforce 연결 앱을 식별합니다. Amazon Kendra
- Salesforce 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 AWS 계정있는지 확인하십시오.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Salesforce 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Salesforce 데이터 원본을 연결할 때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. Amazon Kendra API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Salesforce 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 수 있도록 Amazon Kendra Salesforce 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 Salesforce를 구성하지 않은 경우 참조하십시오. Amazon Kendra 필수 조건

Console

Salesforce에 Amazon Kendra 연결하려면:

1. 에 AWS Management Console 로그인하고 콘솔을 엽니다.Amazon Kendra
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Salesforce 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Salesforce 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 색인을 기준으로 문서를 필터링할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. Salesforce URL - 인덱싱하려는 Salesforce 사이트의 인스턴스 URL을 입력합니다.

b. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

c. 기존 암호를 입력하거나 새 암호를 만들면 AWS Secrets Manager 보안 암호 창이 열립니다.

- 인증 - 암호 만들기 창에 다음 정보를 입력합니다. AWS Secrets Manager

A. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Salesforce-'가 자동으로 추가됩니다.

B. 사용자 이름, 암호, 보안 토큰, 소비자 키, 소비자 보안 암호, 인증 URL - Salesforce 계정에서 생성하고 다운로드한 보안 인증 값을 입력합니다.

> **Note**
>
> 세일즈포스 개발자 에디션을 사용하는 경우 내 도메인 로그인 URL (예: ***https://.my.salesforce.com***) 을 인증 URL로 사용하십시오**https://login.salesforce.com/services/oauth2/token**. MyCompany ***##### #### #### #### ## ## URL# https://test.salesforce.com/services/oauth2/token ## # ## # ### URL (#: -.sandbox.my.salesforce.com) # ######. MyDomainName SandboxName***

C. 인증 저장을 선택합니다.

d. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

e. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

f. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

g. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 첨부 파일 크롤링 - 첨부된 Salesforce 객체를 모두 크롤링하도록 선택합니다.

b. 표준 객체, 첨부 파일이 있는 표준 객체 및 첨부 파일 없는 표준 객체, 지식 문서 - 크롤링하려는 Salesforce 개체 또는 콘텐츠 유형을 선택합니다.

c. 표준 객체, 지식 문서 또는 chatter 피드 중 하나 이상을 인덱싱하기 위한 구성 정보를 제공해야 합니다. 지식 문서를 크롤링하기로 선택한 경우 인덱싱할 지식 문서 유형을 지정해야 합니다. 게시, 보관, 초안 및 첨부 파일을 선택할 수 있습니다.

정규식 필터 - 특정 카탈로그 항목을 포함하도록 정규식 패턴을 지정합니다.

8. 추가 구성의 경우:

- ACL 정보 모든 액세스 제어 목록이 기본적으로 포함됩니다. 액세스 제어 목록을 선택 취소하면 해당 범주의 모든 파일이 공개됩니다.
- 정규식 패턴 - 특정 페이지 및 자산을 포함하거나 제외하기 위한 정규 표현식 패턴을 추가합니다. 최대 100개의 패턴을 추가할 수 있습니다.

동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

9. 다음을 선택합니다.
10. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

    a. 표준 참조 문서, 표준 개체 첨부 파일 및 추가 제안 필드 매핑의 경우 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

    > **Note**
    >
    > _document_body에 대한 인덱스 매핑이 필요합니다. Salesforce ID 필드와 Amazon Kendra _document_id 필드 간의 매핑은 변경할 수 없습니다. 모든 Salesforce 필드를 문서 제목 또는 문서 본문 Amazon Kendra 예약/기본 색인 필드에 매핑할 수 있습니다.
    > Salesforce 필드를 Amazon Kendra 문서 제목 및 문서 본문 필드에 매핑하는 경우 Amazon Kendra는 검색 응답에서 문서 제목 및 본문 필드의 데이터를 사용합니다.

    b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

    c. 다음을 선택합니다.
11. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

세일즈포스에 연결하려면 Amazon Kendra

API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. TemplateConfiguration 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 SALESFORCEV2 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 호스트 URL - Salesforce 인스턴스 호스트 URL을 지정합니다.
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) —Salesforce 계정의 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공하십시오. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "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 역할 —역할을 CreateDataSource 호출하여 Secrets Manager 비밀에 액세스할 권한을 제공하고 IAM Salesforce 커넥터 및 에 필요한 공개 API를 호출할 RoleArn 시기를 지정합니다. Amazon Kendra자세한 내용은 Salesforce 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - CreateDataSource를 호출할 때 VpcConfiguration을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 특정 문서, 계정, 캠페인, 사례, 연락처, 잠재 고객, 기회, 솔루션, 작업, 그룹, 채팅 및 사용자 지정 개체 파일을 포함할지 또는 제외할지 여부를 지정할 수 있습니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 필드 매핑 - Salesforce 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 _document_body 합니다. 다른 모든 필드는 선택 사항입니다.

> **Note**
>
> _document_body에 대한 인덱스 매핑이 필요합니다. Salesforce ID 필드와 Amazon Kendra _document_id 필드 간의 매핑은 변경할 수 없습니다. 모든 Salesforce 필드를 문서 제목 또는 문서 본문 Amazon Kendra 예약/기본 색인 필드에 매핑할 수 있습니다.
> Salesforce 필드를 Amazon Kendra 문서 제목 및 문서 본문 필드에 매핑하는 경우 Amazon Kendra는 검색 응답에서 문서 제목 및 본문 필드의 데이터를 사용합니다.

구성해야 할 기타 중요한 JSON 키 목록은 Salesforce 템플릿 스키마를 참조하십시오.

자세히 알아보기

Salesforce 데이터 Amazon Kendra 소스와의 통합에 대한 자세한 내용은 다음을 참조하십시오.

- 에 대한 업데이트된 Salesforce 커넥터 (V2) 발표 Amazon Kendra

ServiceNow

ServiceNow IT 서비스, 티켓 시스템 및 지원과 같은 조직 수준의 워크플로를 만들고 관리할 수 있는 클라우드 기반 서비스 관리 시스템을 제공합니다. ServiceNow 카탈로그, 기술 문서, 인시던트 및 첨부 파일을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

Amazon Kendra 콘솔, TemplateConfigurationAPI 또는 API를 Amazon Kendra 사용하여 ServiceNow 데이터 원본에 연결할 수 있습니다. ServiceNowConfiguration

Amazon Kendra ServiceNow 커넥터에는 두 가지 버전이 있습니다. 각 버전에 지원되는 기능은 다음과 같습니다.

ServiceNow 커넥터 V1.0/ API ServiceNowConfiguration

- 필드 매핑
- ServiceNow 인스턴스 버전: 런던, 기타
- 포함/제외 필터

ServiceNow 커넥터 V2.0/API TemplateConfiguration

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- ServiceNow 인스턴스 버전: 로마, 샌디에이고, 도쿄, 기타
- Virtual Private Cloud(VPC)

> **Note**
>
> ServiceNow 커넥터 ServiceNowConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. 커넥터 V2.0/ API로 마이그레이션하거나 ServiceNow 커넥터 V2.0/API를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra ServiceNow 데이터 소스 커넥터 문제 해결은 을 참조하십시오. 데이터 소스 문제 해결

주제

- ServiceNow 커넥터 V1.0
- ServiceNow 커넥터 V2.0
- 쿼리로 인덱싱할 문서 지정

ServiceNow 커넥터 V1.0

ServiceNow IT 서비스, 티켓 시스템 및 지원과 같은 조직 수준의 워크플로를 만들고 관리할 수 있는 클라우드 기반 서비스 관리 시스템을 제공합니다. ServiceNow 카탈로그, 기술 문서 및 Amazon Kendra 첨부 파일을 인덱싱하는 데 사용할 수 있습니다.

> **Note**
>
> ServiceNow 커넥터 ServiceNowConfiguration V1.0/API에 대한 지원은 2023년에 종료될 예정입니다. 커넥터 V2.0/ API로 마이그레이션하거나 ServiceNow 커넥터 V2.0/API를 사용하는 것이 좋습니다. TemplateConfiguration

Amazon Kendra ServiceNow 데이터 소스 커넥터 문제 해결은 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra ServiceNow 데이터 소스 커넥터는 다음 기능을 지원합니다.

- ServiceNow 인스턴스 버전: 런던, 기타
- 포함/제외 패턴: 서비스 카탈로그, 지식 문서, 첨부 파일

필수 조건

를 사용하여 ServiceNow 데이터 원본을 Amazon Kendra 인덱싱하려면 먼저 ServiceNow 및 AWS 계정에서 다음과 같이 변경하십시오.

ServiceNow에서 다음이 있는지 확인하세요.

- ServiceNow 관리자 계정을 만들고 ServiceNow 인스턴스를 생성했습니다.
- ServiceNow 인스턴스 URL의 호스트를 복사했습니다. 예를 들어 인스턴스의 URL이 *https://your-domain.service-now.com*인 경우 입력하는 호스트 URL의 형식은 *your-domain.service-now.com*입니다.
- ServiceNow 인스턴스에 연결할 수 있는 사용자 이름과 암호가 포함된 기본 인증 자격 증명을 기록해 두었습니다. Amazon Kendra

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 선택 사항: 사용자 이름, 암호, 클라이언트 ID Amazon Kendra 및 클라이언트 암호를 식별하고 생성할 수 있는 OAuth 2.0 자격 증명 토큰을 구성했습니다. 사용자 이름과 암호는 ServiceNow 지식 기반 및 서비스 카탈로그에 대한 액세스를 제공해야 합니다. 자세한 내용은 OAuth 2.0 인증 ServiceNow 설명서를 참조하십시오.
- 다음 권한을 추가했습니다.
  - kb_category
  - kb_knowledge
  - kb_knowledge_base
  - kb_uc_cannot_read_mtom
  - kb_uc_can_read_mtom
  - sc_catalog
  - sc_category
  - sc_cat_item
  - 시스템_어태치먼트
  - sys_attachment_doc
  - sys_user_role
- 동일한 색인에 사용할 다른 데이터 소스에서 각 문서가 고유한지 확인했습니다. ServiceNow 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하세요. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- ServiceNow 인증 자격 증명을 AWS Secrets Manager 비밀에 저장하고 API를 사용하는 경우 비밀의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 ServiceNow 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

ServiceNow 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 ServiceNow 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 ServiceNow 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 ServiceNow

1. AWS 관리 콘솔에 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 ServiceNow커넥터 V1.0을 선택한 다음 데이터 원본 추가를 선택합니다.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. ServiceNow 호스트 - 호스트 URL을 입력합니다. ServiceNow

b. ServiceNow 버전 —버전을 선택합니다. ServiceNow

c. 사용 사례에 따라 기본 인증, OAuth 2.0 인증 중에서 선택합니다.

d. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 ServiceNow 인증 Secrets Manager 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

i. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-ServiceNow -'가 자동으로 추가됩니다.

ii. 기본 인증을 사용하는 경우 - 계정의 비밀 이름, 사용자 이름, 비밀번호를 입력합니다. ServiceNow

OAuth2 인증을 사용하는 경우 - 계정에서 생성한 비밀 이름, 사용자 이름, 암호, 클라이언트 ID, 클라이언트 암호를 입력합니다. ServiceNow

iii. 저장 및 보안 암호 추가를 선택합니다.

e. IAM 역할 - 기존 역할을 선택하거나 새 IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 IAM 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

f. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 지식 문서 포함 - 지식 문서를 인덱싱하도록 선택합니다.

b. 참조 문서 유형 - 사용 사례에 따라 공개 문서만 포함과 ServiceNow 필터 쿼리를 기반으로 문서 포함 중에서 선택합니다. ServiceNow 필터 쿼리 기반 문서 포함을 선택하는 경우 ServiceNow 계정에서 복사한 필터 쿼리를 입력해야 합니다.

c. 지식 문서 첨부 파일 포함 - 지식 문서 첨부 파일을 인덱싱하도록 선택합니다. 인덱싱할 특정 파일 형식을 선택할 수도 있습니다.

d. 카탈로그 항목 포함 - 카탈로그 항목을 인덱싱하도록 선택합니다.

e. 카탈로그 항목 첨부 파일 포함 - 카탈로그 항목 첨부 파일을 인덱싱하려면 선택합니다. 인덱싱할 특정 파일 형식을 선택할 수도 있습니다.

f. 빈도 Amazon Kendra - 데이터 원본과 동기화하는 빈도.

g. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 참조 문서 및 서비스 카탈로그 - Amazon Kendra 생성된 기본 데이터 원본 필드와 색인에 매핑하려는 추가 제안 필드 매핑 중에서 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra ServiceNow

ServiceNowConfiguration API를 사용하여 다음을 지정해야 합니다.

- 데이터 소스 URL ServiceNow —URL을 지정합니다. 호스트 엔드포인트는 *your-domain.service-now.com*과 같은 형식이어야 합니다.
- 데이터 원본 호스트 인스턴스 ServiceNow —호스트 인스턴스 버전을 또는 LONDON 로 지정합니다. OTHERS
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. ServiceNow

  기본 인증을 사용하는 경우 보안 암호는 다음 키가 있는 JSON 구조에 저장됩니다.

```
{
    "username": "user name",
    "password": "password"
}
```

OAuth2 인증을 사용하는 경우 보안 암호는 다음 키가 있는 JSON 구조에 저장됩니다.

```
{
    "username": "user name",
    "password": "password",
    "clientId": "client id",
    "clientSecret": "client secret"
}
```

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 수 있는 권한을 IAM 역할을 제공하고 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. ServiceNow Amazon Kendra자세한 내용은 ServiceNow 데이터 원본의IAM 역할을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- 필드 매핑 - ServiceNow 데이터 원본 필드를 인덱스 필드에 매핑하도록 Amazon Kendra 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

- 포함 및 제외 필터 - 카탈로그와 지식 문서의 특정 파일 첨부 파일을 포함할지 또는 제외할지 여부를 지정합니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함

필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 인덱싱 파라미터 - 다음을 수행할지 여부를 지정할 수도 있습니다.
  - 지식 문서 및 서비스 카탈로그 또는 이 두 가지를 모두 인덱싱합니다. 참조 문서 및 서비스 카탈로그 항목을 색인화하기로 선택한 경우 색인의 색인 문서 콘텐츠 ServiceNow 필드에 매핑되는 필드의 Amazon Kendra 이름을 제공해야 합니다.
  - 지식 문서 및 카탈로그 항목에 대한 첨부 파일을 인덱싱합니다.
  - 하나 이상의 지식 기반에서 문서를 선택하는 ServiceNow 쿼리를 사용하십시오. 지식 기반은 공개 또는 비공개일 수 있습니다. 자세한 내용을 알아보려면 쿼리로 인덱싱할 문서 지정을 참조하세요.

자세히 알아보기

ServiceNow 데이터 Amazon Kendra 소스와의 통합에 대해 자세히 알아보려면 다음을 참조하십시오.

- Amazon Kendra ServiceNow 온라인 커넥터 시작하기

ServiceNow 커넥터 V2.0

ServiceNow IT 서비스, 티켓 시스템 및 지원과 같은 조직 수준의 워크플로를 만들고 관리할 수 있는 클라우드 기반 서비스 관리 시스템을 제공합니다. ServiceNow 카탈로그, 기술 문서, 인시던트 및 첨부 파일을 Amazon Kendra 인덱싱하는 데 사용할 수 있습니다.

Amazon Kendra ServiceNow 데이터 소스 커넥터 문제 해결에 대한 자세한 내용은 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra ServiceNow 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화
- ServiceNow 인스턴스 버전: 로마, 샌디에이고, 도쿄, 기타
- Virtual Private Cloud(VPC)

필수 조건

를 사용하여 ServiceNow 데이터 원본을 Amazon Kendra 인덱싱하려면 먼저 AND AWS 계정에서 다음과 같이 변경하십시오 ServiceNow .

ServiceNow에서 다음이 있는지 확인하세요.

- 개인용 또는 엔터프라이즈 개발자 인스턴스를 생성하고 관리자 역할을 가진 ServiceNow 인스턴스를 보유했습니다.
- ServiceNow 인스턴스 URL의 호스트를 복사했습니다. 입력하는 호스트 URL의 형식은 *your-domain.service-now.com*입니다. 연결하려면 ServiceNow 인스턴스 URL이 필요합니다 Amazon Kendra.
- ServiceNow 인스턴스에 연결할 수 있도록 Amazon Kendra 사용자 이름과 비밀번호의 기본 인증 자격 증명을 기록해 두었습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 선택 사항: 사용자 이름, 암호, 생성된 클라이언트 ID 및 클라이언트 암호를 Amazon Kendra 사용하여 식별할 수 있는 구성된 OAuth 2.0 클라이언트 자격 증명. 자세한 내용은 OAuth 2.0 인증 ServiceNow 설명서를 참조하십시오.
- 동일한 색인에 사용할 다른 데이터 소스에서 각 문서가 고유한지 확인했습니다. ServiceNow 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 있는지 확인하세요. AWS 계정

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- ServiceNow 인증 자격 증명을 AWS Secrets Manager 비밀에 저장하고 API를 사용하는 경우 비밀의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 ServiceNow 데이터 원본을 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

ServiceNow 데이터 Amazon Kendra 원본에 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 ServiceNow 데이터 원본의 필수 세부 정보를 제공해야 합니다. 아직 구성하지 않은 경우 ServiceNow 을 Amazon Kendra 참조하십시오필수 조건.

Console

Amazon Kendra 연결하려면 ServiceNow

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

> **Note**
>
> 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 커넥터를 선택한 다음 ServiceNow 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 ServiceNow 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

    a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

    b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

    c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

    d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

    e. 다음을 선택합니다.
6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

    a. ServiceNow 호스트 - 호스트 URL을 입력합니다. ServiceNow 입력하는 호스트 URL의 형식은 *your-domain.service-now.com*입니다.

    b. ServiceNow 버전 - ServiceNow 인스턴스 버전을 선택합니다. 로마, 샌디에이고, 도쿄 또는 기타 중에서 선택할 수 있습니다.

    c. 승인 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

    d. 인증 - 기본 인증과 Oauth 2.0 인증 중에서 선택합니다.

    e. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 인증 자격 Secrets Manager 증명을 저장합니다. ServiceNow 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다. 창에 다음 정보를 입력합니다.

        i. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-ServiceNow -'가 자동으로 추가됩니다.

        ii. 기본 인증을 사용하는 경우 - 계정의 비밀 이름, 사용자 이름, 비밀번호를 입력합니다. ServiceNow

OAuth2.0 인증을 사용하는 경우 - 계정에서 생성한 비밀 이름, 사용자 이름, 암호, 클라이언트 ID, 클라이언트 암호를 입력합니다. ServiceNow

iii. 비밀번호를 저장하고 추가하세요.

f. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

g. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

h. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

i. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 지식 문서의 경우 다음 옵션 중에서 선택합니다.

- 지식 문서 - 지식 문서를 인덱싱하도록 선택합니다.
- 지식 문서 첨부 파일 - 지식 문서 첨부 파일을 인덱싱하도록 선택합니다.
- 참조 문서 유형 - 사용 사례에 따른 ServiceNow 필터 쿼리를 기반으로 공개 문서만과 참조 문서 중에서 선택합니다. ServiceNow 필터 쿼리 기반 문서 포함을 선택하는 경우 ServiceNow 계정에서 복사한 필터 쿼리를 입력해야 합니다. 필터 쿼리의 예: `workflow_state=draft^EQ`, `kb_knowledge_base=dfc19531bf2021003f07e2c1ac0739ab^textISNOTEMPTY^EQ`, `article_type=text^active=true^EQ`.

> **Important**
>
> 공개 문서만 크롤링하도록 선택하면 공개 액세스 역할이 할당된 참조 문서만 Amazon Kendra 크롤링합니다. ServiceNow

- 간단한 설명 필터를 기반으로 문서 포함 - 특정 문서를 포함하거나 제외하도록 정규 표현식 패턴을 지정합니다.

b. 서비스 카탈로그 항목의 경우:

- 서비스 카탈로그 항목 - 서비스 카탈로그 항목을 인덱싱하도록 선택합니다.
- 서비스 카탈로그 항목 첨부 파일 - 서비스 카탈로그 항목 첨부 파일을 인덱싱하려면 선택합니다.
- 활성 서비스 카탈로그 항목 - 활성 서비스 카탈로그 항목을 인덱싱하도록 선택합니다.
- 비활성 서비스 카탈로그 항목 - 비활성 서비스 카탈로그 항목을 인덱싱하도록 선택합니다.
- 필터 쿼리 - 인스턴스에 정의된 필터를 기반으로 서비스 카탈로그 항목을 포함하도록 선택합니다. ServiceNow 필터 쿼리의 예: *short_descriptionLIKEAccess^category=2809952237b1300054b6a3549dbe5dd4^nameSTARTSWITHService^active=true^EQ*.
- 간단한 설명 필터를 기반으로 서비스 카탈로그 항목 포함 - 특정 카탈로그 항목을 포함하도록 정규식 패턴을 지정합니다.

c. 인시던트의 경우:

- 인시던트 - 서비스 인시던트를 인덱싱하도록 선택합니다.
- 인시던트 첨부 파일 - 인시던트 첨부 파일을 인덱싱하도록 선택합니다.
- 활성 인시던트 - 활성 인시던트를 인덱싱하도록 선택합니다.
- 비활성 인시던트 - 비활성 인시던트를 인덱싱하도록 선택합니다.
- 활성 인시던트 유형 - 사용 사례에 따라 모든 인시던트, 미해결 인시던트, 미해결 - 미할당 인시던트, 해결된 인시던트 중에서 선택합니다.
- 필터 쿼리 - 인스턴스에 정의된 필터를 기반으로 인시던트를 포함하도록 ServiceNow 선택합니다. 필터 쿼리의 예: *short_descriptionLIKETest^urgency=3^state=1^EQ*, *priority=2^category=software^EQ* .

- 간단한 설명 필터를 기반으로 인시던트 포함 - 특정 인시던트를 포함하도록 정규식 패턴을 지정합니다.

d. 추가 구성의 경우:

- ACL 정보 - 선택한 개체의 액세스 제어 목록이 기본적으로 포함됩니다. 액세스 제어 목록을 선택 취소하면 해당 범주의 모든 파일이 공개됩니다. 선택하지 않은 개체의 경우 ACL 옵션이 자동으로 비활성화됩니다. 공개 문서의 경우 ACL이 적용되지 않습니다.
- 최대 파일 크기 — Amazon Kendra가 크롤링할 파일 크기 제한을 MB 단위로 지정합니다. Amazon Kendra는 사용자가 정의한 크기 제한 내에 있는 파일만 크롤링합니다. 기본 파일 크기는 50MB입니다. 최대 파일 크기는 0MB보다 크고 50MB보다 작거나 같아야 합니다.
- 첨부 정규식 패턴 - 카탈로그, 지식 문서 및 인시던트의 특정 첨부 파일을 포함하거나 제외하는 정규 표현식 패턴을 추가합니다. 최대 100개의 패턴을 추가할 수 있습니다.

e. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

f. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

g. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

a. 기본 필드 매핑 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

연결하려면 Amazon Kendra ServiceNow

TemplateConfigurationAPI를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 SERVICENOWV2 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 호스트 URL - ServiceNow 호스트 인스턴스 버전을 지정합니다. 예: *your-domain.service-now.com*.
- 인증 유형 basicAuth —인스턴스용이든 OAuth2 관계없이 사용하는 인증 유형을 지정합니다. ServiceNow
- ServiceNow 인스턴스 버전 - 사용하는 ServiceNow 인스턴스 (, Tokyo SandiegoRome, 또는) 를 지정합니다. Others
- 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - 계정에서 생성한 인증 자격 증명이 포함된 Secrets Manager 암호의 Amazon 리소스 이름 (ARN) 을 제공합니다. ServiceNow

  기본 인증을 사용하는 경우 보안 암호는 다음 키가 있는 JSON 구조에 저장됩니다.

```
{
    "username": "user name",
```

```
    "password": "password"
}
```

- OAuth2 클라이언트 자격 증명을 사용하는 경우 암호는 다음 키와 함께 JSON 구조에 저장됩니다.

```
{
    "username": "user name",
    "password": "password",
    "clientId": "client id",
    "clientSecret": "client secret"
}
```

- IAM 역할 - `CreateDataSource` Secrets Manager 비밀번호에 액세스할 수 있는 권한을 IAM 역할에 제공하고 커넥터 및 에 필요한 공개 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. ServiceNow Amazon Kendra자세한 내용은 ServiceNow 데이터 원본의IAM 역할을 참조하십시오.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 포함 및 제외 필터 - 지식 문서, 서비스 카탈로그 및 인시던트의 파일 이름과 파일 유형을 사용하여 특정 첨부 파일을 포함할지 또는 제외할지 여부를 지정할 수 있습니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 색인화할 특정 문서 - ServiceNow 쿼리를 사용하여 개인 지식 베이스를 비롯한 하나 이상의 지식 기반에서 원하는 문서를 지정할 수 있습니다. 지식 기반에 대한 액세스 권한은 ServiceNow 인스턴스에 연결하는 데 사용하는 사용자에 따라 결정됩니다. 자세한 내용을 알아보려면 쿼리로 인덱싱할 문서 지정을 참조하세요.
- 인덱싱 파라미터 - 다음을 수행할지 여부를 지정할 수도 있습니다.

- 지식 문서, 서비스 카탈로그, 인시던트 또는 이 세 가지를 모두 인덱싱합니다. 참조 자료, 서비스 카탈로그 항목 및 인시던트를 색인화하기로 선택한 경우 색인의 색인 문서 내용 ServiceNow 필드에 매핑되는 필드의 이름을 제공해야 합니다. Amazon Kendra
- 지식 문서 및 서비스 카탈로그 항목과 인시던트에 대한 첨부 파일을 인덱싱합니다.
- `short description` 필터 패턴에 따라 지식 문서, 서비스 카탈로그 항목 및 인시던트를 포함합니다.
- 활성 및 비활성 서비스 카탈로그 항목 및 인시던트를 필터링하도록 선택합니다.
- 인시던트 유형을 기준으로 인시던트를 필터링하도록 선택합니다.
- ACL을 크롤링해야 하는 개체를 선택합니다.
- ServiceNow 쿼리를 사용하여 비공개 지식 베이스를 비롯한 하나 이상의 지식 기반에서 원하는 문서를 지정할 수 있습니다. 지식 기반에 대한 액세스 권한은 ServiceNow 인스턴스에 연결하는 데 사용하는 사용자에 따라 결정됩니다. 자세한 내용을 알아보려면 쿼리로 인덱싱할 문서 지정을 참조하세요.

- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 필드 매핑 - ServiceNow 데이터 원본 필드를 인덱스 필드에 매핑하도록 선택합니다. Amazon Kendra 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문이 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 ServiceNow 템플릿 스키마를 참조하십시오.

자세히 알아보기

ServiceNow 데이터 Amazon Kendra 소스와의 통합에 대해 자세히 알아보려면 다음을 참조하십시오.

- 업데이트된 ServiceNow 커넥터 (V2) Amazon Kendra발표 시작하기 Amazon Kendra

쿼리로 인덱싱할 문서 지정

ServiceNow 쿼리를 사용하여 Amazon Kendra 색인에 포함하려는 문서를 지정할 수 있습니다. 쿼리를 사용할 때 비공개 지식 기반을 비롯한 여러 지식 기반을 지정할 수 있습니다. 지식창고 액세스 권한은 ServiceNow 인스턴스에 연결하는 데 사용하는 사용자에 따라 결정됩니다.

쿼리를 작성하려면 ServiceNow 쿼리 빌더를 사용합니다. 빌더를 사용하여 쿼리를 만들고 쿼리가 올바른 문서 목록을 반환하는지 테스트할 수 있습니다.

ServiceNow 콘솔을 사용하여 쿼리를 만들려면

1. ServiceNow 콘솔에 로그인합니다.
2. 왼쪽 메뉴에서 지식, 문서를 차례로 선택한 다음 모두를 선택합니다.
3. 페이지 상단에서 필터 아이콘을 선택합니다.
4. 쿼리 빌더를 사용하여 쿼리를 생성합니다.
5. 쿼리가 완료되면 쿼리를 마우스 오른쪽 버튼으로 클릭하고 쿼리 복사를 선택하여 쿼리 빌더에서 쿼리를 복사합니다. 에서 사용하려면 이 쿼리를 Amazon Kendra저장하세요.

쿼리를 복사할 때 쿼리 파라미터를 변경하지 않도록 하세요. 쿼리 매개 변수가 인식되지 않는 경우 매개 변수를 빈 것으로 ServiceNow 처리하고 결과를 필터링하는 데 사용하지 않습니다.

Slack

Slack은 사용자가 다양한 공개 및 비공개 채널을 통해 메시지와 첨부 파일을 보낼 수 있는 엔터프라이즈 커뮤니케이션 앱입니다. Slack 공개 및 비공개 채널, 봇 및 아카이브 메시지, 파일 및 첨부 파일, 다

이렉트 및 그룹 메시지를 인덱싱하는 데 사용할 Amazon Kendra 수 있습니다. 필터링할 특정 콘텐츠를 선택할 수도 있습니다.

> **Note**
>
> Amazon Kendra 이제 업그레이드된 Slack 커넥터를 지원합니다.
> 콘솔이 자동으로 업그레이드되었습니다. 콘솔에서 새로 만드는 모든 커넥터는 업그레이드된 아키텍처를 사용합니다. API를 사용하는 경우 이제 TemplateConfiguration개체 대신 개체를 사용하여 커넥터를 구성해야 합니다. `SlackConfiguration`
> 이전 콘솔과 API 아키텍처를 사용하여 구성된 커넥터는 구성된 대로 계속 작동합니다. 하지만 편집하거나 업데이트할 수는 없습니다. 커넥터 구성을 편집하거나 업데이트하려면 새 커넥터를 만들어야 합니다.
> 커넥터 워크플로를 업그레이드된 버전으로 마이그레이션하는 것이 좋습니다. 이전 아키텍처를 사용하여 구성된 커넥터에 대한 지원은 2024년 6월에 종료될 예정입니다.

Amazon Kendra 콘솔 또는 API를 Amazon Kendra 사용하여 Slack 데이터 소스에 연결할 수 있습니다. TemplateConfiguration

Amazon Kendra Slack 데이터 소스 커넥터의 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Slack 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 전체 및 점진적 콘텐츠 동기화

- Virtual Private Cloud(VPC)

필수 조건

Slack 데이터 소스를 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Slack과 계정에서 다음과 같이 변경하세요. AWS

Slack에서 다음 사항을 갖추었는지 확인하세요.

- 슬랙 봇 사용자 OAuth 토큰 또는 슬랙 사용자 OAuth 토큰을 구성했습니다. 두 토큰 중 하나를 선택하여 Slack 데이터 소스에 연결할 Amazon Kendra 수 있습니다. 인증 자격 증명으로 사용하려면 토큰이 필요합니다. 자세한 내용은 액세스 토큰에 대한 Slack 설명서를 참조하세요.

> **Note**
>
> 봇 토큰을 Slack 보안 인증의 일부로 사용하는 경우 다이렉트 메시지 및 그룹 메시지를 인덱싱할 수 없으며 인덱싱하려는 채널에 봇 토큰을 추가해야 합니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- Slack 워크스페이스 메인 페이지 URL에서 Slack 워크스페이스 팀 ID를 기록해 두었습니다. 예: *https://app.slack.com/client/T0123456789/...* 여기에서 *T0123456789*는 팀 ID입니다.
- 다음과 같은 Oauth 범위/권한이 추가되었습니다.

사용자 토큰 범위	봇 토큰 범위
• channels:history • channels:read • emoji:read • files:read • groups:history	• channels:history • 채널:관리 • channels:read • 대화. 연결: 관리 • 대화. 연결: 읽기

사용자 토큰 범위	봇 토큰 범위
• groups:read • im:history • im:read • mpim:history • mpim:read • team:read • users.profile:read • users:read • 사용자:읽기.이메일	• files:read • groups:history • groups:read • im:history • im:read • mpim:history • mpim:read • 반응: 읽기 • team:read • usergroups:read • users.profile:read • users:read • 사용자:읽기. 이메일

- Slack 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

귀하의 문서에 다음이 있는지 확인하세요 AWS 계정.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Slack 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 암호가 없는 경우 Slack 데이터 소스를 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 암호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Slack 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Slack 데이터 소스의 필수 세부 정보를 제공해야 합니다. 아직 Slack을 구성하지 않은 경우 을 Amazon Kendra참조하십시오. 필수 조건

Console

Slack에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Slack 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2(해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Slack 커넥터를 선택하십시오.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

   c. 기본 언어 - 문서에서 색인을 필터링할 언어를 선택합니다. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. Slack 워크스페이스의 경우 팀 ID - Slack 워크스페이스의 팀 ID입니다. 팀 ID는 Slack 워크스페이스 메인 페이지 URL에서 찾을 수 있습니다. 예: *https://app.slack.com/client/T0123456789/...* 여기에서 *T0123456789*는 팀 ID입니다.

b. 승인 — ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록 (ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

c. AWS Secrets Manager 비밀 - 기존 암호를 선택하거나 새 암호를 생성하여 Slack Secrets Manager 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름. 암호 이름에 접두사 AmazonKendra '-Slack-'가 자동으로 추가됩니다.

B. Slack 토큰의 경우 - Slack을 구성한 인증 자격 증명 값을 입력합니다.

ii. 비밀번호를 저장하고 추가하세요.

d. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

e. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

f. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

g. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 콘텐츠 유형 선택 - 크롤링하려는 Slack 엔티티 또는 콘텐츠 형식을 선택합니다. 모든 채널, 공개 채널, 비공개 채널, 그룹 메시지, 비공개 메시지 중에서 선택할 수 있습니다.

b. 크롤링 시작일 선택 —콘텐츠 크롤링을 시작하려는 날짜를 입력합니다.

c. 추가 구성 - 봇 및 보관된 메시지를 포함하도록 선택하고 정규 표현식 패턴을 사용하여 특정 콘텐츠를 포함하거나 제외합니다.

> **Note**
>
> 채널 ID와 채널 이름 모두에 포함하도록 선택하면 Amazon Kendra Slack 커넥터는 채널 이름보다 채널 ID를 우선시합니다.
> 특정 비공개 및 그룹 메시지를 포함하도록 선택한 경우 Amazon Kendra Slack 커넥터는 모든 비공개 및 그룹 메시지를 무시하고 지정한 비공개 및 그룹 메시지만 크롤링합니다.

d. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정, 삭제된 동기화: 데이터 소스가 색인과 동기화될 때마다 신규, 수정, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

e. 동기화 실행 일정에서 빈도 - 데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

f. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

   b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

   c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Slack에 Amazon Kendra 연결하려면

TemplateConfiguration API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다. 다음 정보를 제공해야 합니다.

- 데이터 소스 - TemplateConfigurationJSON 스키마를 사용할 SLACK 때와 같이 데이터 소스 유형을 지정합니다. 또한 API를 TEMPLATE 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- Slack 워크스페이스 팀 ID - Slack 메인 페이지 URL에서 복사한 Slack 팀 ID입니다.
- 시작 날짜 - Slack 워크스페이스 팀에서 데이터를 크롤링하기 시작한 날짜입니다. 날짜는 다음 형식을 따라야 합니다. yyyy-mm-dd
- 동기화 모드 - 데이터 원본 콘텐츠가 변경될 때 색인을 업데이트하는 방법을 Amazon Kendra 지정합니다. 데이터 원본을 처음으로 동기화하는 경우 기본적으로 모든 콘텐츠가 크롤링되고 색인됩니다. Amazon Kendra 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다. 다음 중에서 선택할 수 있습니다.
  - FORCED_FULL_CRAWL모든 콘텐츠를 새로 인덱싱하려면 데이터 원본이 색인과 동기화될 때마다 기존 콘텐츠를 교체해야 합니다.
  - FULL_CRAWL데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화하도록 합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
  - CHANGE_LOG데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

- 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.
- 비밀 Amazon 리소스 이름 (ARN) - Slack 계정의 인증 자격 증명이 포함된 Secrets Manager 비밀의 Amazon 리소스 이름 (ARN) 을 제공합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "slackToken": "token"
}
```

- IAM 역할 —역할에 Secrets Manager 보안 액세스 권한을 제공하고 Slack IAM 커넥터 및 에 필요한 퍼블릭 API를 호출하기 위해 호출할 `RoleArn` 시기를 지정합니다. `CreateDataSource` Amazon Kendra자세한 내용은 Slack 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 특정 채널 —공개 또는 비공개 채널별로 필터링하고 ID를 기준으로 특정 채널을 지정합니다.
- 채널 및 메시지 유형 - 공개 및 비공개 채널, 그룹 및 다이렉트 메시지, 봇 및 보관된 메시지를 Amazon Kendra 인덱싱할지 여부. 봇 토큰을 Slack 보안 인증의 일부로 사용하는 경우 인덱싱하려는 채널에 봇 토큰을 추가해야 합니다. 봇 토큰을 사용하여 다이렉트 메시지와 그룹 메시지를 인덱싱할 수 없습니다.
- 되돌아보기 —Slack 커넥터가 마지막 커넥터 동기화 전까지 지정된 시간까지 업데이트되거나 삭제된 콘텐츠를 크롤링하도록 `lookBack` 매개 변수를 구성할 수 있습니다.
- 포함 및 제외 필터 - 특정 Slack 콘텐츠를 포함할지 제외할지를 지정합니다. 봇 토큰을 Slack 보안 인증의 일부로 사용하는 경우 인덱싱하려는 채널에 봇 토큰을 추가해야 합니다. 봇 토큰을 사용하여 다이렉트 메시지와 그룹 메시지를 인덱싱할 수 없습니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 필드 매핑 - Slack 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

> **Note**
>
> 문서를 Amazon Kendra 검색하려면 문서 본문 필드 또는 문서에 해당하는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Slack 템플릿 스키마를 참조하세요.

자세히 알아보기

Slack 데이터 Amazon Kendra 소스와의 통합에 대해 자세히 알아보려면 다음을 참조하십시오.

- Amazon Kendra Slack 커넥터를 사용한 지능형 검색으로 Slack 워크스페이스의 지식을 넓히세요

Zendesk

Zendesk는 기업이 고객 지원 상호 작용을 자동화하고 향상시키는 데 도움이 되는 고객 관계 관리 시스템입니다. Zendesk Support 티켓, 티켓 댓글, 티켓 첨부 파일, 헬프 센터 문서, 문서 댓글, 문서 댓글 첨부 파일, Guide 커뮤니티 주제, 커뮤니티 게시물, 커뮤니티 게시물 댓글을 인덱싱하는 데 사용할 Amazon Kendra 수 있습니다.

특정 조직 내에만 있는 티켓을 인덱싱하려면 조직 이름을 기준으로 필터링할 수 있습니다. Zendesk에서 데이터 크롤링을 시작하려는 날짜를 크롤링 날짜로 설정하도록 선택할 수도 있습니다.

Amazon Kendra 콘솔과 API를 사용하여 Zendesk 데이터 소스에 연결할 Amazon Kendra 수 있습니다. TemplateConfiguration

Amazon Kendra Zendesk 데이터 소스 커넥터의 문제를 해결하려면 을 참조하십시오. 데이터 소스 문제 해결

주제

- 지원되는 기능
- 필수 조건
- 연결 지침
- 자세히 알아보기

지원되는 기능

Amazon Kendra Zendesk 데이터 소스 커넥터는 다음 기능을 지원합니다.

- 필드 매핑
- 사용자 액세스 제어
- 포함/제외 필터
- 변경 로그, 전체 및 증분 콘텐츠 동기화
- Virtual Private Cloud(VPC)

필수 조건

Zendesk 데이터 소스를 Amazon Kendra 인덱싱하는 데 사용하려면 먼저 Zendesk와 계정에서 이러한 변경을 하세요. AWS

Zendesk에서 다음 사항을 갖추었는지 확인하세요.

- Zendesk Suite(Professional/Enterprise) 관리자 계정을 만들었습니다.
- Zendesk 호스트 URL을 기록해 두었습니다. *#: https://{sub-domain}.zendesk.com/.*

> **Note**
>
> (온프레미스/서버) AWS Secrets Manager 는 에 포함된 엔드포인트 정보가 데이터 소스 구성 세부 정보에 지정된 엔드포인트 정보와 동일한지 Amazon Kendra 확인합니다. 이렇게 하면 사용자가 작업을 수행할 권한이 없지만 구성된 보안 암호에 액세스하여 작업을 수행하는 데 Amazon Kendra 를 프록시로 사용하는 보안 문제인 혼동된 대리자 문제를 방지하는 데 도

움이 됩니다. 나중에 엔드포인트 정보를 변경하는 경우 새 보안 암호를 생성하여 이 정보를 동기화해야 합니다.

- 클라이언트 ID, 클라이언트 암호, 사용자 이름 및 암호를 포함하는 OAuth 2.0 토큰을 구성했습니다. 인증 자격 증명으로 사용하려면 OAuth 2.0 토큰이 필요합니다. 자세한 내용은 OAuth 2.0 토큰 구성에 관한 Zendesk 설명서를 참조하세요.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

- 다음 OAuth 2.0 범위가 추가되었습니다.
  - 읽기
- 선택 사항: Amazon Kendra 에 연결하도록 허용하는 SSL 인증서를 설치했습니다.
- Zendesk 및 동일한 인덱스에 사용할 다른 여러 데이터 소스에서 각 문서가 고유한지 확인했습니다. 인덱스에 사용하려는 각 데이터 소스에는 데이터 소스 전체에서 동일한 문서가 포함되어서는 안 됩니다. 문서 ID는 인덱스 전체에 적용되며 인덱스별로 고유해야 합니다.

에 다음이 AWS 계정있는지 확인하세요.

- Amazon Kendra 색인을 만들고 API를 사용하는 경우 색인 ID를 기록해 두었습니다.
- 데이터 소스에 대한 IAM 역할을 만들고 API를 사용하는 경우 역할의 IAM ARN을 기록했습니다.

> **Note**
>
> 인증 유형과 자격 증명을 변경하는 경우 올바른 AWS Secrets Manager 비밀 ID에 액세스하려면 IAM 역할을 업데이트해야 합니다.

- Zendesk 보안 인증 정보를 AWS Secrets Manager 보안 암호에 저장했고 API를 사용하는 경우 보안 암호의 ARN을 기록했습니다.

> **Note**
>
> 보안 인증 정보와 보안 암호를 정기적으로 새로 고치거나 교체하는 것이 좋습니다. 보안을 위해 필요한 액세스 수준만 제공하세요. 데이터 소스, 커넥터 버전 1.0 및 2.0(해당하는 경우) 간에 보안 인증 정보와 보안 암호를 재사용하지 않는 것이 좋습니다.

기존 IAM 역할이나 비밀번호가 없는 경우 Zendesk 데이터 소스를 연결할 Amazon Kendra때 콘솔을 사용하여 새 IAM 역할과 Secrets Manager 비밀번호를 만들 수 있습니다. API를 사용하는 경우 기존 IAM 역할 및 Secrets Manager 암호의 ARN과 인덱스 ID를 제공해야 합니다.

연결 지침

Zendesk 데이터 소스에 Amazon Kendra 연결하려면 데이터에 액세스할 Amazon Kendra 수 있도록 Zendesk 데이터 원본의 필수 세부 정보를 제공해야 합니다. Zendesk를 위해 Amazon Kendra아직 구성하지 않았다면 을 참조하세요. 필수 조건

Console

Zendesk에 Amazon Kendra 연결하려면

1. 에 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 인덱스를 선택한 다음 인덱스 목록에서 사용할 인덱스를 선택합니다.

   > **Note**
   >
   > 인덱스 설정에서 사용자 액세스 제어 설정을 구성하거나 편집할 수 있습니다.

3. 시작하기 페이지에서 데이터 소스 추가를 선택합니다.
4. 데이터 원본 추가 페이지에서 Zendesk 커넥터를 선택한 다음 커넥터 추가를 선택합니다. 버전 2 (해당하는 경우) 를 사용하는 경우 “V2.0" 태그가 있는 Zendesk 커넥터를 선택하세요.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 정보를 입력합니다.

   a. 이름 및 설명에서 데이터 소스 이름에 데이터 소스의 이름을 입력합니다. 하이픈은 포함할 수 있지만 공백은 포함할 수 없습니다.

   b. (선택 사항) 설명 - 데이터 소스에 대한 선택적 설명을 입력합니다.

c. 기본 언어 - 문서를 필터링하여 색인에 사용할 언어를 선택하세요. 달리 지정하지 않는 한, 언어는 영어로 기본 설정됩니다. 문서 메타데이터에 지정된 언어가 선택한 언어보다 우선합니다.

d. 태그의 새 태그 추가 - 리소스를 검색 및 필터링하거나 비용을 추적할 수 있는 선택적 태그를 포함합니다. AWS

e. 다음을 선택합니다.

6. 액세스 및 보안 정의 페이지에서 다음 정보를 입력합니다.

a. Zendesk URL - Zendesk URL을 입력합니다. 예: *https://{sub-domain}.zendesk.com/.*

b. 권한 부여 —ACL이 있고 액세스 제어에 사용하려는 경우 문서에 대한 액세스 제어 목록(ACL) 정보를 켜거나 끕니다. ACL은 사용자와 그룹이 액세스할 수 있는 문서를 지정합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

c. AWS Secrets Manager 비밀 - 기존 비밀번호를 선택하거나 새 비밀번호를 만들어 Zendesk Secrets Manager 인증 자격 증명을 저장합니다. 새 암호를 만들기로 선택하면 AWS Secrets Manager 보안 암호 창이 열립니다.

i. AWS Secrets Manager 보안 암호 생성 창에 다음 정보를 입력합니다.

A. 보안 암호 이름 - 보안 암호의 이름. 접두사 AmazonKendra '-Zendesk-'가 비밀 이름에 자동으로 추가됩니다.

B. 클라이언트 ID, 클라이언트 비밀번호, 사용자 이름, 비밀번호의 경우 —Zendesk에 구성된 인증 자격 증명 값을 입력합니다.

ii. 비밀번호를 저장하고 추가하세요.

d. Virtual Private Cloud(VPC) - VPC를 사용하기로 선택할 수 있습니다. 그렇다면 서브넷과 VPC 보안 그룹을 추가해야 합니다.

e. 아이덴티티 크롤러 - 아이덴티티 크롤러를 활성화할지 여부를 지정합니다. Amazon Kendra ID 크롤러는 문서에 대한 액세스 제어 목록 (ACL) 정보를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링합니다. 문서에 대한 ACL이 있고 ACL을 사용하기로 선택한 경우 ID 크롤러를 활성화하여 검색 결과에 대한 Amazon Kendra사용자 컨텍스트 필터링을 구성할 수도 있습니다. 그렇지 않으면 ID 크롤러를 끄면 모든 문서를 공개적으로 검색할 수 있습니다. 문서에 대한 액세스 제어를 사용하고 싶은데 ID 크롤러가 꺼져 있는 경우 PutPrincipalMappingAPI를 사용하여 사용자 컨텍스트 필터링을 위한 사용자 및 그룹 액세스 정보를 업로드할 수도 있습니다.

f. IAM 역할 - 기존 역할을 선택하거나 새 IAM IAM 역할을 생성하여 리포지토리 자격 증명에 액세스하고 콘텐츠를 인덱싱합니다.

> **Note**
>
> IAM 색인에 사용되는 역할은 데이터 원본에 사용할 수 없습니다. 기존 역할을 인덱스나 FAQ에 사용하는지 확실하지 않은 경우 새 역할 생성을 선택하여 오류를 방지하세요.

g. 다음을 선택합니다.

7. 동기화 설정 구성 페이지에 다음 정보를 입력합니다.

a. 콘텐츠 선택: 티켓, 헬프 센터 문서, 커뮤니티 주제 등에서 크롤링하려는 콘텐츠 유형을 선택합니다.

b. 조직 이름 - 콘텐츠를 필터링하려면 Zendesk 조직 이름을 입력합니다.

c. 동기화 시작 날짜 - 콘텐츠 크롤링을 시작하려는 날짜를 입력합니다.

d. 정규식 패턴 - 특정 페이지 및 자산을 포함하거나 제외하기 위한 정규 표현식 패턴을 추가합니다. 최대 100개의 패턴을 추가할 수 있습니다.

e. 동기화 모드 - 데이터 소스 콘텐츠가 변경될 때 인덱스를 업데이트하는 방법을 선택합니다. 데이터 원본을 처음으로 동기화하는 경우 Amazon Kendra 기본적으로 모든 콘텐츠가 크롤링되고 인덱싱됩니다. 동기화 모드 옵션으로 전체 동기화를 선택하지 않더라도 초기 동기화에 실패한 경우 데이터의 전체 동기화를 실행해야 합니다.

- 전체 동기화: 모든 콘텐츠를 새로 인덱싱하여 데이터 소스가 색인과 동기화될 때마다 기존 콘텐츠를 대체합니다.
- 신규, 수정된 동기화: 데이터 소스가 색인과 동기화될 때마다 새 콘텐츠와 수정된 콘텐츠만 인덱싱합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.
- 새 콘텐츠, 수정된 콘텐츠, 삭제된 동기화: 데이터 원본이 색인과 동기화될 때마다 새 콘텐츠, 수정된 콘텐츠, 삭제된 콘텐츠만 색인화합니다. Amazon Kendra 데이터 원본의 메커니즘을 사용하여 콘텐츠 변경 사항을 추적하고 마지막 동기화 이후 변경된 콘텐츠를 인덱싱할 수 있습니다.

f. 빈도에 대한 동기화 실행 일정 —데이터 원본 콘텐츠를 동기화하고 색인을 업데이트하는 빈도를 선택합니다.

g. 다음을 선택합니다.

8. 필드 매핑 설정 페이지에 다음 정보를 입력합니다.

   a. 기본 데이터 원본 필드 - Amazon Kendra 생성된 기본 데이터 원본 필드 중에서 색인에 매핑하려는 필드를 선택합니다.

   b. 필드 추가 - 사용자 지정 데이터 소스 필드를 추가하려면 매핑할 인덱스 필드 이름과 필드 데이터 유형을 만듭니다.

   c. 다음을 선택합니다.

9. 검토 및 생성 페이지에서 입력한 정보가 정확한지 확인한 다음 데이터 소스 추가를 선택합니다. 이 페이지에서 정보를 편집하도록 선택할 수도 있습니다. 데이터 소스가 성공적으로 추가된 후 데이터 소스 페이지에 데이터 소스가 표시됩니다.

API

Zendesk에 Amazon Kendra 연결하려면

API를 사용하여 데이터 소스 스키마의 JSON을 지정해야 합니다 TemplateConfiguration. 다음 정보를 제공해야 합니다.

- 데이터 원본 - TemplateConfigurationJSON 스키마를 사용할 `ZENDESK` 때와 같이 데이터 원본 유형을 지정합니다. 또한 API를 `TEMPLATE` 호출할 때와 같이 데이터 소스를 지정하십시오. CreateDataSource
- 호스트 URL - 연결 구성 또는 리포지토리 엔드포인트 세부 정보의 일부로 Zendesk 호스트 URL을 제공합니다. 예: *`https://yoursubdomain.zendesk.com`*.
- 변경 로그 - Zendesk 데이터 소스 변경 로그 메커니즘을 사용하여 색인에서 문서를 업데이트해야 하는지 여부를 Amazon Kendra 결정할지 여부입니다.

  > **Note**
  >
  > Amazon Kendra 가 모든 문서를 스캔하지 않도록 하려면 변경 로그를 사용하세요. 변경 로그가 크면 변경 로그를 처리하는 것보다 Zendesk 데이터 원본의 문서를 스캔하는 데 걸리는 시간이 더 Amazon Kendra 적을 수 있습니다. Zendesk 데이터 소스를 인덱스와 처음으로 동기화하는 경우 모든 문서가 스캔됩니다.

- 비밀 Amazon 리소스 이름 (ARN): Zendesk 계정의 인증 자격 증명이 포함된 Secrets Manager 비밀의 Amazon 리소스 이름 (ARN) 을 입력합니다. 보안 암호는 다음 키가 있는 JSON 구조로 저장됩니다.

```
{
    "hostUrl": "https://yoursubdomain.zendesk.com",
    "clientId": "client ID",
    "clientSecret": "Zendesk client secret",
    "userName": "Zendesk user name",
    "password": "Zendesk password"
}
```

- IAM 역할 `CreateDataSource` Secrets Manager —비밀에 액세스할 권한을 가진 IAM 역할을 제공하고 Zendesk 커넥터 및 에 필요한 공개 API를 호출하도록 호출할 `RoleArn` 시기를 지정합니다. Amazon Kendra자세한 내용은 Zendesk 데이터 소스에 대한IAM 역할을 참조하세요.

다음 선택적 기능도 추가할 수 있습니다.

- Virtual Private Cloud(VPC) - `CreateDataSource`를 호출할 때 `VpcConfiguration`을 지정합니다. 자세한 정보는 를 사용하도록 Amazon Kendra 구성하기 Amazon VPC을 참조하세요.
- 문서/콘텐츠 유형 —크롤링 여부를 지정하세요.
  - 지원 티켓, 티켓 설명 및/또는 티켓 설명 첨부 파일
  - 도움말 센터 기사, 기사 첨부 파일, 기사 설명
  - 가이드 커뮤니티 주제, 게시물 또는 게시물 댓글
- 포함 및 제외 필터 - 특정 Slack 콘텐츠를 포함할지 제외할지를 지정합니다. 봇 토큰을 Slack 보안 인증의 일부로 사용하는 경우 인덱싱하려는 채널에 봇 토큰을 추가해야 합니다. 봇 토큰을 사용하여 다이렉트 메시지와 그룹 메시지를 인덱싱할 수 없습니다.

> **Note**
>
> 대부분의 데이터 소스는 필터라고 하는 포함 또는 제외 패턴인 정규 표현식 패턴을 사용합니다. 포함 필터를 지정하는 경우 포함 필터와 일치하는 콘텐츠만 인덱싱됩니다. 포함 필터와 일치하지 않는 문서는 인덱싱되지 않습니다. 포함 및 제외 필터를 지정하는 경우 제외 필터와 일치하는 문서는 포함 필터와 일치하더라도 인덱싱되지 않습니다.

- 사용자 컨텍스트 필터링 및 액세스 제어 — 문서에 대한Amazon Kendra ACL이 있는 경우 문서의 액세스 제어 목록 (ACL) 을 크롤링합니다. ACL 정보는 검색 결과를 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링하는 데 사용됩니다. 자세한 내용은 사용자 컨텍스트 필터링을 참조하세요.

- 필드 매핑 - Zendesk 데이터 소스 필드를 Amazon Kendra 인덱스 필드에 매핑하기로 선택합니다. 자세한 내용을 알아보려면 데이터 소스 필드 매핑을 참조하세요.

  > **Note**
  >
  > 문서를 Amazon Kendra 검색하려면 문서에 해당하는 문서 본문 필드 또는 문서 본문 필드가 필요합니다. 데이터 원본의 문서 본문 필드 이름을 색인 필드 이름에 매핑해야 `_document_body` 합니다. 다른 모든 필드는 선택 사항입니다.

구성해야 할 기타 중요한 JSON 키 목록은 Zendesk 템플릿 스키마를 참조하세요.

자세히 알아보기

Zendesk 데이터 Amazon Kendra 소스와 통합하는 방법에 대해 자세히 알아보려면 다음을 참조하세요.

- 지능형 검색으로 Zendesk의 인사이트를 찾아보세요 Amazon Kendra

데이터 소스 필드 매핑

Amazon Kendra 데이터 원본 커넥터는 데이터 원본의 문서 또는 콘텐츠 필드를 색인의 필드에 매핑할 수 있습니다. Amazon Kendra 기본적으로 각 커넥터는 특정 데이터 소스 필드를 크롤링하도록 설계되었습니다. 기본 데이터 소스 필드와 해당 속성은 변경하거나 사용자 지정할 수 없습니다. Amazon Kendra 콘솔에서 편집할 수 없는 기본 필드 및 기본 필드 속성은 회색으로 표시됩니다.

Amazon Kendra 또한 커넥터를 사용하면 데이터 원본의 사용자 지정 문서 또는 콘텐츠 필드를 색인의 사용자 지정 필드에 매핑할 수 있습니다. 예를 들어, 데이터 소스에 문서의 부서 정보가 포함된 “dept”라는 필드가 있는 경우, “Department”라는 인덱스 필드에 매핑할 수 있습니다. 이렇게 하면 문서를 쿼리할 때 필드를 사용할 수 있습니다.

다음과 같은 Amazon Kendra 예약된 필드나 공통 필드를 매핑할 수도 `_created_at` 있습니다. 데이터 원본에 “creation_date”라는 필드가 있는 경우 이 필드를 이라는 동일한 Amazon Kendra 예약 필드에 매핑할 수 있습니다. `_created_at` Amazon Kendra 예약된 필드에 대한 자세한 내용은 문서 속성 또는 필드를 참조하십시오.

대부분의 데이터 소스에 대해 필드를 매핑할 수 있습니다. 다음 데이터 소스에 대한 필드 매핑을 만들 수 있습니다.

- Adobe Experience Manager
- Alfresco
- Aurora (MySQL)
- Aurora (PostgreSQL)
- Amazon FSx (윈도우)
- Amazon FSx (NetApp ONTAP)
- Amazon RDS/Aurora
- Amazon RDS (Microsoft SQL Server)
- Amazon RDS (MySQL)
- Amazon RDS (Oracle)
- Amazon RDS (PostgreSQL)
- Amazon Kendra 웹 크롤러
- Amazon WorkDocs
- Box
- Confluence
- Dropbox
- Drupal
- GitHub
- Google Workspace Drives
- Gmail
- IBM Db2
- Jira
- Microsoft Exchange
- 마이크로소프트 OneDrive
- 마이크로소프트 SharePoint
- Microsoft Teams
- Microsoft SQL Server
- Microsoft Yammer
- MySQL

- Oracle Database
- PostgreSQL
- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

문서를 S3 버킷 또는 S3 데이터 소스에 저장하는 경우 JSON 메타데이터 파일을 사용하여 필드를 지정합니다. 자세한 내용은 S3 데이터 소스 커넥터를 참고하세요.

데이터 소스 필드를 인덱스 필드에 매핑하는 과정은 다음과 같이 3단계입니다.

1. 인덱스를 생성합니다. 자세한 내용은 인덱스 생성을 참조하세요.
2. 인덱스를 업데이트하여 필드를 추가합니다.
3. 데이터 원본을 만들고 필드 매핑을 포함하여 예약된 필드와 사용자 지정 필드를 Amazon Kendra 인덱스 필드에 매핑하세요.

색인을 업데이트하여 사용자 지정 필드를 추가하려면 콘솔을 사용하여 데이터 원본 필드 매핑을 편집하고 사용자 지정 필드를 추가하거나 API를 사용하십시오. UpdateIndex 인덱스에 총 500개의 사용자 지정 필드를 추가할 수 있습니다.

데이터베이스 데이터 소스의 경우 데이터베이스 열의 이름이 예약된 필드의 이름과 일치하면 필드와 열이 자동으로 매핑됩니다.

UpdateIndexAPI를 사용하면 를 사용하여 예약 및 사용자 지정 필드를 추가할 수 있습니다. `DocumentMetadataConfigurationUpdates`

다음 JSON 예제는 `DocumentMetadataConfigurationUpdates`를 사용하여 “Department”라는 필드를 인덱스에 추가합니다.

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

```
]
```

필드를 생성할 때 해당 필드가 검색에 사용되는 방식을 설정할 수 있습니다. 사용자는 다음 중에서 선택할 수 있습니다.

- Displayable - 쿼리 응답에서 필드를 반환할지 여부를 결정합니다. 기본값은 `true`입니다.
- Facetable - 필드를 사용하여 패싯을 생성할 수 있음을 나타냅니다. 기본값은 `false`입니다.
- Searchable - 필드가 검색에 사용되는지 여부를 결정합니다. 기본값은 문자열 필드의 경우 `true`이고 숫자 및 날짜 필드의 경우 `false`입니다.
- Sortable - 필드를 사용하여 쿼리의 응답을 정렬할 수 있음을 나타냅니다. 날짜, 숫자 및 문자열 필드에 대해서만 설정할 수 있습니다. 문자열 목록 필드에는 설정할 수 없습니다.

다음 JSON 예제는 `DocumentMetadataConfigurationUpdates`를 사용하여 “Department”라는 필드를 인덱스에 추가하고 facetable로 표시합니다.

```
"DocumentMetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE",
       "Search": {
           "Facetable": true
       }
   }
]
```

Amazon Kendra 예약된 문서 필드 또는 공통 문서 필드 사용

UpdateIndex API를 사용하면 예약된 인덱스 필드 이름을 `DocumentMetadataConfigurationUpdates` 사용하고 지정하여 해당 문서 속성/필드 이름에 매핑되는 Amazon Kendra 예약 또는 공통 필드를 만들 수 있습니다. 사용자 지정 필드도 생성할 수 있습니다. 데이터 소스 커넥터를 사용하는 경우 대부분의 커넥터에는 데이터 소스 문서 필드를 인덱스 필드에 매핑하는 필드 매핑이 포함됩니다. Amazon Kendra 콘솔을 사용하는 경우 데이터 소스를 선택하고 편집 작업을 선택한 다음 데이터 소스 구성을 위한 필드 매핑 섹션 옆으로 이동하여 필드를 업데이트합니다.

필드를 displayable, facetable, searchable, sortable로 설정하도록 `Search` 객체를 구성할 수 있습니다. 필드의 순위 순서, 부스트 기간 또는 부스팅에 적용할 기간, 최신성, 중요도 값 및 특정 필드 값에 매핑된 중요도 값을 설정하도록 `Relevance` 객체를 구성할 수 있습니다. 콘솔을 사용하는 경우 탐색 메뉴

에서 패싯 옵션을 선택하여 필드에 대한 검색 설정을 지정할 수 있습니다. 관련성 조정을 설정하려면 탐색 메뉴에서 인덱스를 검색하는 옵션을 선택하고 쿼리를 입력한 다음 사이드 패널 옵션을 사용하여 검색 관련성을 조정합니다. 필드를 생성한 후에는 필드 유형을 변경할 수 없습니다.

Amazon Kendra 사용할 수 있는 다음과 같은 예약된 문서 필드 또는 공통 문서 필드가 있습니다.

- `_authors` - 문서 내용을 책임지는 한 명 이상의 작성자 목록.
- `_category` - 문서를 특정 그룹에 배치하는 범주.
- `_created_at` - 문서가 생성된 ISO 8601 형식의 날짜 및 시간. 예를 들어, 2012-03-25T12:30:10+01:00은 중부 유럽 시간으로 2012년 3월 25일 오후 12시 30분 10초에 대한 ISO 8601 날짜/시간 형식입니다.
- `_data_source_id` - 문서가 포함된 데이터 소스의 식별자.
- `_document_body` - 작업 문서의 내용.
- `_document_id` - 문서의 고유 식별자.
- `_document_title` - 문서의 제목.
- `_excerpt_page_number` - 문서 발췌문이 나타나는 PDF 파일의 페이지 번호. 2020년 9월 8일 이전에 인덱스를 만든 경우 이 속성을 사용하려면 먼저 문서를 다시 인덱싱해야 합니다.
- `_faq_id` - 질문-답변 유형 문서(FAQ)인 경우 FAQ의 고유 식별자.
- `_file_type` - 문서의 파일 형식(예: pdf 또는 doc).
- `_last_updated_at` - 문서가 마지막으로 업데이트된 ISO 8601 형식의 날짜 및 시간. 예를 들어, 2012-03-25T12:30:10+01:00은 중부 유럽 시간으로 2012년 3월 25일 오후 12시 30분 10초에 대한 ISO 8601 날짜/시간 형식입니다.
- `_source_uri` - 문서가 제공된 URI. 예를 들면, 회사 웹 사이트에 있는 문서의 URI.
- `_version` - 문서의 특정 버전을 나타내는 식별자.
- `_view_count` - 문서가 조회된 횟수.
- `_language_code`(문자열) - 문서에 적용되는 언어의 코드. 언어를 지정하지 않으면 영어가 기본값으로 사용됩니다. 코드를 포함하여 지원되는 언어에 대한 자세한 내용은 영어 이외의 언어로 문서 추가를 참조하세요.

사용자 지정 필드의 경우 예약된 필드 또는 공통 필드를 만들 때와 마찬가지로 `UpdateIndex` API와 `DocumentMetadataConfigurationUpdates`를 사용하여 이러한 필드를 만듭니다. 사용자 지정 필드에 적절한 데이터 유형을 설정해야 합니다. 콘솔을 사용하는 경우 데이터 소스를 선택하고 편집 작업을 선택한 다음 데이터 소스 구성을 위한 필드 매핑 섹션 옆으로 이동하여 필드를 업데이트합니다. 일

부 데이터 소스는 새 필드 또는 사용자 지정 필드 추가를 지원하지 않습니다. 필드를 생성한 후에는 필드 유형을 변경할 수 없습니다.

사용자 지정 필드에 설정할 수 있는 유형은 다음과 같습니다.

- 날짜
- 숫자
- String
- 문자열 목록

BatchPutDocumentAPI를 사용하여 색인에 문서를 추가한 경우 문서의 필드/속성을 `Attributes` 나열하고 객체를 사용하여 필드를 생성합니다. `DocumentAttribute`

Amazon S3 데이터 소스에서 색인된 문서의 경우 필드 정보가 포함된 JSON 메타데이터 파일을 사용하여 필드를 만듭니다.

지원되는 데이터베이스를 데이터 소스로 사용하는 경우 필드 매핑 옵션을 사용하여 필드를 구성할 수 있습니다.

영어 이외의 언어로 문서 추가

문서를 여러 언어로 인덱싱할 수 있습니다. 언어를 지정하지 않으면 Amazon Kendra 는 기본적으로 문서를 영어로 인덱싱합니다. 문서 메타데이터에 있는 문서의 언어 코드를 필드로 포함합니다. 문서의 `_language_code` 필드에 대한 자세한 내용은 필드 매핑 및 사용자 지정 속성을 참조하세요.

호출할 때 데이터 소스에 있는 모든 문서의 언어 코드를 지정할 수 있습니다. CreateDataSource 메타데이터 필드에 지정된 언어 코드가 없는 문서의 경우, 데이터 소스 수준에서 모든 문서에 지정된 언어 코드를 사용하여 문서가 인덱싱됩니다. 콘솔에서는 데이터 소스 수준에서만 지원되는 언어로 문서를 인덱싱할 수 있습니다. 데이터 소스로 이동한 다음 데이터 소스 세부 정보 지정 페이지로 이동하여 언어 드롭다운에서 언어를 선택합니다.

지원되는 언어로 문서를 검색하거나 쿼리할 수도 있습니다. 자세한 내용은 언어 선택을 참조하세요.

다음 언어와 해당 코드가 지원됩니다(언어를 지정하지 않은 경우 영어(en)가 기본적으로 지원됨). 이 표에는 전체 시맨틱 검색을 Amazon Kendra 지원하는 언어와 단순 키워드 일치만 지원하는 언어가 포함되어 있습니다. 전체 시맨틱 검색을 지원하는 언어는 별표로 표시되며 다음 표에서는 굵은 텍스트로 표시됩니다. 영어(기본 언어)도 전체 시맨틱 검색이 지원됩니다.

언어 이름	언어 코드
아랍어	ar
아르메니아어	hy
바스크어	eu
벵골어	bn
불가리아어	bg
카탈루냐어	ca
중국어 — 간체 및 번체*	zh
체코어	cs
덴마크어	da
네덜란드어	nl
핀란드어	fi
프랑스어 — 프랑스어(캐나다) 포함*	fr
갈리시아어	gl
독일어*	de
그리스어	el
힌디어	hi
헝가리어	hu
인도네시아어	id
아일랜드어	ga
이탈리아어	it

언어 이름	언어 코드
일본어*	ja
한국어*	ko
라트비아어	lv
리투아니아어	lt
노르웨이어	no
페르시아어	fa
포르투갈어	pt
포르투갈어(브라질)*	pt-BR
루마니아어	ro
러시아어	ru
소라니	ckb
스페인어 — 스페인어(멕시코) 포함*	es
스웨덴어	sv
터키어	tr

*해당 언어에 대한 시맨틱 검색이 지원됩니다.

시맨틱 검색을 지원하는 언어의 경우 다음 기능이 지원됩니다.

- 단순한 키워드 매칭을 넘어선 문서 관련성.
- 단순한 키워드 매칭 이외의 FAQ.
- Amazon Kendra의 독해력을 기반으로 문서에서 답을 추출합니다.
- 검색 결과의 신뢰도 버킷(매우 높음, 높음, 중간, 낮음).

시맨틱 검색을 지원하지 않는 언어의 경우 문서 관련성 및 FAQ에 대한 단순 키워드 매칭이 지원됩니다.

동의어(사용자 지정 동의어 포함), 점진적 학습 및 피드백, 쿼리 제안은 영어(기본 언어)에만 지원됩니다.

를 사용하도록 Amazon Kendra 구성하기 Amazon VPC

Amazon Kendra 를 사용하여 생성한 가상 사설 클라우드 (VPC) 에 연결하여 사설 클라우드에서 실행되는 데이터 소스에 저장된 콘텐츠를 인덱싱할 수 있습니다. Amazon Virtual Private Cloud 데이터 소스 커넥터를 만들 때 데이터 소스가 포함된 서브넷에 보안 그룹 및 서브넷 식별자를 제공할 수 있습니다. 이 정보를 바탕으로 VPC 내의 데이터 소스와 안전하게 통신하는 데 사용하는 Elastic Network 인터페이스를 Amazon Kendra 생성합니다.

Amazon Kendra 데이터 소스 커넥터를 Amazon VPC 설정하려면 AWS Management Console 또는 CreateDataSourceAPI 작업을 사용할 수 있습니다. 콘솔을 사용하는 경우 커넥터 구성 프로세스 중에 VPC를 연결합니다.

> **Note**
>
> 이 Amazon VPC 기능은 Amazon Kendra 데이터 소스 커넥터를 설정할 때 선택 사항입니다. 공용 인터넷에서 데이터 원본에 접근할 수 있는 경우 Amazon VPC 기능을 활성화하지 않아도 됩니다. 모든 Amazon Kendra 데이터 소스 커넥터가 지원하는 것은 아닙니다 Amazon VPC.

데이터 원본이 퍼블릭 인터넷에서 실행되지 않고 Amazon VPC 있고 퍼블릭 인터넷에서 액세스할 수 없는 경우, 먼저 가상 사설망 (VPN) 을 사용하여 데이터 소스를 VPC에 연결합니다. 그런 다음 Amazon VPC 및 AWS Virtual Private Network 를 Amazon Kendra 조합하여 데이터 소스를 연결할 수 있습니다. VPN 설정에 대한 자세한 내용은 AWS VPN 설명서를 참조하십시오.

주제

- Amazon Kendra 커넥터 Amazon VPC 지원 구성
- 연결할 Amazon Kendra 데이터 소스 설정 Amazon VPC
- VPC의 데이터베이스에 연결
- VPC 연결 문제 해결

Amazon Kendra 커넥터 Amazon VPC 지원 구성

Amazon Kendra 커넥터와 함께 Amazon VPC 사용하도록 구성하려면 다음 단계를 수행하십시오.

단계

- 단계 1. 에 대한 Amazon VPC 서브넷을 생성합니다. Amazon Kendra
- 단계 2. 에 대한 Amazon VPC 보안 그룹을 생성합니다. Amazon Kendra
- 단계 3. 외부 데이터 소스를 구성하고 Amazon VPC

단계 1. 에 대한 Amazon VPC 서브넷을 생성합니다. Amazon Kendra

데이터 원본에 액세스하는 데 사용할 Amazon Kendra 수 있는 기존 Amazon VPC 서브넷을 만들거나 선택합니다. 준비된 서브넷은 다음 AWS 리전 및 가용 영역 중 하나에 있어야 합니다.

- 미국 서부(오레곤)/us-west-2—usw2-az1, usw2-az1, usw2-az2, usw2-az3
- 미국 동부(버니지아)/us-east-1—use1-az1, use1-az2, use1-az4
- 미국 동부(오하이오)/us-east-2—use2-az1, use2-az2, use2-az2, use2-az3
- 아시아 태평양(도쿄)/ap-northeast-1—apne1-az1, apne1-az2, apne1-az4
- 아시아 태평양(뭄바이)/ap-south-1—aps1-az1, aps1-az2, aps1-az3
- 아시아 태평양(싱가포르)/ap-southeast1-az1, apse1-az1, apse1-az2, apse1-az3
- 아시아 태평양(시드니)/ap-southeast-2—apse2-az1, apse2-az2, apse2-az3
- 캐나다(중부)/ca-central-1—cac1-az1, cac1-az2, cac1-az4
- 유럽(아일랜드)/eu-west-1—euw1-az1, uew1-az2, euw1-az3
- 유럽(런던)/eu-west-2—usw2-az1, usw2-az2, usw2-az2, usw2-az3

커넥터에 제공한 서브넷에서 데이터 원본에 액세스할 수 있어야 합니다. Amazon Kendra

Amazon VPC 서브넷 구성 방법에 대한 자세한 내용은 Amazon VPC 사용 설명서의 사용자 Amazon VPC 서브넷을 참조하십시오.

둘 이상의 서브넷 간에 연결을 Amazon Kendra 라우팅해야 하는 경우 여러 서브넷을 준비할 수 있습니다. 예를 들어, 데이터 원본이 포함된 서브넷의 IP 주소가 부족합니다. 이 경우 충분한 IP 주소를 갖고 첫 번째 서브넷에 연결된 추가 서브넷을 제공할 Amazon Kendra 수 있습니다. 여러 서브넷을 나열하는 경우 서브넷이 서로 통신할 수 있어야 합니다.

단계 2. 에 대한 Amazon VPC 보안 그룹을 생성합니다. Amazon Kendra

Amazon Kendra 데이터 소스 커넥터를 연결하려면 VPC에서 할당할 보안 그룹을 하나 이상 준비해야 합니다. Amazon VPC Amazon Kendra 보안 그룹은 에서 만든 Elastic Network 인터페이스에 연결됩니다 Amazon Kendra. 이 네트워크 인터페이스는 서브넷에 액세스할 Amazon Kendra 때 들어오고 나가는 인바운드 및 아웃바운드 트래픽을 제어합니다. Amazon VPC

보안 그룹의 아웃바운드 규칙이 Amazon Kendra 데이터 소스 커넥터의 트래픽이 동기화할 서브넷 및 데이터 원본에 액세스하는 것을 허용하는지 확인하십시오. 예를 들어 MySQL 커넥터를 사용하여 데이터베이스에서 동기화할 수 있습니다MySQL. 기본 포트를 사용하는 경우 보안 그룹은 데이터베이스를 실행하는 호스트의 포트 3306에 대한 Amazon Kendra 액세스를 허용해야 합니다.

사용할 기본 보안 그룹을 다음 값으로 구성하는 것이 좋습니다. Amazon Kendra

- 인바운드 규칙 - 이 필드를 비워 두면 모든 인바운드 트래픽이 차단됩니다.
- 아웃바운드 규칙 - 모든 아웃바운드 트래픽을 허용하는 규칙을 하나 추가하여 데이터 소스에서 동기화 요청을 시작할 Amazon Kendra 수 있도록 합니다.
  - IP 버전 — IPv4
  - 유형 — 모든 트래픽
  - 프로토콜 — 모든 트래픽
  - 포트 범위 — 전체
  - 목적지 — 0.0.0.0/0

Amazon VPC 보안 그룹을 구성하는 방법에 대한 자세한 내용은 Amazon VPC 사용 설명서의 보안 그룹 규칙을 참조하십시오.

단계 3. 외부 데이터 소스를 구성하고 Amazon VPC

외부 데이터 원본에 액세스할 수 있는 올바른 권한 구성 및 네트워크 설정이 있는지 확인하세요. Amazon Kendra 각 커넥터 페이지의 사전 요구 사항 섹션에서 데이터 원본을 구성하는 방법에 대한 자세한 지침을 찾을 수 있습니다.

또한 Amazon VPC 설정을 확인하고 할당할 서브넷에서 외부 데이터 원본에 연결할 수 있는지 확인하십시오. Amazon Kendra 이렇게 하려면 동일한 서브넷에 동일한 보안 그룹을 포함하는 Amazon EC2 인스턴스를 만들고 이 인스턴스에서 데이터 원본에 대한 액세스를 테스트하는 것이 좋습니다. Amazon EC2 자세한 내용은 Amazon VPC 연결 문제 해결을 참조하십시오.

연결할 Amazon Kendra 데이터 소스 설정 Amazon VPC

에서 Amazon Kendra 새 데이터 소스를 추가할 때 선택한 데이터 소스 커넥터가 이 Amazon VPC 기능을 지원하는 경우 해당 기능을 사용할 수 있습니다.

AWS Management Console 또는 Amazon Kendra API를 사용하여 Amazon VPC 활성화된 상태로 새 Amazon Kendra 데이터 원본을 설정할 수 있습니다. 구체적으로, CreateDataSourceAPI 작업을 사용한 다음 `VpcConfiguration` 파라미터를 사용하여 다음 정보를 제공하십시오.

- `SubnetIds`— 서브넷 식별자 Amazon VPC 목록
- `SecurityGroupIds`— 보안 그룹 식별자 목록 Amazon VPC

콘솔을 사용하는 경우 커넥터 구성 중에 필요한 Amazon VPC 정보를 제공합니다. 콘솔을 사용하여 커넥터의 Amazon VPC 기능을 활성화하려면 먼저 Amazon VPC를 선택합니다. 그런 다음 Amazon VPC 서브넷의 식별자와 Amazon VPC 보안 그룹의 식별자를 제공합니다. Amazon VPC 구성에서 생성한 Amazon VPC 서브넷 및 Amazon VPC 보안 그룹을 선택하거나 기존 서브넷을 사용할 수 있습니다.

주제

- Amazon VPC 식별자 보기
- 데이터 소스 IAM 역할 확인

Amazon VPC 식별자 보기

서브넷 및 보안 그룹의 식별자는 콘솔에서 구성됩니다. Amazon VPC 식별자를 보려면 다음 절차를 사용하십시오.

서브넷 식별자를 보려면

1. AWS Management Console 로그인하고 https://console.aws.amazon.com/vpc/ 에서 Amazon VPC 콘솔을 엽니다.
2. 탐색 창에서 서브넷을 선택합니다.
3. 서브넷 목록에서 데이터베이스 서버가 포함된 서브넷을 선택합니다.
4. 세부 정보 탭에서 서브넷 ID 필드의 식별자를 기록해 둡니다.

보안 그룹 식별자를 보려면

1. AWS Management Console 로그인하고 https://console.aws.amazon.com/vpc/ 에서 Amazon VPC 콘솔을 엽니다.
2. 탐색 창에서 보안 그룹을 선택합니다.
3. 보안 그룹 목록에서 식별자를 사용하려는 그룹을 선택합니다.
4. 세부 정보 탭에서 보안 그룹 ID 필드에 있는 식별자를 기록해 둡니다.

데이터 소스 IAM 역할 확인

데이터 소스 AWS Identity and Access Management IAM (커넥터) 역할에 액세스 권한이 포함되어 있는지 확인하세요 Amazon VPC.

콘솔을 사용하여 역할에 대한 새 역할을 생성하는 경우 는 사용자를 대신하여 IAM 역할에 올바른 권한을 Amazon Kendra 자동으로 추가합니다. IAM API를 사용하거나 기존 IAM 역할을 사용하는 경우 역할에 액세스 권한이 포함되어 있는지 확인하세요 Amazon VPC. 권한이 올바른지 확인하려면 VPC의 IAM 역할을 참조하십시오.

다른 Amazon VPC 서브넷을 사용하도록 기존 데이터 소스를 수정할 수 있습니다. 하지만 데이터 원본 커넥터가 제대로 작동하려면 데이터 원본의 IAM 역할을 확인하고 필요한 경우 변경 내용을 반영하도록 수정하십시오. Amazon Kendra

VPC의 데이터베이스에 연결

다음 예제는 가상 사설 클라우드 (VPC) 에서 실행되는 MySQL 데이터베이스를 연결하는 방법을 보여줍니다. 이 예제에서는 기본 VPC로 시작하고 데이터베이스를 만들어야 MySQL 한다고 가정합니다. VPC가 이미 있는 경우 다음과 같이 구성되어 있는지 확인하세요. 데이터베이스가 있는 경우 새 MySQL 데이터베이스를 만드는 대신 데이터베이스를 사용할 수 있습니다.

단계

- 1단계: VPC 구성
- 2단계: 보안 그룹 생성 및 구성
- 3단계: 데이터베이스 생성
- 4단계: 데이터 소스 커넥터 만들기

1단계: VPC 구성

서브넷에서 실행되는 MySQL 데이터베이스에 액세스할 수 있는 프라이빗 서브넷과 보안 그룹을 Amazon Kendra 갖도록 VPC를 구성합니다. VPC 구성에서 제공되는 서브넷은 미국 서부 (오레곤) 지역, 미국 동부 (버지니아 북부) 지역 또는 유럽 (아일랜드) 지역에 있어야 합니다.

를 사용하여 VPC를 구성하려면 Amazon VPC

1. AWS Management Console 로그인하고 https://console.aws.amazon.com/vpc/ 에서 Amazon VPC 콘솔을 엽니다.
2. 탐색 창에서 라우팅 테이블을 선택한 다음 라우팅 테이블 생성을 선택합니다.
3. 이름 필드에 를 입력합니다**`Private subnet route table`**. VPC 드롭다운에서 VPC를 선택한 다음 라우팅 테이블 생성을 선택합니다. 닫기를 선택하여 라우팅 테이블 목록으로 돌아갑니다.
4. 탐색 창에서 NAT 게이트웨이를 선택한 다음 NAT 게이트웨이 생성을 선택합니다.
5. 서브넷 드롭다운에서 퍼블릭 서브넷인 서브넷을 선택합니다. 서브넷 ID를 기록해 둡니다.
6. 탄력적 IP 주소가 없는 경우 새 EIP 생성을 선택하고 NAT 게이트웨이 생성을 선택한 다음 닫기를 선택합니다.
7. 탐색 창에서 라우팅 테이블을 선택합니다.
8. 라우팅 테이블 목록에서 3단계에서 만든 프라이빗 서브넷 라우팅 테이블을 선택합니다. 작업에서 경로 편집을 선택합니다.
9. 라우팅 추가를 선택합니다. 인터넷으로 나가는 모든 트래픽을 **`0.0.0.0/0`** 허용하려면 목적지로 를 입력합니다. 대상에서 NAT 게이트웨이를 선택한 다음, 4단계에 생성한 게이트웨이를 선택합니다. 변경 내용 저장을 선택한 다음 [닫기] 를 선택합니다.
10. 작업 메뉴에서 서브넷 연결 편집을 선택합니다.
11. 비공개로 설정하려는 서브넷을 선택합니다. 앞서 언급한 NAT 게이트웨이가 있는 서브넷은 선택하지 마세요. 작업을 완료하면 연결 저장을 선택합니다.

2단계: 보안 그룹 생성 및 구성

다음으로, 데이터베이스를 위한 보안 그룹을 구성합니다.

보안 그룹 생성 및 구성하기

1. AWS Management Console 로그인하고 https://console.aws.amazon.com/vpc/ 에서 Amazon VPC 콘솔을 엽니다.
2. VPC에 대한 설명에서 IPv4 CIDR을 기록해 두세요.

3. 탐색 창에서 보안 그룹을 선택한 다음 보안 그룹 생성을 선택합니다.
4. 보안 그룹 이름에 **DataSourceInboundSecurityGroup**를 입력합니다. 설명을 입력한 다음 목록에서 VPC를 선택합니다. 보안 그룹 생성을 선택한 다음 [닫기] 를 선택합니다.
5. 인바운드 규칙 탭을 선택합니다.
6. 인바운드 규칙 편집을 선택한 다음 규칙 추가를 선택합니다.
7. 데이터베이스의 경우 포트 범위의 포트 번호를 입력합니다. 예를 들어, HTTPS의 **3306** 경우 이고 HTTPS의 **443** 경우 입니다. MySQL 소스에 VPC의Classless Inter-Domain Routing(CIDR)을 입력합니다. 규칙 저장을 선택한 다음 닫기를 선택합니다.

보안 그룹은 VPC 내의 모든 사용자가 데이터베이스에 연결할 수 있도록 허용하고 인터넷으로의 아웃바운드 연결을 허용합니다.

3단계: 데이터베이스 생성

문서를 보관할 데이터베이스를 만들거나 기존 데이터베이스를 사용할 수 있습니다.

MySQL데이터베이스 생성 방법에 대한 지침은 을 참조하십시오 MySQL.

4단계: 데이터 소스 커넥터 만들기

VPC를 구성하고 데이터베이스를 만든 후 데이터베이스의 데이터 소스 커넥터를 만들 수 있습니다. Amazon Kendra 지원하는 데이터베이스 커넥터에 대한 자세한 내용은 지원되는 커넥터를 참조하십시오.

데이터베이스의 경우 VPC, VPC에서 만든 프라이빗 서브넷, VPC에 만든 보안 그룹을 구성해야 합니다.

VPC 연결 문제 해결

가상 사설 클라우드 (VPC) 연결에 문제가 발생하는 경우 IAM 권한, 보안 그룹 설정 및 서브넷의 라우팅 테이블이 올바르게 구성되어 있는지 확인하십시오.

데이터 소스 커넥터 동기화 실패의 잠재적 원인 중 하나는 할당된 서브넷에서 데이터 소스에 도달하지 못할 수 있기 때문입니다. Amazon Kendra 이 문제를 해결하려면 동일한 설정으로 Amazon EC2 인스턴스를 생성하는 것이 좋습니다. Amazon VPC 그런 다음 REST API 호출 또는 다른 방법 (데이터 원본의 특정 유형에 따라 다름) 을 사용하여 이 Amazon EC2 인스턴스에서 데이터 원본에 액세스해 보십시오.

생성한 Amazon EC2 인스턴스에서 데이터 원본에 성공적으로 액세스하면 이 서브넷에서 데이터 원본에 연결할 수 있다는 의미입니다. 따라서 동기화 문제는 사용자가 데이터 원본에 액세스할 수 없는 것과 관련이 없습니다. Amazon VPC

VPC 구성에서 Amazon EC2 인스턴스에 액세스할 수 없고 생성한 인스턴스로 Amazon EC2 인스턴스를 검증할 수 없는 경우 추가 문제 해결이 필요합니다. 예를 들어 연결 문제와 관련된 오류로 인해 동기화가 실패한 Amazon S3 커넥터가 있는 경우 커넥터에 할당한 것과 동일한 Amazon VPC 구성으로 Amazon EC2 인스턴스를 설정할 수 있습니다 Amazon S3 . 그런 다음 이 Amazon EC2 인스턴스를 사용하여 올바르게 Amazon VPC 설정되었는지 테스트하십시오.

다음은 Amazon S3 데이터 소스와의 Amazon VPC 연결 문제를 해결하기 위해 Amazon EC2 인스턴스를 설정하는 예제입니다.

주제

1단계: 인스턴스 Amazon EC2 시작

1. AWS Management Console 로그인하고 https://console.aws.amazon.com/ec2/ 에서 Amazon EC2 콘솔을 엽니다.
2. 인스턴스 시작을 선택합니다.
3. 네트워크 설정을 선택한 다음 편집을 선택한 후 다음을 수행하십시오.
   a. 할당한 것과 동일한 VPC와 서브넷을 선택합니다. Amazon Kendra
   b. 방화벽 (보안 그룹) 의 경우 기존 보안 그룹 선택을 선택합니다. 그런 다음 할당한 보안 그룹을 선택합니다 Amazon Kendra.

      > **Note**
      > 보안 그룹은 아웃바운드 트래픽을 허용해야 합니다 Amazon S3.

   c. 퍼블릭 IP 자동 할당을 비활성화로 설정합니다.
   d. 고급 세부 정보에서 다음을 수행하십시오.

- IAM 인스턴스 프로필의 경우 새 IAM 프로필 생성을 선택하여 IAM 인스턴스 프로필을 생성하고 인스턴스에 연결합니다. 프로필에 액세스 권한이 있는지 확인하십시오. Amazon S3 자세한 내용은 Amazon EC2 인스턴스에 Amazon S3 버킷 액세스 권한을 부여하려면 어떻게 해야 합니까? 를 참조하십시오. in AWS re:Post.
- 다른 모든 설정은 기본값으로 둡니다.

e. Amazon EC2 인스턴스를 검토하고 실행합니다.

2단계: Amazon EC2 인스턴스에 연결

Amazon EC2 인스턴스가 실행되면 인스턴스 세부 정보 페이지로 이동하여 인스턴스에 연결합니다. 이렇게 하려면 Linux 인스턴스용 사용 Amazon EC2 설명서의 EC2 Instance Connect Endpoint를 사용하여 퍼블릭 IPv4 주소 없이 인스턴스에 연결에 나와 있는 단계를 사용하십시오.

3단계: 액세스 테스트 Amazon S3

Amazon EC2 인스턴스 터미널에 연결한 후 AWS CLI 명령을 실행하여 이 프라이빗 서브넷에서 Amazon S3 버킷으로의 연결을 테스트합니다.

Amazon S3 액세스를 테스트하려면 필드에 다음 AWS CLI 명령어를 입력합니다. AWS CLI `aws s3 ls`

AWS CLI 명령이 실행된 후 다음을 검토하십시오.

- 필요한 IAM 권한을 올바르게 설정하고 Amazon S3 설정이 올바르면 Amazon S3 버킷 목록이 표시될 것입니다.
- 와 같은 `Access Denied` 권한 오류가 표시되면 VPC 구성은 올바르지만 IAM 권한 또는 Amazon S3 버킷 정책에 문제가 있는 것일 수 있습니다.

명령 제한 시간이 초과되면 VPC 설정이 잘못되어 Amazon EC2 인스턴스가 서브넷에서 Amazon S3에 액세스할 수 없기 때문에 연결 시간이 초과되었을 수 있습니다. VPC를 재구성하고 다시 시도하세요.

인덱스, 데이터 소스 또는 일괄 업로드된 문서 삭제

이 섹션에서는 인덱스, 인덱스에 있는 문서의 데이터 소스 리포지토리 또는 일괄 업로드한 인덱스의 문서를 삭제하는 방법을 보여줍니다.

주제

인덱스 삭제

인덱스를 더 이상 사용하지 않을 때 Amazon Kendra에서 인덱스를 삭제할 수 있습니다. 예를 들어 다음과 같은 경우 인덱스를 삭제하세요.

- 인덱스를 더 이상 사용하지 않으므로 AWS 계정에 부과되는 비용을 줄이고 싶습니다. Amazon Kendra 인덱스에 대한 쿼리 여부와 상관없이 인덱스를 실행하는 동안 요금이 발생합니다.
- 인덱스를 다른 Amazon Kendra 에디션에 맞게 재구성하려고 합니다. 기존 인덱스를 삭제한 다음, 다른 에디션으로 새 인덱스를 만듭니다.
- 계정의 최대 인덱스 수에 도달했으므로 할당량을 초과하고 싶지 않습니다. 기존 인덱스를 삭제하고 새 인덱스를 추가합니다. 생성할 수 있는 최대 인덱스 수에 대한 자세한 내용은 할당량을 참조하세요.

인덱스를 삭제하려면 콘솔, AWS Command Line Interface, AWS CloudFormation 스크립트 또는 `DeleteIndex` API를 사용합니다. 인덱스를 삭제하면 인덱스와 모든 관련 데이터 소스 및 문서 데이터가 제거됩니다. 인덱스를 삭제해도 원본 문서는 리포지토리에서 제거되지 않습니다.

인덱스 삭제는 비동기식 작업입니다. 인덱스 삭제를 시작하면 인덱스 상태가 `DELETING`으로 변경됩니다. 인덱스와 관련된 모든 정보가 제거될 때까지 `DELETING` 상태를 유지합니다. 인덱스가 삭제되면 ListIndices API에 대한 호출 결과에 더 이상 표시되지 않습니다. 삭제된 인덱스의 식별자를 사용하여 DescribeIndex API를 호출하면 `ResourceNotFound` 예외가 발생합니다.

인덱스를 삭제하려면(콘솔)

1. AWS Management Console에 로그인하고 https://console.aws.amazon.com/kendra/에서 Amazon Kendra 콘솔을 엽니다.

2. 탐색 창에서 인덱스를 선택한 후 삭제할 인덱스를 선택합니다.
3. 삭제를 선택하여 선택한 인덱스를 삭제합니다.

인스턴스를 삭제하려면(CLI)

- AWS CLI에서 다음 명령을 사용합니다. 이 명령은 Linux 및 macOS용으로 형식이 지정됩니다. Windows를 사용하는 경우 Unix 줄 연속 문자(\)를 캐럿(^)으로 바꿉니다.

```
aws kendra delete-index \
   --id index-id
```

데이터 소스 삭제

데이터 소스에 포함된 정보를 Amazon Kendra 인덱스에서 제거하려는 경우 데이터 소스를 삭제합니다. 예를 들어 다음과 같은 경우 데이터 소스를 삭제합니다.

- 데이터 소스가 잘못 구성되었습니다. 데이터 소스를 삭제하고 데이터 소스 삭제가 완료될 때까지 기다린 다음 데이터 소스를 다시 만듭니다.
- 한 데이터 소스에서 다른 데이터 소스로 문서를 마이그레이션했습니다. 원본 데이터 소스를 삭제하고 새 위치에 다시 만듭니다.
- 인덱스의 데이터 소스 수 한도에 도달했습니다. 기존 데이터 소스 중 하나를 삭제하고 새로 추가합니다. 만들 수 있는 데이터 소스 수에 대한 자세한 내용은 할당량 단원을 참조하세요.

데이터 소스를 삭제하려면 콘솔, AWS Command Line Interface(AWS CLI), `DeleteDataSource` API 또는 AWS CloudFormation 스크립트를 사용합니다. 데이터 소스를 삭제하면 데이터 소스에 대한 모든 정보가 인덱스에서 제거됩니다. 데이터 소스의 동기화만 중지하려면 데이터 소스의 동기화 일정을 “온디맨드 실행”으로 변경합니다.

데이터 소스 삭제는 비동기식 작업입니다. 데이터 소스 삭제를 시작하면 데이터 소스 상태가 `DELETING`으로 변경됩니다. 데이터 소스와 관련된 정보가 제거될 때까지 `DELETING` 상태를 유지합니다. 데이터 소스가 삭제된 후에는 ListDataSources API에 대한 호출 결과에 더 이상 표시되지 않습니다. 삭제된 데이터 소스의 식별자를 사용하여 DescribeDataSource API를 호출하면 `ResourceNotFound` 예외가 발생합니다.

> **Note**
>
> 데이터 소스에서 특정 문서를 삭제한 후 전체 데이터 소스를 삭제하거나 인덱스를 다시 동기화하는 작업은 삭제하려는 문서 수에 따라 최대 1시간 이상 걸릴 수 있습니다.

데이터 소스(콘솔)를 삭제하려면

1. AWS Management Console에 로그인하고 https://console.aws.amazon.com/kendra/에서 Amazon Kendra 콘솔을 엽니다.
2. 탐색 창에서 인덱스를 선택한 후 삭제할 데이터 소스를 포함하는 인덱스를 선택합니다.
3. 탐색 창에서 데이터 소스를 선택합니다.
4. 제거할 데이터 소스를 선택합니다.
5. 삭제를 선택하여 데이터 소스를 삭제합니다.

데이터 소스를 삭제하려면(CLI)

- AWS Command Line Interface에서 다음 명령을 사용합니다. 이 명령은 Linux 및 macOS용으로 형식이 지정됩니다. Windows를 사용하는 경우 Unix 줄 연속 문자(\)를 캐럿(^)으로 바꿉니다.

```
aws kendra delete-data-source \
   --id data-source-id \
   --index-id index-id
```

데이터 소스를 삭제하면 Amazon Kendra가 데이터 소스에 대해 저장된 모든 정보가 제거됩니다. Amazon Kendra가 인덱스에 저장된 모든 문서 데이터, 데이터 소스와 관련된 모든 실행 기록 및 지표를 제거합니다. 데이터 소스를 삭제해도 스토리지에서 원본 문서가 제거되지는 않습니다.

Amazon Kendra가 데이터 소스를 삭제하는 동안 `DescribeIndex` API에서 반환되는 문서 수에 데이터 소스의 문서가 포함될 수 있습니다. Amazon Kendra가 데이터 소스를 삭제하는 동안 데이터 소스의 문서가 검색 결과에 표시될 수 있습니다.

`DeleteDataSource` API를 호출하거나 콘솔에서 데이터 소스를 삭제하도록 선택하는 즉시 Amazon Kendra가 데이터 소스의 리소스를 해제합니다. 데이터 소스 수를 한도 미만으로 줄이기 위해 데이터 소스를 삭제하려는 경우 새 데이터 소스를 즉시 만들 수 있습니다.

데이터 소스를 삭제한 다음, 문서 데이터에 다른 데이터 소스를 만드는 경우 새 데이터 소스를 동기화하기 전에 첫 번째 데이터 소스가 삭제될 때까지 기다립니다.

Amazon Kendra와 동기화 중인 데이터 소스를 삭제할 수 있습니다. 동기화가 중지되고 데이터 소스가 제거됩니다. 데이터 소스가 삭제되는 동안 동기화를 시작하려고 하면 `ConflictException` 예외가 발생합니다.

관련 인덱스가 `DELETING` 상태에 있는 경우 데이터 소스를 삭제할 수 없습니다. 인덱스를 삭제하면 해당 인덱스의 모든 데이터 소스가 삭제됩니다. 해당 인덱스의 데이터 소스가 `DELETING` 상태에 있는 동안 인덱스 삭제를 시작할 수 있습니다.

두 개의 데이터 소스가 동일한 문서를 가리키는 경우(예: 두 데이터 소스가 동일한 Amazon S3 버킷을 가리키는 경우) 데이터 소스 중 하나가 삭제되면 인덱스의 문서가 일관되지 않을 수 있습니다. 두 데이터 소스가 동일한 문서를 참조하는 경우 문서 데이터 사본 하나만 인덱스에 저장됩니다. 데이터 소스 하나를 제거하면 문서의 인덱스 데이터가 제거됩니다. 다른 데이터 소스는 문서가 제거되었음을 인식하지 못하므로 Amazon Kendra가 다음에 동기화할 때 문서를 제대로 다시 인덱싱하지 않습니다. 동일한 문서 위치를 가리키는 데이터 소스가 두 개 있는 경우, 두 데이터 소스를 모두 삭제한 다음, 하나를 다시 만들어야 합니다.

일괄 업로드된 문서 삭제

BatchDeleteDocument API를 사용하여 인덱스에서 직접 문서를 삭제할 수 있습니다. 콘솔을 사용하여 문서를 직접 삭제할 수 없습니다. 콘솔을 사용하는 경우 데이터 소스 리포지토리에서 특정 문서를 삭제하고 인덱스와 다시 동기화하거나 전체 데이터 소스 커넥터를 삭제할 수 있습니다.

`BatchDeleteDocument`를 사용하여 인덱스에서 문서를 삭제하는 것은 비동기식 작업입니다. `BatchDeleteDocument` API를 호출한 후에는 BatchGetDocumentStatus API를 사용하여 문서 삭제 진행 상황을 모니터링합니다. 인덱스에서 문서가 삭제되면 Amazon Kendra가 상태로 `NOT_FOUND`를 반환합니다.

> **Note**
>
> `BatchDeleteDocument`를 사용하여 인덱스에서 문서를 삭제하는 경우 삭제하려는 문서 수에 따라 최대 1시간 이상 걸릴 수 있습니다.

인덱스에서 일괄 업로드된 문서를 삭제하려면(CLI)

- AWS Command Line Interface에서 다음 명령을 사용합니다. 이 명령은 Linux 및 macOS용으로 형식이 지정됩니다. Windows를 사용하는 경우 Unix 줄 연속 문자(\)를 캐럿(^)으로 바꿉니다.

```
aws kendra batch-delete-document \
   --index-id index-id \
   --document-id-list 'doc-id-1' 'doc-id-2'
```

수집 중에 문서 보강

문서 수집 프로세스 중 콘텐츠 및 문서 메타데이터 필드 또는 속성을 변경할 수 있습니다. Amazon Kendra의 사용자 지정 문서 보강 기능을 사용하면 문서를 Amazon Kendra로 수집할 때 문서 속성 및 콘텐츠를 생성, 수정 또는 삭제할 수 있습니다. 즉, 필요에 따라 데이터를 조작하고 수집할 수 있습니다.

이 기능을 사용하면 문서가 처리되고 Amazon Kendra로 수집되는 방식을 제어할 수 있습니다. 예를 들어 문서를 Amazon Kendra로 수집하는 동안 문서 메타데이터에서 개인 식별 정보를 스크러빙할 수 있습니다.

이 기능을 사용할 수 있는 또 한 가지 방법은 AWS Lambda의 Lambda 함수를 호출하여 이미지의 광학 문자 인식(OCR), 텍스트 번역 및 기타 검색 또는 분석을 위한 데이터 준비 작업을 실행하는 것입니다. 예를 들어, 이미지에서 OCR을 실행하는 함수를 호출할 수 있습니다. 함수는 이미지의 텍스트를 해석하고 각 이미지를 텍스트 문서로 취급할 수 있습니다. 우편으로 고객 설문조사를 받아 설문조사를 이미지로 저장하는 회사는 이러한 이미지를 Amazon Kendra에 텍스트 문서로 수집할 수 있습니다. 그러면 회사는 Amazon Kendra에서 중요한 고객 설문조사 정보를 검색할 수 있습니다.

기본 작업을 사용하여 데이터의 첫 번째 구문 분석으로 적용한 다음, Lambda 함수를 사용하여 데이터에 더 복잡한 작업을 적용할 수 있습니다. 예를 들어, 기본 작업을 사용하여 문서 메타데이터 필드 'Customer_ID'에서 모든 값을 간단히 제거한 다음, Lambda 함수를 적용하여 문서의 텍스트 이미지에서 텍스트를 추출할 수 있습니다.

사용자 지정 문서 보강 작동 방식

사용자 지정 문서 보강의 전체 프로세스는 다음과 같습니다.

1. 사용자 지정 문서 보강은 데이터 소스를 만들거나 업데이트할 때 구성하거나 문서를 Amazon Kendra로 직접 인덱싱합니다.
2. Amazon Kendra는 인라인 구성 또는 기본 로직을 적용하여 데이터를 변경합니다. 자세한 내용은 the section called “메타데이터를 변경하는 기본 작업” 섹션을 참조하세요.
3. 고급 데이터 조작을 구성하기로 선택한 경우 Amazon Kendra가 원본, 원시 문서 또는 구조화되고 파싱된 문서에 이를 적용할 수 있습니다. 자세한 내용은 the section called “Lambda 함수: 메타데이터 또는 콘텐츠 추출 및 변경” 섹션을 참조하세요.
4. 변경된 문서는 Amazon Kendra에 수집됩니다.

구성이 유효하지 않으면 이 프로세스의 어느 시점에서든 Amazon Kendra에 오류가 발생합니다.

CreateDataSource, UpdateDataSource 또는 BatchPutDocument API를 호출하면 사용자 지정 문서 보강 구성을 제공합니다. `BatchPutDocument`를 호출하는 경우 각 요청마다 사용자 지정 문서 보강을 구성해야 합니다. 콘솔을 사용하는 경우 인덱스를 선택한 다음, 문서 보강을 선택하여 사용자 지정 문서 보강을 구성합니다.

콘솔에서 문서 보강을 사용하는 경우, API를 사용하는 것처럼 기본 작업만 구성하거나 Lambda 함수만 구성하거나 둘 다 구성하도록 선택할 수 있습니다. 콘솔 단계에서 다음을 선택하여 기본 작업을 구성하지 않고 Lambda 함수만 구성하도록 선택할 수 있습니다. 여기에는 원본 (추출 전) 데이터에 적용할지 아니면 구조화된 (추출 후) 데이터에 적용할지 여부가 포함됩니다. 콘솔에서 모든 단계를 완료해야만 구성을 저장할 수 있습니다. 모든 단계를 완료하지 않으면 문서 구성이 저장되지 않습니다.

메타데이터를 변경하는 기본 작업

기본 로직을 사용하여 문서 필드와 콘텐츠를 조작할 수 있습니다. 여기에는 필드의 값 제거, 조건을 사용한 필드 값 수정 또는 필드 생성이 포함됩니다. 기본 로직을 사용하여 조작할 수 있는 범위를 넘어서는 고급 조작의 경우 Lambda 함수를 호출합니다. 자세한 내용은 the section called “Lambda 함수: 메타데이터 또는 콘텐츠 추출 및 변경” 섹션을 참조하세요.

기본 로직을 적용하려면 DocumentAttributeTarget 객체를 사용하여 조작하려는 대상 필드를 지정합니다. 속성 키를 제공합니다. 예를 들어, 'Department' 키는 문서와 연결된 모든 부서 이름을 포함한 필드 또는 속성입니다. 특정 조건이 충족되는 경우 대상 필드에 사용할 값을 지정할 수도 있습니다. 이 조건은 DocumentAttributeCondition 객체를 사용하여 설정합니다. 예를 들어 'Source_URI' 필드의 URI 값에 'financial'이 포함된 경우 대상 필드인 'Department'를 문서의 대상 값인 'Finance'로 미리 채웁니다. 대상 문서 속성 값도 삭제할 수 있습니다.

콘솔을 사용하여 기본 로직을 적용하려면 인덱스를 선택한 다음, 탐색 메뉴에서 문서 보강을 선택합니다. 기본 작업 구성으로 이동하여 문서 필드 및 콘텐츠에 기본 조작을 적용합니다.

다음은 기본 로직을 사용하여 'Customer_ID'라는 문서 필드의 모든 고객 식별 번호를 제거하는 예제입니다.

예 1: 문서와 관련된 고객 식별 번호 제거

기본 조작이 적용되기 전의 데이터.

Document_ID	Body_Text	Customer_ID
1	Lorem Ipsum.	CID1234
2	Lorem Ipsum.	CID1235

Document_ID	Body_Text	Customer_ID
3	Lorem Ipsum.	CID1236

기본 조작이 적용된 이후의 데이터.

Document_ID	Body_Text	Customer_ID
1	Lorem Ipsum.	
2	Lorem Ipsum.	
3	Lorem Ipsum.	

다음은 기본 로직을 사용하여 'Department'라는 필드를 만들고 'Source_URI' 필드의 정보에 기반한 부서 이름을 이 필드에 미리 채우는 예제입니다. 이 경우는 'Source_URI' 필드의 URI 값에 'financial'이 포함된 경우 대상 필드인 'Department'를 문서의 대상 값인 'Finance'로 미리 채우는 조건을 사용합니다.

예 2: '부서' 필드를 만들고 조건을 사용하여 문서와 관련된 부서 이름으로 필드를 미리 채웁니다.

기본 조작이 적용되기 전의 데이터.

Document_ID	Body_Text	Source_URI
1	Lorem Ipsum.	financial/1
2	Lorem Ipsum.	financial/2
3	Lorem Ipsum.	financial/3

기본 조작이 적용된 이후의 데이터.

Document_ID	Body_Text	Source_URI	Department
1	Lorem Ipsum.	financial/1	Finance
2	Lorem Ipsum.	financial/2	Finance

Document_ID	Body_Text	Source_URI	Department
3	Lorem Ipsum.	financial/3	Finance

> **Note**
>
> Amazon Kendra는 아직 인덱스 필드로 생성되지 않은 대상 문서 필드를 생성할 수 없습니다. 인덱스 필드를 만든 후 `DocumentAttributeTarget`을 사용하여 문서 필드를 생성할 수 있습니다. 그러면 Amazon Kendra가 새로 만든 문서 메타데이터 필드를 인덱스 필드에 매핑합니다.

다음 코드는 문서와 관련된 고객 식별 번호를 제거하도록 기본 데이터 조작을 구성하는 예제입니다.

Console

고객 식별 번호를 제거하도록 기본 데이터 조작을 구성하려면

1. 왼쪽 탐색 창의 인덱스에서 문서 보강을 선택한 다음, 문서 보강 추가를 선택합니다.
2. 기본 작업 구성 페이지의 드롭다운에서 문서 필드 및 콘텐츠를 변경하려는 데이터 소스를 선택합니다. 그런 다음 드롭다운에서 문서 필드 이름 'Customer_ID'를 선택하고 드롭다운에서 인덱스 필드 이름 'Customer_ID'를 선택한 다음, 드롭다운에서 대상 작업 삭제를 선택합니다. 그런 다음 기본 작업 추가를 선택합니다.

CLI

고객 식별 번호를 제거하도록 기본 데이터 조작을 구성하려면

```
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

고객 식별 번호를 제거하도록 기본 데이터 조작을 구성하려면

```
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

고객 식별 번호를 제거하도록 기본 데이터 조작을 구성하려면

```
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");
    }
}
```

Lambda 함수: 메타데이터 또는 콘텐츠 추출 및 변경

Lambda 함수를 사용하여 문서 필드와 콘텐츠를 조작할 수 있습니다. 기본 로직을 넘어 고급 데이터 조작을 적용하려는 경우에 유용합니다. 이미지에서 텍스트를 해석하고 각 이미지를 텍스트 문서로 처리하는 광학 문자 인식(OCR) 사용을 예로 들 수 있습니다. 또는 특정 시간대의 현재 날짜-시간을 검색하고 날짜 필드에 빈 값이 있을 경우 날짜-시간을 삽입하는 경우에도 해당합니다.

기본 로직을 먼저 적용한 다음 Lambda 함수를 사용하여 데이터를 추가로 조작하거나 그 반대로 적용할 수 있습니다. Lambda 함수만 적용하도록 선택할 수도 있습니다.

Amazon Kendra는 Lambda 함수를 호출하여 수집 프로세스 중에 CustomDocumentEnrichmentConfiguration의 일부로 고급 데이터 조작을 적용할 수 있습니다. Lambda 함수를 실행하고 Amazon S3 버킷에 액세스하여 데이터 조작의 출력을 저장할 권한이 포함된 역할을 지정합니다. IAM 액세스 역할을 참조하세요.

Amazon Kendra는 원본, 원시 문서 또는 구문 분석된 정형 문서에 Lambda 함수를 적용할 수 있습니다. 원본 또는 원시 데이터를 가져와 PreExtractionHookConfiguration을 사용하여 데이터 조작을 적용하는 Lambda 함수를 구성할 수 있습니다. 또한 PostExtractionHookConfiguration을 사용하여 구조화된 문서를 가져와 데이터 조작을 적용하는 Lambda 함수를 구성할 수 있습니다. Amazon Kendra는 문서 메타데이터와 텍스트를 추출하여 문서를 구조화합니다. Lambda 함수는 필수 요청 및 응답 구조를 따라야 합니다. 자세한 내용은 the section called “Lambda 함수에 대한 데이터 계약” 섹션을 참조하세요.

콘솔에서 Lambda 함수를 구성하려면 인덱스를 선택한 다음 탐색 메뉴에서 문서 보강을 선택합니다. Lambda 함수 구성으로 이동하여 Lambda 함수를 구성합니다.

`PreExtractionHookConfiguration` 및 `PostExtractionHookConfiguration`에 대해 각각 Lambda 함수를 하나씩만 구성할 수 있습니다. 그러나 Lambda 함수는 필요한 다른 함수를 호출할 수 있습니다. `PreExtractionHookConfiguration` 및 `PostExtractionHookConfiguration` 또는 둘 중 하나를 구성할 수 있습니다. `PreExtractionHookConfiguration`의 Lambda 함수는 실행 시간 5분을 초과할 수 없고 `PostExtractionHookConfiguration`의 Lambda 함수는 실행 시간 1분을 초과할 수 없습니다. 사용자 지정 문서 보강을 구성하면 당연히 구성하지 않은 경우보다 Amazon Kendra에 문서를 수집하는 시간이 더 오래 걸립니다.

조건이 충족되는 경우에만 Lambda 함수를 호출하도록 Amazon Kendra를 구성할 수 있습니다. 예를 들어, 빈 날짜/시간 값이 있는 경우 Amazon Kendra가 현재 날짜/시간을 삽입하는 함수를 호출해야 한다는 조건을 지정할 수 있습니다.

다음은 Lambda 함수를 사용하여 이미지의 텍스트를 해석하는 OCR을 실행하고 이 텍스트를 'Document_Image_Text'라는 필드에 저장하는 예제입니다.

예 1: 이미지에서 텍스트를 추출하여 텍스트 문서 생성

고급 조작이 적용되기 전의 데이터.

Document_ID	Document_Image
1	image_1.png
2	image_2.png
3	image_3.png

고급 조작이 적용된 이후의 데이터.

Document_ID	Document_Image	Document_Image_Text
1	image_1.png	우편으로 보낸 설문조사 응답
2	image_2.png	우편으로 보낸 설문조사 응답
3	image_3.png	우편으로 보낸 설문조사 응답

다음은 Lambda 함수를 사용하여 빈 날짜 값에 현재 날짜-시간을 삽입하는 예제입니다. 날짜 필드 값이 'null'인 경우 이를 현재 날짜-시간으로 대체한다는 조건을 사용합니다.

예 2: Last_Update 필드의 빈 값을 현재 날짜-시간으로 교체.

고급 조작이 적용되기 전의 데이터.

Document_ID	Body_Text	Last_Updated
1	Lorem Ipsum.	2020년 1월 1일
2	Lorem Ipsum.	
3	Lorem Ipsum.	2020년 7월 1일

고급 조작이 적용된 이후의 데이터.

Document_ID	Body_Text	Last_Updated
1	Lorem Ipsum.	2020년 1월 1일
2	Lorem Ipsum.	2021년 12월 1일
3	Lorem Ipsum.	2020년 7월 1일

다음 코드는 원시, 원본 데이터에 대한 고급 데이터 조작을 위해 Lambda 함수를 구성하는 예제입니다.

Console

원시, 원본 데이터에 대한 고급 데이터 조작을 위해 Lambda 함수를 구성하려면

1. 왼쪽 탐색 창의 인덱스에서 문서 보강을 선택한 다음, 문서 보강 추가를 선택합니다.
2. Lambda 함수 구성 페이지의 사전 추출용 Lambda 섹션에서 드롭다운의 Lambda 함수 ARN 및 Amazon S3 버킷을 선택합니다. 드롭다운에서 새 역할을 생성하는 옵션을 선택하여 IAM 액세스 역할을 추가합니다. 이렇게 하면 문서 보강을 만드는 데 필요한 Amazon Kendra 권한이 생성됩니다.

CLI

원시, 원본 데이터에 대한 고급 데이터 조작을 위해 Lambda 함수를 구성하려면

```
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

원시, 원본 데이터에 대한 고급 데이터 조작을 위해 Lambda 함수를 구성하려면

```
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

원시, 원본 데이터에 대한 고급 데이터 조작을 위해 Lambda 함수를 구성하려면

```
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");
    }
}
```

Lambda 함수에 대한 데이터 계약

고급 데이터 조작을 위한 Lambda 함수는 Amazon Kendra 데이터 계약과 상호 작용합니다. 이 계약은 Lambda 함수의 필수 요청 및 응답 구조입니다. Lambda 함수가 이러한 구조를 따르지 않으면 Amazon Kendra에 오류가 발생합니다.

`PreExtractionHookConfiguration`의 Lambda 함수는 다음과 같은 요청 구조를 예상해야 합니다.

```
{
    "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>
}
```

CustomDocumentAttribute 구조를 포함하는 metadata 구조는 다음과 같습니다.

```
{
    "attributes": [<CustomDocumentAttribute<]
}

CustomDocumentAttribute
{
    "name": <str>,
    "value": <CustomDocumentAttributeValue>
}

CustomDocumentAttributeValue
{
    "stringValue": <str>,
    "integerValue": <int>,
    "longValue": <long>,
    "stringListValue": list<str>,
    "dateValue": <str>
}
```

PreExtractionHookConfiguration에 대한 Lambda 함수는 다음 응답 구조를 준수해야 합니다.

```
{
    "version": <str>,
    "dataBlobStringEncodedInBase64": <str>, //In the case of a data blob
    "s3ObjectKey": <str>, //In the case of an S3 bucket
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

PostExtractionHookConfiguration의 Lambda 함수는 다음과 같은 요청 구조를 예상해야 합니다.

```
{
    "version": <str>,
    "s3Bucket": <str>,
    "s3ObjectKey": <str>,
```

```
    "metadata": <Metadata>
}
```

PostExtractionHookConfiguration에 대한 Lambda 함수는 다음 응답 구조를 준수해야 합니다.

```
PostExtractionHookConfiguration Lambda Response
{
    "version": <str>,
    "s3ObjectKey": <str>,
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

변경된 문서가 Amazon S3버킷에 업로드됩니다. 변경된 문서는 the section called “구조화된 문서 형식”에 표시된 형식을 따라야 합니다.

구조화된 문서 형식

Amazon Kendra는 구조화된 문서를 지정된 Amazon S3 버킷에 업로드합니다. 구조화된 문서는 다음 형식을 따릅니다.

```
Kendra document

{
   "textContent": <TextContent>
}

TextContent
{
  "documentBodyText": <str>
}
```

데이터 계약을 준수하는 Lambda 함수의 예

다음 Python 코드는 원시 또는 원본 문서의 본문 콘텐츠와 메타데이터 필드 _authors, _document_title의 고급 조작을 적용하는 Lambda 함수의 예입니다.

본문 콘텐츠가 Amazon S3 버킷에 있는 경우

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

본문 콘텐츠가 데이터 블롭에 있는 경우

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

다음 Python 코드는 구조화되었거나 구문 분석된 문서의 본문 콘텐츠와 메타데이터 필드 _authors, _document_title의 고급 조작을 적용하는 Lambda 함수의 예입니다.

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

인덱스 검색

Amazon Kendra 색인을 검색하려면 쿼리 API를 사용합니다. `Query` API는 애플리케이션에서 사용하는 인덱싱된 문서에 대한 정보를 반환합니다. 이 섹션에서는 쿼리를 만들고, 필터를 수행하고, `Query` API에서 받은 응답을 해석하는 방법을 보여줍니다.

색인을 생성하는 데 사용한 문서를 검색하려면 Amazon Kendra AMAZON을 Amazon Lex사용하십시오. KendraSearchIntent. 를 Amazon Kendra 사용하여 Amazon Lex구성하는 예제는 Amazon Kendra 인덱스용 FAQ 봇 만들기를 참조하십시오.

주제

인덱스 쿼리.

색인을 검색할 때 는 문서에 대해 제공한 모든 정보를 Amazon Kendra 사용하여 입력한 검색어와 가장 관련이 있는 문서를 결정합니다. Amazon Kendra 고려해야 할 몇 가지 항목은 다음과 같습니다.

- 문서의 텍스트 또는 본문.
- 문서의 제목.
- 검색 가능으로 표시한 사용자 지정 텍스트 필드.
- 지정한 날짜 필드는 문서의 “최신성”을 결정하는 데 사용해야 합니다.
- 관련 정보를 제공할 수 있는 기타 모든 필드.

Amazon Kendra 또한 검색에 설정했을 수 있는 모든 필드/속성 필터를 기반으로 응답을 필터링할 수 있습니다. 예를 들어, “부서”라는 사용자 지정 필드가 있는 경우 “법률”이라는 부서의 문서만 반환하도록 응답을 필터링할 수 있습니다. 자세한 내용은 사용자 지정 필드 또는 속성을 참조하세요.

반환된 검색 결과는 각 문서에 대해 Amazon Kendra 결정되는 관련성에 따라 정렬됩니다. 결과는 페이지로 구분되므로 사용자에게 한 번에 한 페이지씩 표시될 수 있습니다.

인덱싱한 문서를 검색하려면 Amazon Kendra AMAZON을 Amazon Lex사용하십시오. KendraSearchIntent. 를 Amazon Kendra 사용하여 Amazon Lex구성하는 예제는 Amazon Kendra 인덱스용 FAQ 봇 만들기를 참조하십시오.

다음 예제는 인덱스를 검색하는 방법을 보여줍니다. Amazon Kendra 쿼리에 가장 적합한 검색 결과 유형 (답변, 문서, 질문-답변) 을 결정합니다. 쿼리에 특정 유형의 검색 응답 (답변, 문서, 질문-답변) 을 Amazon Kendra 반환하도록 구성할 수 없습니다.

쿼리 응답에 대한 자세한 내용은 쿼리 응답 및 응답 유형를 참조하세요.

필수 조건

Query API를 사용하여 인덱스를 쿼리하기 전에:

- 인덱스에 필요한 권한을 설정하고 데이터 소스에 연결하거나 문서를 일괄 업로드하세요. 자세한 내용은 IAM 역할을 참조하세요. API를 호출하여 인덱스 및 데이터 소스 커넥터를 생성하거나 문서를 일괄 업로드할 때는 역할의 Amazon 리소스 이름을 사용합니다.
- AWS Command Line Interface, SDK를 설정하거나 콘솔로 이동합니다. Amazon Kendra 자세한 내용은 Amazon Kendra설정을 참조하십시오.
- 인덱스를 만들고 문서의 데이터 소스에 연결하거나 문서를 일괄 업로드하세요. 자세한 내용은 인덱스 생성 및 데이터 소스 커넥터 생성을 참조하세요.

인덱스 검색 (콘솔)

Amazon Kendra 콘솔을 사용하여 색인을 검색하고 테스트할 수 있습니다. 쿼리를 만들고 결과를 볼 수 있습니다.

콘솔로 인덱스를 검색하려면

1. http://console.aws.amazon.com/kendra/ 에서 AWS Management Console 로그인하고 Amazon Kendra 콘솔을 엽니다.
2. 탐색 창에서 인덱스를 선택합니다.

3. 인덱스를 선택합니다.
4. 탐색 메뉴에서 인덱스 검색 옵션을 선택합니다.
5. 텍스트 상자에 쿼리를 입력한 후 Enter 키를 누릅니다.
6. Amazon Kendra 검색 결과를 반환합니다.

사이드 패널에서 전구 아이콘을 선택하여 검색에 대한 쿼리 ID를 얻을 수도 있습니다.

인덱스 검색 (SDK)

Python 또는 Java로 인덱스를 검색하려면

- 다음 예제에서는 인덱스를 검색합니다. `query`의 값을 검색 쿼리로, `index_id` 또는 `indexId`를 검색하려는 인덱스의 인덱스 식별자로 변경합니다.

  Query API를 호출할 때 응답 요소의 일부로 검색에 대한 쿼리 ID를 가져올 수도 있습니다.

  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");
        }
    }
}
```

인덱스 검색 (Postman)

Postman을 사용하여 Amazon Kendra 색인을 쿼리하고 테스트할 수 있습니다.

Postman을 사용하여 인덱스를 검색하려면

1. Postman에서 새 컬렉션을 만들고 요청 유형을 POST로 설정합니다.
2. 엔드포인트 URL을 입력합니다. 예: https://kendra.<region>.amazonaws.com.
3. 승인 탭을 선택하고 다음 정보를 입력합니다.

   - 유형 - AWS 서명을 선택합니다.
   - AccessKey—사용자를 생성할 때 생성된 액세스 키를 입력합니다. IAM
   - SecretKey—사용자를 생성할 때 생성된 비밀 키를 입력합니다. IAM
   - AWS 지역 - 색인의 지역을 입력합니다. 예: us-west-2.
   - 서비스 이름 - kendra를 입력합니다. 대소문자를 구분하므로 소문자여야 합니다.

> **Warning**
>
> 서비스 이름을 잘못 입력하거나 소문자를 사용하지 않는 경우 전송을 선택하여 요청을 보내면 “자격 증명 범위가 올바른 서비스 'kendra'로 지정되어야 합니다.”라는 오류가 발생합니다.
> 또한 올바른 액세스 키 및 보안 암호 키를 입력했는지도 확인해야 합니다.

4. 헤더 탭을 선택하고 다음 키 및 값 정보를 입력합니다.

   - 키: X-Amz-Target

     값: com.amazonaws.kendra. AWSKendraFrontendService. 쿼리
   - 키: Content-Encoding

     값: amz-1.0
5. 본문 탭을 선택하고 다음을 수행합니다.

   - 요청 본문의 원시 JSON 유형을 선택합니다.
   - 인덱스 ID와 쿼리 텍스트가 포함된 JSON을 입력합니다.

```
{
    "IndexId": "index-id",
    "QueryText": "enter a query here"
}
```

> **Warning**
>
> JSON에서 올바른 들여쓰기를 사용하지 않는 경우 오류가 발생합니다: "". SerializationException JSON의 들여쓰기를 확인하세요.

6. 전송(오른쪽 상단 근처)을 선택합니다.

고급 쿼리 구문으로 검색

고급 쿼리 구문이나 연산자를 사용하여 단순 키워드나 자연어 쿼리보다 더 구체적인 쿼리를 만들 수 있습니다. 여기에는 범위, 부울, 와일드카드 등이 포함됩니다. 연산자를 사용하면 쿼리에 더 많은 컨텍스트를 제공하고 검색 결과를 더 세분화할 수 있습니다.

Amazon Kendra 다음 연산자를 지원합니다.

- 부울: 검색을 제한하거나 넓히는 로직. 예를 들어, `amazon AND sports`는 두 용어가 모두 포함된 문서만 검색하도록 검색을 제한합니다.
- 괄호: 중첩된 쿼리 용어를 우선 순위에 따라 읽습니다. 예를 들어, `(amazon AND sports) NOT rainforest`는 `(amazon AND sports)`를 `NOT rainforest`보다 먼저 읽습니다.
- 범위: 날짜 또는 숫자 범위 값. 범위는 포함, 제외 또는 제한 없음일 수 있습니다. 예를 들어, 해당 날짜를 포함하여 2020년 1월 1일에서 2020년 12월 31일 사이에 마지막으로 업데이트된 문서를 검색할 수 있습니다.
- 필드: 특정 필드를 사용하여 검색을 제한합니다. 예를 들어, '위치' 필드에 '미국'이 있는 문서를 검색할 수 있습니다.
- 와일드카드: 텍스트 문자열과 부분적으로 일치합니다. 예를 들어, `Cloud*` 일치할 수 CloudFormation 있습니다. Amazon Kendra 현재는 후행 와일드카드만 지원합니다.
- 정확한 따옴표: 텍스트 문자열과 정확히 일치합니다. `"Amazon Kendra" "pricing"`을 포함하는 문서를 예로 들 수 있습니다.

위 연산자를 여럿 조합하여 사용할 수 있습니다.

연산자를 과도하게 사용하거나 매우 복잡한 쿼리를 사용하면 쿼리 지연 시간에 영향을 줄 수 있습니다. 와일드카드는 지연 시간 측면에서 가장 비용이 많이 드는 연산자입니다. 일반적으로 사용하는 용어와 연산자가 많을수록 지연 시간에 미치는 잠재적 영향도 커집니다. 지연 시간에 영향을 미치는 다른 요인으로는 색인된 문서의 평균 크기, 색인 크기, 검색 결과에 대한 필터링, 색인의 전체 부하 등이 있습니다. Amazon Kendra

불

부울 연산자 `AND`, `OR`, `NOT`를 사용하여 여러 단어를 결합하거나 제외할 수 있습니다.

다음은 부울 연산자를 사용하는 예제입니다.

`amazon AND sports`

텍스트에 '아마존'과 '스포츠'라는 용어가 모두 포함된 검색 결과(예: Amazon Prime 비디오 스포츠 또는 기타 유사한 콘텐츠)를 반환합니다.

`sports OR recreation`

텍스트에 '스포츠' 또는 '레크리에이션' 또는 둘 다 포함된 검색 결과를 반환합니다.

amazon NOT rainforest

텍스트에 '아마존'이라는 용어는 포함되지만 '열대우림'이라는 용어는 포함되지 않은 검색 결과를 반환합니다. 이는 아마존 열대우림이 아니라 아마존이라는 회사에 관한 문서를 검색하기 위한 것입니다.

괄호

괄호를 사용하여 우선순위에 따라 중첩된 단어를 쿼리할 수 있습니다. 괄호는 쿼리를 읽는 Amazon Kendra 방법을 나타냅니다.

다음은 괄호 연산자를 사용하는 예제입니다.

(amazon AND sports) NOT rainforest

텍스트에 '아마존'과 '스포츠'라는 용어가 모두 포함되지만 '열대우림'이라는 용어는 포함되지 않은 문서를 반환합니다. 아마존 열대우림의 어드벤처 스포츠가 아닌 Amazon Prime 비디오 스포츠 또는 기타 유사한 콘텐츠를 검색하기 위한 것입니다. 괄호는 `amazon AND sports`가 `NOT rainforest`보다 먼저 읽혀야 된다는 지시입니다. 쿼리가 다음과 같이 읽히면 안 됩니다. `amazon AND (sports NOT rainforest)`

(amazon AND (sports OR recreation)) NOT rainforest

'스포츠' 또는 '레크리에이션'이라는 용어 또는 둘 다와 '아마존'이라는 용어가 포함된 문서를 반환합니다. 하지만 '열대우림'이라는 용어는 포함되지 않습니다. 아마존 열대우림의 어드벤처 스포츠가 아닌 Amazon Prime 비디오 스포츠 또는 레크리에이션을 검색하기 위한 것입니다. 괄호는 `NOT rainforest`가 먼저 읽히지만, '아마존'과 결합하기 전에 `sports OR recreation`가 읽혀야 한다는 지시입니다. 쿼리가 다음과 같이 읽히면 안 됩니다. `amazon AND (sports OR (recreation NOT rainforest))`

범위

값 범위를 사용하여 검색 결과를 필터링할 수 있습니다. 속성과 범위 값을 지정합니다. 날짜 또는 숫자 유형일 수 있습니다.

날짜 범위는 다음 형식이어야 합니다.

- Epoch
- YYYY
- YYYY-mm
- YYYY-mm-dd
- YYYY-mm-dd'T'HH

범위의 하위 값과 상위 값을 포함할지 또는 제외할지도 지정할 수 있습니다.

다음은 부울 연산자를 사용하는 예제입니다.

`_processed_date:>2019-12-31 AND _processed_date:<2021-01-01`

2020년(2019년 12월 31일 이후부터 2021년 1월 1일 이전까지)에 처리된 문서를 반환합니다.

`_processed_date:>=2020-01-01 AND _processed_date:<=2020-12-31`

2020년(2020년 1월 1일부터 2020년 12월 31일까지)에 처리된 문서를 반환합니다.

`_document_likes:<1`

좋아요가 없거나 사용자 피드백이 없는 문서(좋아요가 1개 미만)인 문서를 반환합니다.

범위를 주어진 범위 값을 포함하는 것으로 처리할지 아니면 제외로 처리할지 지정할 수 있습니다.

Inclusive

`_last_updated_at:[2020-01-01 TO 2020-12-31]`

2020년 12월 1일과 2020년 12월 31일을 포함하여 2020년에 마지막으로 업데이트된 문서를 반환합니다.

Exclusive

`_last_updated_at:{2019-12-31 TO 2021-01-01}`

2019년 12월 31일과 2021년 1월 1일을 제외하고 2020년에 마지막으로 업데이트된 문서를 반환합니다.

포함하지도 제외하지도 않는, 제한 없는 범위의 경우 < and > 연산자를 사용하면 됩니다. 예제: `_last_updated_at:>2019-12-31 AND _last_updated_at:<2021-01-01`

필드

특정 필드의 값을 충족하는 문서만 반환하도록 검색을 제한할 수 있습니다. 필드는 모든 유형이 될 수 있습니다.

다음은 필드 수준 컨텍스트 연산자를 사용하는 예제입니다.

`status:"Incomplete" AND financial_year:2021`

상태가 미완성인 2021 회계연도의 문서를 반환합니다.

`(sports OR recreation) AND country:"United States" AND level:"professional"`

미국 내 프로 스포츠 또는 레크리에이션에 관한 문서를 반환합니다.

와일드카드

와일드카드 연산자를 사용하여 다양한 단어와 구문을 포함하도록 검색 범위를 넓힐 수 있습니다. 이는 이름 변형을 검색할 때 유용합니다. Amazon Kendra 현재는 후행 와일드카드만 지원합니다. 후행 와일드카드의 접두사 문자 수는 2보다 커야 합니다.

다음은 와일드카드 연산자를 사용하는 예제입니다.

`Cloud*`

및 와 같은 CloudFormation 변형이 포함된 문서를 반환합니다. CloudWatch

`kendra*aws`

kendra.amazonaws 같은 변형을 포함하는 문서를 반환합니다.

`kendra*aws*`

kendra.amazonaws.com 같은 변형을 포함하는 문서를 반환합니다.

정확한 따옴표

따옴표를 사용하여 텍스트와 정확히 일치하는 항목을 검색할 수 있습니다.

다음은 따옴표를 사용하는 예제입니다.

`"Amazon Kendra" "pricing"`

'Amazon Kendra' 문구와 '가격'이라는 용어가 모두 포함된 문서를 반환합니다. 결과를 반환하려면 문서에 'Amazon Kendra'와 '가격'이 모두 포함되어야 합니다.

`"Amazon Kendra" "pricing" cost`

'Amazon Kendra' 문구와 '가격'이라는 용어, 선택적으로 '비용'이라는 용어가 모두 포함된 문서를 반환합니다. 문서에 'Amazon Kendra'와 '가격'이 모두 포함되어야 결과에 표시되지만 '비용'은 반드시 포함되지 않을 수 있습니다.

잘못된 쿼리 구문

Amazon Kendra 쿼리 구문에 문제가 있거나 쿼리가 현재 지원되지 않는 경우 경고를 표시합니다 Amazon Kendra. 자세한 내용은 쿼리 경고에 대한 API 문서를 참조하세요.

다음 쿼리는 잘못된 쿼리 구문의 예입니다.

`_last_updated_at:<2021-12-32`

잘못된 날짜. Amazon Kendra에서 사용하는 그레고리력에는 32일이 존재하지 않습니다.

`_view_count:ten`

잘못된 숫자 값. 숫자 값을 나타내려면 숫자를 사용해야 합니다.

`nonExistentField:123`

잘못된 필드 검색. 필드 검색을 사용하려면 필드가 존재해야 합니다.

`Product:[A TO D]`

잘못된 범위. 범위에는 숫자 값이나 날짜를 사용해야 합니다.

`OR Hello`

잘못된 부울. 연산자는 용어와 함께 사용하고 용어 사이에 배치해야 합니다.

언어로 검색

지원되는 언어로 문서를 검색할 수 있습니다. 에 언어 코드를 AttributeFilter전달하여 필터링된 문서를 선택한 언어로 반환합니다. 지원되는 언어로 쿼리를 입력할 수 있습니다.

언어를 지정하지 않으면 기본적으로 영어로 문서를 Amazon Kendra 쿼리합니다. 코드를 포함하여 지원되는 언어에 대한 자세한 내용은 영어 이외의 언어로 문서 추가를 참조하세요.

콘솔에서 지원되는 언어로 문서를 검색하려면 인덱스를 선택한 다음 탐색 메뉴에서 인덱스 검색 옵션을 선택합니다. 검색 설정을 선택한 다음 언어 드롭다운에서 언어를 선택하여 문서를 반환할 언어를 선택합니다.

다음 예에서는 스페인어로 문서를 검색하는 방법을 보여줍니다.

콘솔에서 스페인어로 인덱스를 검색하려면

1. AWS Management Console 로그인하고 http://console.aws.amazon.com/kendra/ 에서 Amazon Kendra 콘솔을 엽니다.
2. 탐색 메뉴에서 인덱스를 선택하고 해당 인덱스를 선택합니다.
3. 탐색 메뉴에서 인덱스 검색 옵션을 선택합니다.

4. 검색 설정에서 언어 드롭다운을 선택하고 스페인어를 선택합니다.
5. 텍스트 상자에 쿼리를 입력한 후 Enter 키를 누릅니다.
6. Amazon Kendra 검색 결과를 스페인어로 반환합니다.

CLI, Python 또는 Java를 사용하여 스페인어로 인덱스를 검색하려면

- 다음 예제에서는 스페인어로 인덱스를 검색합니다. `searchString` 값을 검색 쿼리로 변경하고 `indexID` 값을 검색하려는 인덱스의 식별자로 변경합니다. 스페인어의 언어 코드는 `es`입니다. 이 코드를 사용자의 언어 코드로 바꿀 수 있습니다.

  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");
```

```
        }
    }
}
```

구절 검색

Retrieve API를 검색 증강 생성(RAG) 시스템의 검색기로 사용할 수 있습니다.

RAG 시스템은 생성형 인공 지능을 사용하여 질문에 답하는 애플리케이션을 구축합니다. RAG 시스템은 검색기와 대규모 언어 모델(LLM)로 구성됩니다. 쿼리가 주어지면 검색기는 문서 모음에서 가장 관련성이 높은 텍스트 청크를 식별하고 이를 LLM에 공급하여 가장 유용한 답변을 제공합니다. 그런 다음 LLM은 관련 텍스트 청크 또는 구절을 분석하고 쿼리에 대한 포괄적인 응답을 생성합니다.

Retrieve API는 구절이라고 하는 텍스트 또는 발췌문을 살펴보고 쿼리와 가장 관련성이 높은 상위 구절을 반환합니다.

Query API와 마찬가지로 Retrieve API도 시맨틱 검색을 사용하여 관련 정보를 검색합니다. 시맨틱 검색은 검색 쿼리의 컨텍스트와 인덱싱된 문서에서 사용 가능한 모든 정보를 고려합니다. 하지만 기본적으로 Query API는 최대 100개 토큰 단어로 구성된 발췌문 구절만 반환합니다. Retrieve API를 사용하면 최대 200개의 토큰 단어로 된 긴 구절과 의미상 관련이 있는 최대 100개의 구절을 검색할 수 있습니다. 여기에는 인덱스의 질문-답변 또는 FAQ 유형 응답은 포함되지 않습니다. 구절은 여러 문서 및 동일한 문서의 여러 부분에서 의미론적으로 추출할 수 있는 텍스트 발췌문입니다. 극단적인 경우 Retrieve API를 사용하여 문서에서 구절이 전혀 생성되지 않는 경우 Query API와 해당 응답 유형을 대신 사용할 수 있습니다.

또한 Retrieve API를 사용하여 다음을 수행할 수 있습니다.

- 인덱스 수준에서 부스팅을 재정의
- 문서 필드 또는 속성을 기준으로 필터링
- 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 필터링
- 검색된 구절 결과에 대한 신뢰도 점수 버킷을 확인합니다. 신뢰도 버킷은 응답과 쿼리의 관련성을 Amazon Kendra 가 얼마나 확신하는지 나타내는 상대적 순위를 제공합니다.

> **Note**
>
> 신뢰도 점수 버킷은 현재 영어로만 제공됩니다.

답변에 유용한 추가 정보를 제공할 수 있는 특정 필드를 포함시킬 수도 있습니다.

Retrieve API는 현재 Query API에서 지원하는 모든 기능을 지원하지는 않습니다. 고급 쿼리 구문을 사용한 쿼리, 쿼리에 대한 맞춤법 수정 제안, 패싯, 검색 쿼리 자동 완성을 위한 쿼리 제안, 증분 학습 등의 기능은 지원되지 않습니다. 단, 모든 기능이 API에 적용되는 것은 아닙니다. Retrieve 향후 Retrieve API 릴리스는 이 가이드에 문서화될 예정입니다.

Retrieve API는 인덱스에 설정한 쿼리 용량 단위 수를 공유합니다. 단일 용량 단위에 포함되는 항목 및 인덱스의 기본 용량에 대한 자세한 내용은 용량 조정을 참조하세요.

> **Note**
>
> Amazon Kendra 개발자 에디션을 사용하는 경우 용량을 추가할 수 없으며 Amazon Kendra Enterprise Edition을 사용할 때만 용량을 추가할 수 있습니다. Developer 및 Enterprise Edition에 포함된 내용에 대한 자세한 내용은 Amazon Kendra 에디션을 참조하세요.

다음은 Retrieve API를 사용하여 "how does amazon kendra work?" 쿼리의 인덱스에 있는 문서에서 가장 관련성이 높은 상위 100개 구절을 검색하는 예제입니다.

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");
        }
    }
}
```

인덱스 찾아보기

검색 쿼리를 입력하지 않고도 속성이나 패싯별로 문서를 찾아볼 수 있습니다. Amazon Kendra 인덱스 찾아보기를 사용하면 사용자가 특정 쿼리를 염두에 두지 않고 인덱스를 자유롭게 탐색하여 문서를 검색할 수 있습니다. 또한 이를 통해 사용자는 검색의 시작점으로 인덱스를 광범위하게 탐색할 수 있습니다.

인덱스 찾아보기는 정렬 유형이 있는 문서 속성 또는 패싯별로 검색하는 경우에만 사용할 수 있습니다. 인덱스 찾아보기를 사용하여 전체 인덱스를 검색할 수는 없습니다. 쿼리 텍스트가 없는 경우 문서 속성 필터 또는 패싯과 정렬 유형을 Amazon Kendra 요청합니다.

쿼리 API를 사용하여 인덱스를 탐색할 수 있게 하려면 Facet 및 AttributeFilter를 포함해야 합니다. SortingConfiguration 콘솔에서 인덱스 탐색을 허용하려면 탐색 메뉴의 인덱스에서 인덱스를 선택한 다음 인덱스 검색 옵션을 선택합니다. 검색 상자에서 Enter 키를 두 번 누릅니다. 검색 결과 필터링 드롭다운을 선택하여 필터를 선택하고 정렬 드롭다운을 선택하여 정렬 유형을 선택합니다.

다음은 스페인어로 된 문서에 대한 인덱스를 문서 작성 날짜의 내림차순으로 찾아보는 예입니다.

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");
        }
    }
}
```

검색 결과 추천

사용자가 특정 쿼리를 실행하면 검색 결과에 특정 문서를 추천할 수 있습니다. 이렇게 하면 결과가 사용자에게 더 잘 보이고 눈에 띄도록 만들 수 있습니다. 추천 결과는 일반적인 결과 목록과 분리되어 검색 페이지 상단에 표시됩니다. 여러 쿼리에 대해 여러 다른 문서를 추천해 실험하거나 특정 문서가 제대로 표시되도록 할 수 있습니다.

특정 쿼리를 특정 문서에 매핑하여 결과에 추천할 수 있습니다. 쿼리에 정확히 일치하는 항목이 있는 경우 검색 결과에 하나 이상의 특정 문서가 추천됩니다.

예를 들어, 사용자가 '신제품 2023' 쿼리를 실행하는 경우 '새 소식' 및 '출시 예정'이라는 제목의 문서를 선택하여 검색 결과 페이지 상단에 추천하도록 지정할 수 있습니다. 이렇게 하면 신제품에 대한 이러한 문서가 마땅히 받아야 할 가시성을 확보할 수 있습니다.

Amazon Kendra 검색 결과 페이지 상단에 표시할 결과가 이미 선택된 경우 검색 결과가 중복되지 않습니다. 추천 검색 결과가 다른 모든 검색 결과보다 우선하여 추천되는 경우 첫 번째 결과로 순위가 다시 매겨지지 않습니다.

특정 결과를 추천하려면 쿼리에 포함된 키워드나 구문을 사용하여 쿼리를 일부만 일치시키는 것이 아니라 전체 텍스트 검색어와 정확히 일치하도록 지정해야 합니다. 예를 들어 추천 결과 집합에서 'Kendra' 쿼리만 지정하는 경우 'Kendra는 의미론적으로 결과 순위를 어떻게 매깁니까?'와 같은 쿼리는 추천 결과를 렌더링하지 않습니다. 추천 검색결과는 범위가 너무 넓은 검색어가 아니라 특정 검색어에 맞게 설계되었습니다. Amazon Kendra 키워드 유형 쿼리를 자연스럽게 처리하여 검색 결과에서 가장 유용한 문서의 순위를 매기므로 단순 키워드를 기반으로 결과가 과도하게 표시되지 않습니다.

사용자가 자주 사용하는 특정 쿼리가 있는 경우 추천 결과로 해당 쿼리를 지정할 수 있습니다. 예를 들어 Amazon Kendra 분석을 사용하여 상위 쿼리를 살펴보고 'Kendra는 의미론적으로 결과 순위를

어떻게 매기나요?'와 같은 특정 쿼리를 찾는 경우를 예로 들 수 있습니다. 그리고 '켄드라 시맨틱 검색'이 자주 사용되므로 이러한 쿼리는 'search 101'이라는 제목의 문서를 표시하는 데 유용할 수 있습니다.Amazon Kendra

Amazon Kendra 추천 검색결과에 대한 쿼리는 대소문자를 구분하지 않는 것으로 취급합니다. Amazon Kendra 쿼리를 소문자로 변환하고 후행 공백 문자를 단일 공백으로 바꿉니다. Amazon Kendra 추천 결과에 대한 쿼리를 지정할 때와 마찬가지로 다른 모든 문자와 일치합니다.

CreateFeaturedResultsSetAPI를 사용하여 특정 쿼리에 매핑할 추천 결과 세트를 생성합니다. 콘솔을 사용하는 경우 인덱스를 선택한 다음 탐색 메뉴에서 추천 결과를 선택하여 추천 결과 세트를 생성합니다. 인덱스당 추천 결과 세트 최대 50개, 세트당 추천할 문서 최대 4개, 추천 결과 세트당 최대 49개의 쿼리 텍스트를 생성할 수 있습니다. 지원 팀에 문의하여 이러한 제한 증가를 요청할 수 있습니다.

여러 추천 결과 세트에서 동일한 문서를 선택할 수 있습니다. 하지만 여러 세트에 동일한 일치 쿼리 텍스트를 사용해서는 안 됩니다. 추천 결과에 지정하는 쿼리는 각 인덱스의 추천 결과 세트별로 고유해야 합니다.

추천 문서를 최대 4개까지 선택할 때 문서 순서를 정렬할 수 있습니다. API를 사용하는 경우 추천 문서를 나열하는 순서는 추천 결과에 표시된 순서와 동일합니다. 콘솔을 사용하는 경우 결과에 추천할 문서를 선택할 때 문서 순서를 간단히 끌어서 놓을 수 있습니다.

추천 결과를 구성할 때는 특정 사용자와 그룹이 특정 문서에 액세스할 수 있고 다른 사용자와 그룹은 액세스할 수 없는 액세스 제어가 여전히 유효합니다. 이는 사용자 컨텍스트 필터링의 경우에도 마찬가지입니다. 예를 들어 사용자 A는 회사 기밀 문서에 액세스해서는 안 되는 '인턴' 회사 그룹에 속해 있습니다. 사용자 A가 회사 기밀 문서를 추천하는 쿼리를 입력하면 사용자 A의 검색 결과에 이 문서가 추천되지 않습니다. 이는 검색 결과 페이지의 다른 모든 결과에서도 마찬가지입니다. 태그를 사용하면 액세스를 제어하는 Amazon Kendra 리소스인 추천 결과 세트에 대한 액세스를 제어할 수도 있습니다.

다음은 “신제품 2023", “신제품 출시” 쿼리를 “새 소식”(doc-id-1) 및 “출시 예정”(doc-id-2) 이라는 제목의 문서에 매핑하여 추천 결과 세트를 만드는 예제입니다.

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.")
```

HTML에 대한 표 형식 검색

Amazon Kendra의 표 검색 기능은 HTML 문서에 포함된 표에서 답변을 검색하고 추출할 수 있습니다. 색인을 검색할 때 쿼리와 관련이 있는 경우 표에서 발췌한 내용을 Amazon Kendra 포함시키고 유용한 정보를 제공하세요.

Amazon Kendra 표의 유용한 정보를 포함하여 문서 본문 텍스트 내의 모든 정보를 살펴봅니다. 예를 들어, 인덱스에는 운영 비용, 수입 및 기타 재무 정보에 대한 표가 포함된 비즈니스 보고서가 포함됩니다. 질문의 경우 “2020-2022년의 연간 운영 비용은 얼마입니까?” 는 관련 테이블 열 “운영 (백만 달러)" 및 “회계 연도”와 2020년, 2021년, 2022년의 소득 값이 포함된 테이블 행을 포함하는 테이블에서 발췌한 내용을 반환할 Amazon Kendra 수 있습니다. 이 표 발췌문은 문서 제목, 전체 문서 링크 및 포함하려는 기타 문서 필드와 함께 결과에 포함됩니다.

정보가 표의 한 셀에 있든 여러 셀에 있든 상관없이 검색 결과에 표 발췌문을 표시할 수 있습니다. 예를 들어, Amazon Kendra 다음과 같은 각 쿼리에 맞게 조정된 표 발췌문을 표시할 수 있습니다.

- '2020년 최고 이자율 신용카드'
- “2020-2022년 최고 이자율 신용카드”
- “2020-2022년 최고 이자율 신용카드 3개”
- “이자율이 10% 미만인 신용카드”
- “사용 가능한 모든 저금리 신용카드”

Amazon Kendra 쿼리와 가장 관련이 있는 테이블 셀 또는 셀을 강조 표시합니다. 가장 관련성이 높은 셀과 해당 행, 열, 열 이름이 검색 결과에 표시됩니다. 표 발췌문에는 쿼리와 관련된 표의 셀 수와 원본 표에서 사용할 수 있는 열 수에 따라 최대 5개의 열과 3개의 행이 표시됩니다. 표 발췌문에는 가장 관련성이 높은 상위 셀이 다음으로 가장 관련성이 높은 셀과 함께 표시됩니다.

응답에는 표 답변이 쿼리와 얼마나 관련이 있는지를 보여주는 신뢰도 버킷(MEDIUM, HIGH, VERY_HIGH)이 포함됩니다. 표의 셀 값이 신뢰도 VERY_HIGH인 경우, 해당 값이 '상위 답변'이 되어 강조 표시됩니다. 신뢰도가 HIGH인 표 셀 값의 경우 해당 값이 강조 표시됩니다. 신뢰도가 MEDIUM인 표 셀 값의 경우 해당 값이 강조 표시되지 않습니다. 표 답변에 대한 전체 신뢰도가 응답에 반환됩니다. 예를 들어, 표에 대부분 HIGH 신뢰도인 표 셀이 포함된 경우, 표 답변에 대한 응답에 반환된 전체 신뢰도는 HIGH 신뢰도입니다.

기본적으로 표에는 문서의 다른 구성 요소보다 더 높은 중요도나 가중치가 부여되지 않습니다. 문서 내에서 표가 쿼리와 관련성이 약간만 있지만 관련성이 높은 단락이 있는 경우 해당 단락의 일부를 Amazon Kendra 반환합니다. 검색 결과에는 동일한 문서나 다른 문서에서 가능한 최선의 답변과 가장 유용한 정보를 제공하는 콘텐츠가 표시됩니다. 표에 대한 신뢰도가 신뢰도 `MEDIUM` 밑으로 떨어지면 표 발췌문은 응답에 반환되지 않습니다.

기존 인덱스에서 표 형식 검색을 사용하려면 콘텐츠를 다시 인덱싱해야 합니다.

Amazon Kendra 표 형식 검색은 동의어 (사용자 정의 동의어 포함) 를 지원합니다. Amazon Kendra table 태그 내에 HTML 테이블이 있는 영어 문서만 지원합니다.

다음 예제는 쿼리 결과에 포함된 표 발췌문을 보여줍니다. 표 발췌문을 비롯한 쿼리 응답이 포함된 샘플 JSON을 보려면 쿼리 응답 및 유형을 참조하세요.

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");
        }
    }
}
```

쿼리 제안

Amazon Kendra 쿼리 제안은 사용자가 검색 쿼리를 더 빠르게 입력하고 검색 결과를 안내하는 데 도움이 될 수 있습니다.

Amazon Kendra 다음 중 하나를 기반으로 사용자와 관련된 쿼리를 제안합니다.

- 쿼리 기록 또는 쿼리 로그의 인기 쿼리
- 문서 필드/속성의 내용

SuggestionTypes를 QUERY 또는 DOCUMENT_ATTRIBUTES로 설정하고 GetQuerySuggestions를 호출하여 쿼리 기록 또는 문서 필드 사용에 대한 기본 설정을 지정할 수 있습니다. 기본적으로 는 쿼리 기록을 Amazon Kendra 사용하여 제안을 기반으로 합니다. 전화를 걸 때 쿼리 기록과 문서 필드가 모두 활성화되어 UpdateQuerySuggestionsConfig있고 문서 필드를 사용하도록 SuggestionTypes 환경설정을 지정하지 않은 경우 쿼리 기록이 Amazon Kendra 사용됩니다.

콘솔을 사용하는 경우 쿼리 기록 또는 문서 필드를 기반으로 쿼리를 제안할 수 있습니다. 먼저 인덱스를 선택한 다음 탐색 메뉴의 보강에서 쿼리 제안을 선택합니다. 그런 다음 쿼리 제안 구성을 선택합니다. 쿼리 제안을 구성하면 오른쪽 패널에서 쿼리 기록 또는 문서 필드를 선택하고 검색 창에 검색 쿼리를 입력할 수 있는 검색 콘솔로 이동합니다.

기본적으로 쿼리 기록과 문서 필드를 사용하는 쿼리 제안은 모두 추가 비용 없이 활성화됩니다. UpdateQuerySuggestionsConfig API를 사용하여 언제든지 이러한 유형의 쿼리 제안을 비활성화할 수 있습니다. 쿼리 기록을 기반으로 쿼리 제안을 비활성화하려면 UpdateQuerySuggestionsConfig를 호출할 때 Mode를 DISABLED로 설정합니다. 문서 필드를 기반으로 하는 쿼리 제안을 비활성화하려면 문서 필드 구성에서 AttributeSuggestionsMode를 INACTIVE로 설정한 다음 UpdateQuerySuggestionsConfig>를 호출하세요. 콘솔을 사용하는 경우 쿼리 제안 설정에서 쿼리 제안을 비활성화할 수 있습니다.

쿼리 제안은 대소문자를 구분하지 않습니다. Amazon Kendra 쿼리 접두사와 제안된 쿼리를 소문자로 변환하고, 작은따옴표와 큰따옴표를 모두 무시하고, 여러 개의 공백 문자를 단일 공백으로 바꿉니다. Amazon Kendra 다른 모든 특수 문자와 그대로 일치합니다. Amazon Kendra 사용자가 2자 미만 또는 60자 이상을 입력하면 추천 항목이 표시되지 않습니다.

주제

- 쿼리 기록을 사용한 쿼리 제안
- 문서 필드를 사용한 쿼리 제안
- 제안에서 특정 쿼리 또는 문서 필드 콘텐츠 차단

쿼리 기록을 사용한 쿼리 제안

주제

쿼리 기록 또는 쿼리 로그의 인기 검색어를 기반으로 사용자와 관련된 검색어를 제안하도록 선택할 수 있습니다. Amazon Kendra 사용자가 검색하고 이러한 쿼리를 통해 학습한 모든 쿼리를 사용하여 사용자에게 제안을 제공합니다. Amazon Kendra 사용자가 검색어를 입력하기 시작하면 인기 검색어를 추천해 줍니다. Amazon Kendra 쿼리의 접두사 또는 처음 몇 글자가 사용자가 쿼리로 입력하기 시작한 것과 일치하면 쿼리를 제안합니다.

예를 들어, 사용자가 '예정된 이벤트'라는 쿼리를 입력하기 시작합니다. Amazon Kendra 는 쿼리 기록을 통해 많은 사용자가 '예정된 이벤트 2050'을 여러 번 검색했다는 사실을 알게 되었습니다. 사용자는 검색창 바로 아래에 '예정된 이벤트 2050'이 표시되어 검색 쿼리가 자동으로 완성되는 것을 볼 수 있습니다. 사용자가 이 쿼리 제안을 선택하면 검색 결과에 '새 이벤트: 2050년에 무슨 일이 벌어지고 있는가' 문서가 반환됩니다.

적합한 쿼리를 Amazon Kendra 선택하여 사용자에게 제안하는 방법을 지정할 수 있습니다. 예를 들어 검색어 제안은 최소 10명의 순 사용자 (기본값 3명) 가 검색했고, 최근 30일 이내에 검색되었으며, 차단 목록에 있는 단어나 문구를 포함하지 않도록 지정할 수 있습니다. Amazon Kendra 쿼리에는 검색 결과가 하나 이상 있고 4자 이상의 단어가 하나 이상 포함되어 있어야 합니다.

제안할 쿼리를 선택하기 위한 설정

UpdateQuerySuggestionsConfig API를 사용하여 제안할 쿼리를 선택하기 위해 다음 설정을 구성할 수 있습니다.

- 모드 - 쿼리 기록을 사용하는 쿼리 제안은 ENABLED 또는 LEARN_ONLY 입니다. Amazon Kendra 는 기본적으로 쿼리 제안을 활성화합니다. LEARN_ONLY는 쿼리 제안을 끕니다. 끄면 제안을 Amazon Kendra 계속 학습하지만 사용자에게 검색어를 제안하지는 않습니다.
- 쿼리 로그 기간 - 쿼리 로그 기간의 쿼리가 얼마나 최신인가. 이 기간은 오늘부터 지난 날까지의 일수를 나타내는 정수 값입니다.
- 사용자 정보가 없는 쿼리 - 모든 쿼리를 포함하도록 TRUE로 설정하거나 사용자 정보가 포함된 쿼리만 포함하도록 FALSE로 설정합니다. 사용자가 쿼리를 실행할 때 검색 애플리케이션에 사용자 ID와 같은 사용자 정보가 포함된 경우 이 설정을 사용할 수 있습니다. 기본적으로 이 설정은 쿼리와 관련

된 특정 사용자 정보가 없는 경우 쿼리를 필터링하지 않습니다. 하지만 이 설정을 사용하면 사용자 정보가 포함된 쿼리를 기반으로 제안만 할 수 있습니다.

- 순 사용자 - 쿼리를 사용자에게 제안하기 위해 해당 쿼리를 검색한 최소 순 사용자 수입니다. 이 숫자는 정수 값입니다.
- 쿼리 수 - 쿼리를 사용자에게 제안하기 위해 해당 쿼리가 검색되어야 하는 최소 횟수입니다. 이 숫자는 정수 값입니다.

이러한 설정은 쿼리를 인기 쿼리로 선택하여 사용자에게 제안하는 방식에 영향을 줍니다. 설정을 조정하는 방법은 특정 요구 사항에 따라 달라집니다. 예를 들면 다음과 같습니다.

- 사용자가 보통 한 달에 평균 한 번 검색하는 경우 쿼리 로그 기간의 일수를 30일로 설정할 수 있습니다. 이 설정을 사용하면 사용자의 최근 쿼리 대부분을 기간이 만료되기 전에 캡처할 수 있습니다.
- 사용자 정보가 포함된 쿼리 수가 적고 작은 샘플 크기를 기준으로 쿼리를 제안하지 않으려면 모든 사용자를 포함하도록 쿼리를 설정할 수 있습니다.
- 인기 있는 쿼리를 10명 이상의 순 사용자가 검색하고 100회 이상 검색한다고 정의하면 순 사용자를 10으로 설정하고 쿼리 수를 100으로 설정합니다.

> **Warning**
>
> 설정 변경 내용이 즉시 적용되지 않을 수 있습니다. DescribeQuerySuggestionsConfig API를 사용하여 설정 변경을 추적할 수 있습니다. 업데이트된 설정이 적용되는 데 걸리는 시간은 업데이트한 내용과 인덱스의 검색 쿼리 수에 따라 달라집니다. Amazon Kendra 는 설정을 변경하거나 차단 목록을 적용한 후 24시간마다 제안을 자동으로 업데이트합니다.

CLI

쿼리 제안을 검색하려면

```
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
```

쿼리 제안을 업데이트하려면

예를 들어 쿼리 로그 기간과 쿼리를 검색해야 하는 최소 횟수를 변경하려면:

```
aws kendra update-query-suggestions-config \
 --index-id index-id \
 --query-log-look-back-window-in-days 30 \
 --minimum-query-count 100
```

Python

쿼리 제안을 검색하려면

```
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.")
```

쿼리 제안을 업데이트하려면

예를 들어 쿼리 로그 기간과 쿼리를 검색해야 하는 최소 횟수를 변경하려면:

```
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.")
```

쿼리 기록을 보존하면서 제안 지우기

ClearQuerySuggestions API를 사용하여 쿼리 제안을 지울 수 있습니다. 제안을 지우면 기존 쿼리 제안만 삭제되고 쿼리 기록의 쿼리는 삭제되지 않습니다. 제안을 지우면 제안을 지운 시점부터 쿼리 로그에 추가된 새 쿼리를 기반으로 새 제안을 Amazon Kendra 학습합니다.

CLI

쿼리 제안을 지우려면

```
aws  kendra clear-query-suggestions \
 --index-id index-id
```

Python

쿼리 제안을 지우려면

```
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.")
```

제안 없음

쿼리에 대한 제안이 표시되지 않으면 다음 이유 중 하나일 수 있습니다.

- 색인에 Amazon Kendra 학습할 수 있는 쿼리가 충분하지 않습니다.
- 쿼리 제안 설정이 너무 엄격하여 대부분의 쿼리가 제안에서 필터링됩니다.
- 최근에 제안을 지웠지만 새 제안을 익히려면 새 쿼리가 누적되기까지 Amazon Kendra 아직 시간이 필요합니다.

DescribeQuerySuggestionsConfig API를 사용하여 현재 설정을 확인할 수 있습니다.

문서 필드를 사용한 쿼리 제안

주제

- 제안할 필드를 선택하기 위한 설정
- 문서 필드의 사용자 제어

문서 필드의 콘텐츠를 기반으로 사용자와 관련된 쿼리를 제안하도록 선택할 수 있습니다. 쿼리 기록을 사용하여 다른 인기 있는 관련 쿼리를 제안하는 대신, 쿼리를 자동 완성하는 데 유용한 문서 필드에 포함된 정보를 사용할 수 있습니다. Amazon Kendra 사용자 쿼리로 설정되고 사용자 `Suggestable` 쿼리와 밀접하게 일치하는 필드에서 관련 콘텐츠를 찾습니다. 그런 다음 사용자가 검색어를 입력하기 시작할 때 이 콘텐츠를 Amazon Kendra 추천합니다.

예를 들어 제안의 기반으로 제목 필드를 지정하고 사용자가 'How amazon ken... '이라는 검색어를 입력하기 시작하는 경우 검색을 자동 완료하도록 가장 관련성이 높은 제목인 Amazon Kendra 'How works'를 제안할 수 있습니다. 사용자는 검색창 바로 아래에 Amazon Kendra '작업 방법'이 표시되어 검색 쿼리가 자동으로 완성되는 것을 볼 수 있습니다. 사용자가 이 검색어 제안을 선택하면 검색 결과에 Amazon Kendra '작업 방법' 문서가 반환됩니다.

쿼리 제안을 위한 필드 구성의 일부로 필드를 `Suggestable`로 설정하면 `String` 및 `StringList` 유형의 모든 문서 필드의 콘텐츠를 사용하여 쿼리를 제안할 수 있습니다. 차단 목록을 사용하여 특정 단어나 문구가 포함된 제안된 문서 필드가 사용자에게 표시되지 않도록 할 수도 있습니다. 하나의 차단 목록을 사용할 수 있습니다. 차단 목록은 쿼리 제안을 어떻게 설정하든(쿼리 기록 또는 문서 필드 사용) 상관없이 적용됩니다.

제안할 필드를 선택하기 위한 설정

AttributeSuggestionsConfig를 사용하고 UpdateQuerySuggestionsConfig API를 호출하여 인덱스 수준에서 설정을 업데이트하도록 다음과 같이 제안할 문서 필드 선택 설정을 구성할 수 있습니다.

- 필드/속성 제안 모드 - 문서 필드를 사용하는 쿼리 제안은 `ACTIVE` 또는 `INACTIVE`입니다. Amazon Kendra 는 기본적으로 쿼리 제안을 활성화합니다.
- 제안 가능한 필드/속성 - 제안의 기반이 되는 필드 이름 또는 필드 키입니다. 필드 구성의 일부로 이러한 필드는 `Suggestable`에 대해 `TRUE`로 설정해야 합니다. 인덱스 수준에서 구성을 유지하면서 쿼리 수준에서 필드 구성을 재정의할 수 있습니다. GetQuerySuggestionsAPI를 사용하여 쿼리 `AttributeSuggestionConfig` 수준에서 변경하세요. 쿼리 수준의 이 구성은 인덱스 수준에서 구성을 업데이트할 필요 없이 다양한 문서 필드를 사용하여 빠르게 실험해 보는 데 유용할 수 있습니다.
- 추가 필드/속성 - 쿼리 제안에 대한 응답에 포함하려는 추가 필드입니다. 이러한 필드는 응답에 추가 정보를 제공하는 데 사용되지만 제안의 근거로는 사용되지 않습니다.

> **Warning**
>
> 설정 변경 내용이 즉시 적용되지 않을 수 있습니다. DescribeQuerySuggestionsConfig API를 사용하여 설정 변경을 추적할 수 있습니다. 업데이트된 설정이 적용되는 데 걸리는 시간은 업데이트에 따라 다릅니다. Amazon Kendra 설정을 변경한 후 또는 차단 목록을 적용한 후 24시간마다 제안을 자동으로 업데이트합니다.

CLI

인덱스 수준에서 구성을 변경할 필요 없이 쿼리 제안을 검색하고 쿼리 수준에서 문서 필드 구성을 재정의할 수 있습니다.

```
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
```

쿼리 제안을 업데이트하려면

예를 들어 인덱스 수준에서 문서 필드 구성을 변경하려면:

```
aws kendra update-query-suggestions-config \
 --index-id index-id \
 --attribute-suggestions-config '{"SuggestableConfigList": '[{"SuggestableConfig":
 "_document_title", "Suggestable": true}]', "AttributeSuggestionsMode": "ACTIVE"}'
```

Python

인덱스 수준에서 구성을 변경할 필요 없이 쿼리 제안을 검색하고 쿼리 수준에서 문서 필드 구성을 재정의할 수 있습니다.

```
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.")
```

쿼리 제안을 업데이트하려면

예를 들어 인덱스 수준에서 문서 필드 구성을 변경하려면:

```
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.")
```

문서 필드의 사용자 제어

쿼리 제안의 기반으로 사용할 문서 필드에 사용자 컨텍스트 필터링을 적용할 수 있습니다. 이는 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 문서 필드 정보를 필터링합니다. 인턴이 회사 포털을 검색하지만 회사 일급 기밀 문서에 액세스할 수 없는 경우를 예로 들 수 있습니다. 따라서 일급 기밀 문서의 제목이나 기타 제안 가능한 필드를 기반으로 제안된 쿼리는 인턴에게 표시되지 않습니다.

어떤 사용자와 그룹에 어떤 문서에 대한 액세스 권한이 할당되는지 정의하는 액세스 제어 목록(ACL)을 사용하여 문서를 인덱싱할 수 있습니다. 그런 다음 쿼리 제안을 위해 문서 필드에 사용자 컨텍스트 필터링을 적용할 수 있습니다. 현재 인덱스에 설정된 사용자 컨텍스트 필터링은 쿼리 제안에 대한 문서 필드 구성에 적용되는 사용자 컨텍스트 필터링과 동일합니다. 사용자 컨텍스트 필터링은 문서 필드 구성의 일부입니다. AttributeSuggestionsGetConfig를 사용하고 GetQuerySuggestions를 호출합니다.

제안에서 특정 쿼리 또는 문서 필드 콘텐츠 차단

차단 목록은 사용자에게 특정 검색어를 제안하는 Amazon Kendra 것을 중지합니다. 차단 목록은 검색어 제안에서 제외하려는 단어나 구문의 목록입니다. Amazon Kendra 차단 목록에 있는 단어 또는 구문과 정확히 일치하는 단어를 포함하는 쿼리를 제외합니다.

차단 목록을 사용하면 쿼리 기록이나 문서 필드에 흔히 나타나며 Amazon Kendra 가 제안으로 선택할 수 있는 불쾌한 단어나 문구를 차단할 수 있습니다. 차단 목록을 사용하면 공개적으로 공개 또는 발표할 준비가 되지 않은 정보가 포함된 쿼리를 제안하지 Amazon Kendra 못하게 할 수도 있습니다. 예를 들어, 사용자들이 잠재적 신제품의 향후 출시에 대해 자주 문의합니다. 하지만 아직 출시할 준비가 되지 않았으므로 제품을 제안하고 싶지 않을 것입니다. 제품 이름과 제품 정보가 포함된 쿼리는 제안 항목에서 차단할 수 있습니다.

CreateQuerySuggestionsBlockList API를 사용하여 쿼리에 대한 차단 목록을 만들 수 있습니다. 각각의 차단 단어나 문구를 텍스트 파일의 별도 줄에 넣습니다. 그런 다음 Amazon S3 버킷에 텍스트 파일을 업로드하고 파일의 경로나 위치를 입력합니다 Amazon S3. Amazon Kendra 현재는 하나의 차단 목록만 생성할 수 있습니다.

Amazon S3 버킷에서 차단된 단어와 구문의 텍스트 파일을 바꿀 수 있습니다. 에서 Amazon Kendra차단 목록을 업데이트하려면 UpdateQuerySuggestionsBlockListAPI를 사용하세요.

DescribeQuerySuggestionsBlockList API를 사용하여 차단 목록의 상태를 확인할 수 있습니다. `DescribeQuerySuggestionsBlockList`는 다음과 같은 기타 유용한 정보도 제공할 수 있습니다.

- 차단 목록이 마지막으로 업데이트된 시기
- 현재 차단 목록에 있는 단어 또는 문구의 수
- 차단 목록을 만들 때 유용한 오류 메시지

ListQuerySuggestionsBlockLists API를 사용하여 인덱스의 차단 목록 요약 목록을 가져올 수도 있습니다.

차단 목록을 삭제하려면 DeleteQuerySuggestionsBlockListAPI를 사용하세요.

차단 목록에 대한 업데이트는 즉시 적용되지 않을 수 있습니다.
`DescribeQuerySuggestionsBlockList` API를 사용하여 업데이트를 추적할 수 있습니다.

CLI

차단 목록을 만들려면

```
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
```

차단 목록을 업데이트하려면

```
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
```

차단 목록을 삭제하려면

```
aws kendra delete-query-suggestions-block-list \
 --index-id index-id \
 --id block-list-id
```

Python

차단 목록을 만들려면

```
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.")
```

차단 목록을 업데이트하려면

```
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.")
```

차단 목록을 삭제하려면

```
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.")
```

쿼리 맞춤법 검사기

Amazon Kendra 맞춤법 검사기는 쿼리에 대한 맞춤법 수정을 제안합니다. 이렇게 하면 검색 결과가 제로인 경우를 최소화하고 적절한 결과를 반환할 수 있습니다. 맞춤법이 틀린 쿼리를 사용하면 일치하는 결과가 없거나 반환된 문서가 없는 제로 검색 결과를 사용자가 받을 수 있습니다. 또는 맞춤법이 틀린 쿼리로 인해 사용자에게 관련 없는 검색 결과가 표시될 수 있습니다.

맞춤법 검사기는 인덱싱된 문서에 나타나는 단어, 수정된 단어가 맞춤법 틀린 단어와 얼마나 일치하는지에 따라 맞춤법이 틀린 단어의 수정을 제안하도록 설계되었습니다. 예를 들어 인덱싱된 문서에 'statements'라는 단어가 나타나면 'year-end financial statments'라는 쿼리에서 철자가 틀린 단어 'statments'와 거의 일치합니다.

맞춤법 검사기는 원래 쿼리 텍스트에서 맞춤법이 틀린 단어를 대체하는 의도된 단어 또는 수정된 단어를 반환합니다. 예를 들어 'depoying kendre search'는 'deploying Kendra search'를 반환할 수 있습니다. 또한 API에 제공된 오프셋 위치를 사용하여 프런트 엔드 애플리케이션의 쿼리에서 수정하여 반환된 단어를 강조 표시하거나 기울임꼴로 표시할 수 있습니다. 콘솔에서는 수정된 단어가 기본적으로 강조 표시되거나 기울임꼴로 표시됩니다. 예: 'deploying Kendra search'.

인덱싱된 문서에 나타나는 비즈니스 관련 용어나 특수 용어의 경우 맞춤법 검사기는 이러한 용어를 쿼리의 철자 오류로 오해하지 않습니다. 예를 들어, 'amazon macie'는 'amazon mace'로 정정되지 않습니다.

'year-end'와 같이 하이픈이 있는 단어의 경우 맞춤법 검사기는 이러한 단어를 개별 단어로 취급하여 해당 단어의 수정을 제안합니다. 예를 들어 'yaer-end'에 대해 제안된 수정 사항은 'year-end'일 수 있습니다.

`DOCUMENT` 및 `QUESTION_ANSWER` 쿼리 응답 유형의 경우 맞춤법 검사기는 문서 본문의 단어를 기반으로 철자가 틀린 단어를 수정하도록 제안합니다. 문서 본문은 철자가 틀린 단어와 거의 일치하는 수정을 제안하는 데 있어 제목보다 더 신뢰할 수 있습니다. `ANSWER` 쿼리 응답 유형의 경우 맞춤법 검사기는 인덱스의 기본 질문 및 답변 문서에 있는 단어를 기반으로 수정을 제안합니다.

개체를 사용하여 맞춤법 검사기를 활성화할 수 있습니다. SpellCorrectionConfiguration `IncludeQuerySpellCheckSuggestions`를 `TRUE`로 설정했습니다. 맞춤법 검사기는 콘솔에서 기본적으로 활성화됩니다. 기본적으로 콘솔에 내장되어 있습니다.

맞춤법 검사기는 영어뿐만 아니라 여러 언어로 쿼리에 대한 맞춤법 수정을 제안할 수도 있습니다. 맞춤법 검사기에 지원되는 언어 목록은 Amazon Kendra 에서 지원하는 언어를 참조하세요.

기본 제한이 적용된 쿼리 맞춤법 검사기 사용

맞춤법 검사기는 특정 기본값 또는 제한을 적용하여 설계되었습니다. 다음은 맞춤법 수정 제안을 활성화할 때 적용되는 현재 제한 목록입니다.

- 3자 미만 또는 길이가 30자를 초과하는 단어에 대해서는 제안된 맞춤법 교정이 반환될 수 없습니다. 30자 초과 또는 3자 미만을 허용하려면 지원 팀에 문의하세요.
- 맞춤법 수정 제안은 사용자 액세스 제어 또는 사용자 컨텍스트 필터링에 대한 액세스 제어 목록을 기반으로 제안을 제한할 수 없습니다. 맞춤법 수정은 특정 사용자만 사용할 수 있는 단어인지 여부에 관계없이 인덱싱된 문서에 있는 모든 단어를 기반으로 합니다. 쿼리에 대해 제안된 맞춤법 수정에 특정 단어가 나타나지 않도록 하려면 `SpellCorrectionConfiguration`를 활성화하지 마세요.
- 숫자가 포함된 단어에 대해서는 제안된 맞춤법 교정을 반환할 수 없습니다. 예: 'how 2 not br8k ubun2'.
- 맞춤법 수정 제안에는 인덱싱된 문서에 없는 단어를 사용할 수 없습니다.
- 맞춤법 수정 제안은 인덱싱된 문서에서 사용 빈도가 0.01퍼센트 미만인 단어를 사용할 수 없습니다. 0.01% 임계값을 변경하려면 지원 팀에 문의하세요.

검색 필터링 및 패싯

필터를 사용하여 Query API의 검색 결과 또는 응답을 개선할 수 있습니다. 필터는 응답의 문서를 쿼리에 직접 적용되는 문서로 제한합니다. 패싯된 검색 제안을 만들려면 응답에서 특정 문서 속성이나 특정 기준과 일치하지 않는 문서를 필터링하는 부울 논리를 사용하세요. `Query` API의 `Facets` 파라미터를 사용하여 패싯을 지정할 수 있습니다.

인덱싱할 때 사용한 문서를 검색하려면 Amazon Kendra AMAZON을 Amazon Lex사용하십시오. `KendraSearchIntent`. 를 Amazon Kendra 사용하여 Amazon Lex구성하는 예제는 Amazon Kendra 인덱스용 FAQ 봇 만들기를 참조하십시오. 를 사용하여 AttributeFilter응답에 대한 필터를 제공할 수도 있습니다. `AMAZON.KendraSearchIntent` 구성 시 JSON의 쿼리 필터입니다. 콘솔에서 검색 의도를 구성할 때 속성 필터를 제공하려면 의도 편집기로 이동하여 Amazon Kendra 쿼리를 선택하여 JSON에 쿼리 필터를 제공하세요. `AMAZON.KendraSearchIntent`에 대한 자세한 내용은 Amazon Lex 설명서를 참조하세요.

패싯

패싯은 검색 결과 집합의 범위가 지정된 뷰입니다. 예를 들어, 전 세계 도시에 대한 검색 결과를 제공할 수 있습니다. 그러면 관련 도시별로 문서가 필터링됩니다. 또는 특정 작성자의 결과를 표시하는 패싯을 만들 수 있습니다.

문서와 관련된 문서 속성 또는 메타데이터 필드를 패싯으로 사용하여 사용자가 해당 패싯 내의 범주 또는 값을 기준으로 검색하도록 할 수 있습니다. 사용자가 범주나 필드뿐만 아니라 하위 범주나 하위 필드로도 검색할 수 있도록 검색 결과에 중첩된 패싯을 표시할 수도 있습니다.

다음 예에서는 “City” 사용자 지정 속성에 대한 패싯 정보를 가져오는 방법을 보여줍니다.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        Facets = [
            {
                "DocumentAttributeKey" : "City"
            }
        ]
        )
```

중첩된 패싯을 사용하여 검색 범위를 더 좁힐 수 있습니다. 예를 들어, 문서 속성 또는 패싯 “City”에는 “Seattle”이라는 값이 포함됩니다. 또한 문서 속성 또는 패싯 CityRegion ""에는 “시애틀”에 지정된 문서에 대한 “North” 및 “South” 값이 포함됩니다. 문서를 도시뿐만 아니라 도시 내 지역으로도 검색할 수 있도록 검색 결과에 중첩된 패싯을 개수와 함께 표시할 수 있습니다.

중첩된 패싯은 쿼리 지연 시간에 영향을 줄 수 있다는 점에 유의하세요. 일반적으로 사용하는 중첩된 패싯 수가 많을수록 지연 시간에 미치는 잠재적 영향도 커집니다. 지연 시간에 영향을 미치는 다른 요인으로는 인덱싱된 문서의 평균 크기, 인덱스 크기, 매우 복잡한 쿼리, Amazon Kendra 인덱스의 전체 부하 등이 있습니다.

다음 예에서는 "CityRegion" 사용자 정의 속성에 대한 면 정보를 “City” 내에 중첩된 패싯으로 가져오는 방법을 보여 줍니다.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        Facets = [
            {
                "DocumentAttributeKey" : "City",
```

```
                "Facets": [
                    {
                        "DocumentAttributeKey" : "CityRegion"
                    }
                ]
            }
        ]
        )
```

문서 수와 같은 패싯 정보가 FacetResults 응답 배열에 반환됩니다. 이 콘텐츠를 사용하여 애플리케이션에 패싯된 검색 제안을 표시할 수 있습니다. 예를 들어, 문서 속성 “City”에 검색을 적용할 수 있는 도시가 포함되어 있는 경우 해당 정보를 사용하여 도시 검색 목록을 표시할 수 있습니다. 사용자는 도시를 선택하여 검색 결과를 필터링할 수 있습니다. 패싯된 검색을 수행하려면 Query API를 호출하고 선택한 문서 속성을 사용하여 결과를 필터링하세요.

쿼리에는 패싯당 최대 10개의 패싯 값을 표시할 수 있으며, 하나의 패싯 내에는 하나의 중첩된 패싯만 있을 수 있습니다. 이러한 한도를 증가시키려는 경우 지원 팀에 문의하세요. 패싯당 패싯 값 수를 10 미만으로 제한하려는 경우 Facet 객체에서 이를 지정할 수 있습니다.

다음 샘플 JSON 응답은 “City” 문서 속성으로 범위가 지정된 패싯을 보여줍니다. 이 응답에는 패싯 값에 대한 문서 수가 포함됩니다.

```
{
    'FacetResults': [
        {
            'DocumentAttributeKey': 'City',
            'DocumentAttributeValueCountPairs': [
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Dubai'
                    }
                },
                {
                    'Count': 3,
                    'DocumentAttributeValue': {
                        'StringValue': 'Seattle'
                    }
                },
                {
                    'Count': 1,
                    'DocumentAttributeValue': {
```

```
                        'StringValue': 'Paris'
                    }
                }
            ]
        }
    ]
```

또한 도시 내 지역과 같은 중첩된 패싯에 대한 패싯 정보를 표시하여 검색 결과를 추가로 필터링할 수 있습니다.

다음 샘플 JSON 응답은 CityRegion "" 문서 속성으로 범위가 지정된 패싯을 “City” 내의 중첩된 패싯으로 보여줍니다. 이 응답에는 중첩된 패싯 값에 대한 문서 수가 포함됩니다.

```
{
    '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'
                                    }
                                }
                            ]
                        }
                    ]
                }
```

```
        }
    ]
}
```

문자열 목록 필드를 사용하여 패싯을 생성할 경우 반환되는 패싯 결과는 문자열 목록의 내용을 기반으로 합니다. 예를 들어, 문자열 목록 필드에 “dachshund”, “sausage dog”을 나열한 항목과 “husky” 값이 있는 항목 등, 두 항목이 포함된 경우, 세 개의 패싯이 포함된 `FacetResults`가 표시됩니다.

자세한 정보는 쿼리 응답 및 응답 유형을 참조하세요.

문서 속성을 사용하여 검색 결과 필터링

기본적으로 `Query`는 모든 검색 결과를 반환합니다. 응답을 필터링하려면 문서 속성에 대해 논리적 연산을 수행할 수 있습니다. 예를 들어 특정 도시에 대한 문서만 원하는 경우 “City” 및 “State” 사용자 지정 문서 속성을 기준으로 필터링할 수 있습니다. 제공하는 필터에 대해 부울 연산을 만드는 AttributeFilter데 사용합니다.

대부분의 속성은 모든 응답 유형의 응답을 필터링하는 데 사용할 수 있습니다. 그러나 `_excerpt_page_number` 속성은 응답을 필터링할 때 `ANSWER` 응답 유형에만 적용됩니다.

다음 예제는 특정 도시, 시애틀 및 워싱턴 주를 필터링하여 논리적 AND 연산을 수행하는 방법을 보여줍니다.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {'AndAllFilters':
            [
                {"EqualsTo": {"Key": "City","Value": {"StringValue": "Seattle"}}},
                {"EqualsTo": {"Key": "State","Value": {"StringValue": "Washington"}}}
            ]
            }
        )
```

다음 예제는 `Fileformat`, `Author` 또는 `SourceURI` 키 중 하나라도 지정된 값과 일치하는 경우에 대해 논리적 OR 연산을 수행하는 방법을 보여줍니다.

```
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"}}}
            ]
            }
        )
```

StringList 필드의 경우 ContainsAny 또는 ContainsAll 속성 필터를 사용하여 지정된 문자열이 포함된 문서를 반환합니다. 다음 예제는 Locations 사용자 지정 속성에서 값이 “시애틀” 또는 “Portland”인 모든 문서를 반환하는 방법을 보여줍니다.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {
                "ContainsAny": { "Key": "Locations", "Value": { "StringListValue":
 [ "Seattle", "Portland"] }}
            }
        )
```

검색 결과에서 각 문서의 속성을 필터링

Amazon Kendra 검색 결과의 각 문서에 대한 문서 속성을 반환합니다. 응답에 검색 결과의 일부로 포함하려는 특정 문서 속성을 필터링할 수 있습니다. 기본적으로 문서에 할당된 모든 문서 속성이 응답에 반환됩니다.

다음 예에서는 문서에 대한 응답에 _source_uri 및 _author 문서 속성만 포함됩니다.

```
response=kendra.query(
        QueryText = query,
        IndexId = index,
        RequestedDocumentAttributes = ["_source_uri", "_author"]
        )
```

사용자 컨텍스트 필터링

문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 사용자의 검색 결과를 필터링할 수 있습니다. 사용자 토큰, 사용자 ID 또는 사용자 속성을 사용하여 문서를 필터링할 수 있습니다. Amazon Kendra

는 사용자를 그룹에 매핑할 수도 있습니다. AWS IAM Identity Center 를 아이덴티티 스토어/소스로 사용하도록 선택할 수 있습니다.

사용자 컨텍스트 필터링은 문서에 대한 액세스를 제어할 수 있는 이점이 있는 일종의 개인화된 검색입니다. 예를 들어 회사 포털에서 정보를 검색하는 모든 팀이 일급 기밀 회사 문서에 액세스해야 하는 것은 아니며, 이러한 문서가 모든 사용자에게 관련된 것도 아닙니다. 일급 기밀 문서에 대한 액세스 권한을 받은 특정 사용자 또는 팀 그룹만 검색 결과에서 이러한 문서를 볼 수 있습니다.

문서가 인덱싱되면 대부분의 문서에 Amazon Kendra해당하는 ACL (액세스 제어 목록) 이 수집됩니다. ACL은 문서에 대한 액세스를 허용하거나 거부할 사용자 이름 및 그룹 이름을 지정합니다. ACL이 없는 문서는 공개 문서입니다.

Amazon Kendra 대부분의 데이터 원본에 대해 각 문서와 관련된 사용자 또는 그룹 정보를 추출할 수 있습니다. 예를 들어 Quip의 문서에는 문서에 대한 액세스 권한이 부여된 일부 사용자의 '공유' 목록이 포함될 수 있습니다. S3 버킷을 데이터 소스로 사용하는 경우 ACL용 JSON 파일을 제공하고 이 파일에 대한 S3 경로를 데이터 소스 구성의 일부로 포함해야 합니다. 문서를 색인에 직접 추가하는 경우 Principal 개체의 ACL을 BatchPutDocumentAPI의 문서 개체의 일부로 지정합니다.

CreateAccessControlConfigurationAPI를 사용하면 모든 문서를 다시 인덱싱하지 않고도 기존 문서 수준 액세스 제어를 재구성할 수 있습니다. 예를 들어 인덱스에는 특정 직원이나 사용자만 액세스할 수 있는 회사 일급 기밀 문서가 포함되어 있습니다. 이러한 사용자 중 한 명이 퇴사하거나, 일급 기밀 문서에 액세스하지 못하도록 차단해야 하는 팀으로 이동합니다. 이전에 문서를 인덱싱했을 때 해당 사용자가 액세스할 수 있었기 때문에 사용자는 여전히 일급 기밀 문서에 액세스할 수 있습니다. 이 경우 해당 사용자의 액세스를 거부하는 특정 액세스 제어 구성을 만들 수 있습니다. 나중에 사용자가 회사로 돌아와 '일급 기밀' 팀에 다시 합류할 경우 액세스를 허용하도록 액세스 제어 구성을 업데이트하면 됩니다. 상황 변화에 따라 문서에 대한 액세스 제어를 다시 구성할 수 있습니다.

액세스 제어 구성을 특정 문서에 적용하려면 Document 객체에 `AccessControlConfigurationId` 포함된 BatchPutDocument항목을 사용하여 API를 호출합니다. S3 버킷을 데이터 소스로 사용하는 경우 S3 버킷을 `.metadata.json` 로 `AccessControlConfigurationId` 업데이트하고 데이터 소스를 동기화합니다. Amazon Kendra 현재는 API를 사용하여 인덱싱된 S3 데이터 소스 및 문서에 대한 액세스 제어 구성만 지원합니다. `BatchPutDocument`

사용자 토큰별 필터링

인덱스를 쿼리할 때 사용자 토큰을 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링할 수 있습니다. 쿼리를 실행하면 토큰을 Amazon Kendra 추출 및 검증하고, 사용자 및 그룹 정보를 가져와 확인하고, 쿼리를 실행합니다. 공개 문서를 포함하여 사용자가 액세스할 수 있는 모든 문서가 반환됩니다. 자세한 내용은 토큰 기반 사용자 액세스 제어를 참조하세요.

UserContext개체에 사용자 토큰을 제공하고 이를 Query API에 전달합니다.

아래에서는 사용자 토큰을 포함하는 방법을 보여줍니다.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserToken = {
        Token = "token"
    })
```

사용자를 그룹에 매핑할 수 있습니다. 사용자-컨텍스트 필터링을 사용하는 경우 쿼리를 실행할 때 사용자가 속한 모든 그룹을 포함할 필요는 없습니다. PutPrincipalMappingAPI를 사용하여 사용자를 그룹에 매핑할 수 있습니다. PutPrincipalMapping API를 사용하지 않으려면 쿼리를 실행할 때 사용자 이름과 사용자가 속한 모든 그룹을 제공해야 합니다. 객체를 사용하여 IAM Identity Center ID 소스에 있는 그룹 및 사용자의 액세스 수준을 가져올 수도 있습니다. UserGroupResolutionConfiguration

사용자 ID 및 그룹별 필터링

인덱스를 쿼리할 때 사용자 ID와 그룹을 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링할 수 있습니다. 쿼리를 실행하면 사용자 및 그룹 정보를 Amazon Kendra 확인하고 쿼리를 실행합니다. 공개 문서를 포함하여 사용자가 액세스할 수 있는 쿼리와 관련된 모든 문서가 반환됩니다.

사용자 및 그룹이 액세스할 수 있는 데이터 소스별로 검색 결과를 필터링할 수도 있습니다. 그룹이 여러 데이터 소스에 연결되어 있지만 그룹이 특정 데이터 소스의 문서에만 액세스하도록 하려는 경우 데이터 소스를 지정하는 것이 유용합니다. 예를 들어 “연구”, “엔지니어링”, “영업 및 마케팅” 그룹은 모두 데이터 소스 Confluence 및 Salesforce에 저장된 회사 문서에 연결되어 있습니다. 하지만 “영업 및 마케팅” 팀은 Salesforce에 저장된 고객 관련 문서에만 액세스하면 됩니다. 따라서 영업 및 마케팅 사용자가 고객 관련 문서를 검색하면 검색 결과에서 Salesforce의 문서를 볼 수 있습니다. 영업 및 마케팅 부서에 속하지 않는 사용자는 검색 결과에서 Salesforce 문서를 볼 수 없습니다.

UserContext개체에 사용자, 그룹 및 데이터 소스 정보를 제공하고 이 정보를 Query API에 전달합니다. 사용자 ID, 그룹 및 데이터 소스 목록은 사용자, 그룹 및 데이터 소스를 식별하기 위해 보안 주체 객체에 지정한 이름과 일치해야 합니다. `Principal` 객체를 사용하면 문서에 액세스하기 위한 허용 목록 또는 거부 목록에 사용자, 그룹 또는 데이터 소스를 추가할 수 있습니다.

다음 중 하나를 제공해야 합니다.

- 사용자 및 그룹 정보, (선택 사항) 데이터 소스 정보.

- PutPrincipalMappingAPI를 사용하여 사용자를 그룹 및 데이터 소스에 매핑하는 경우 사용자 정보만 제공합니다. 객체를 사용하여 IAM Identity Center 자격 증명 소스에 있는 그룹 및 사용자의 액세스 수준을 가져올 수도 있습니다. UserGroupResolutionConfiguration

이 정보가 쿼리에 포함되지 않은 경우 모든 문서를 Amazon Kendra 반환합니다. 이 정보를 제공하는 경우 사용자 ID, 그룹 및 데이터 소스가 일치하는 문서만 반환됩니다.

다음은 사용자 ID, 그룹 및 데이터 소스를 포함하는 방법을 보여줍니다.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserId = {
        UserId = "user1"
    },
    Groups = {
        Groups = ["Sales and Marketing"]
    },
    DataSourceGroups = {
        DataSourceGroups = [{"DataSourceId" : "SalesforceCustomerDocsGroup", "GroupId":
 "Sales and Marketing"}]
    })
```

사용 속성으로 필터링

인덱스를 쿼리할 때 내장 속성인 _user_id 및 _group_id를 사용하여 문서에 대한 사용자 또는 그룹의 액세스 권한을 기준으로 검색 결과를 필터링할 수 있습니다. 최대 100개의 그룹 식별자를 설정할 수 있습니다. 쿼리를 실행하면 사용자 및 그룹 정보를 Amazon Kendra 확인하고 쿼리를 실행합니다. 공개 문서를 포함하여 사용자가 액세스할 수 있는 쿼리와 관련된 모든 문서가 반환됩니다.

AttributeFilter개체에 사용자 및 그룹 속성을 제공하고 이를 Query API에 전달합니다.

다음 예제는 사용자 ID와 사용자가 속한 “HR” 및 “IT” 그룹을 기반으로 쿼리 응답을 필터링하는 요청을 보여줍니다. 쿼리는 허용 목록에 사용자 또는 “HR” 또는 “IT” 그룹이 있는 모든 문서를 반환합니다. 사용자 또는 그룹 중 하나가 문서에 대한 거부 목록에 있는 경우 문서는 반환되지 않습니다.

```
response = kendra.query(
        QueryText = query,
        IndexId = index,
        AttributeFilter = {
```

```
            "OrAllFilters": [
                {
                    "EqualsTo": {
                        "Key": "_user_id",
                        "Value": {
                            "StringValue": "user1"
                        }
                     }
                },
                {
                    "EqualsTo": {
                        "Key": "_group_ids",
                        "Value": {
                            "StringListValue": ["HR", "IT"]
                        }
                    }
                }
            ]
        }
        )
```

또한 `Principal` 객체에서 그룹이 액세스할 수 있는 데이터 소스를 지정할 수 있습니다.

> **Note**
>
> 사용자 컨텍스트 필터링은 콘텐츠에 대한 인증 또는 권한 부여 제어가 아닙니다. `Query` API로 전송된 사용자 및 그룹에 대해서는 사용자 인증을 수행하지 않습니다. `Query` API로 전송된 사용자 및 그룹 정보가 인증되고 승인되었는지 확인하는 것은 애플리케이션에 달려 있습니다.

각 데이터 소스에 대해 사용자 컨텍스트 필터링이 구현되어 있습니다. 다음 섹션에서는 각 구현에 대해 설명합니다.

주제

- 인덱스에 직접 추가된 문서에 대한 사용자 컨텍스트 필터링
- 자주 묻는 질문에 대한 사용자 컨텍스트 필터링
- 데이터 소스의 사용자 컨텍스트 필터링

인덱스에 직접 추가된 문서에 대한 사용자 컨텍스트 필터링

BatchPutDocumentAPI를 사용하여 색인에 직접 문서를 추가하면 문서 `AccessControlList` 필드에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 문서에 대한 액세스 제어 목록(ACL)을 제공하면 ACL이 문서와 함께 수집됩니다.

보안 주체 객체의 ACL을 `BatchPutDocument` API의 문서 객체의 일부로 지정합니다. 다음 정보를 제공합니다.

- 사용자 또는 그룹이 가져야 하는 액세스 권한 `ALLOW` 또는 `DENY`라고 말할 수 있습니다.
- 개체의 유형입니다. `USER` 또는 `GROUP`라고 말할 수 있습니다.
- 사용자 또는 그룹의 이름입니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

자주 묻는 질문에 대한 사용자 컨텍스트 필터링

색인에 FAQ를 추가하면 FAQ JSON 파일의 `AccessControlList` 개체/필드에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 액세스 제어를 위한 사용자 지정 필드 또는 속성이 포함된 FAQ CSV 파일을 사용할 수도 있습니다.

다음 정보를 제공합니다.

- 사용자 또는 그룹이 가져야 하는 액세스 권한 `ALLOW` 또는 `DENY`라고 말할 수 있습니다.
- 개체의 유형입니다. `USER` 또는 `GROUP`라고 말할 수 있습니다.
- 사용자 또는 그룹의 이름입니다.

자세한 내용은 FAQ 파일을 참조하세요.

데이터 소스의 사용자 컨텍스트 필터링

Amazon Kendra 또한 지원되는 데이터 원본 커넥터에서 사용자 및 그룹 ACL (액세스 제어 목록) 정보를 크롤링합니다. 이는 문서에 대한 사용자 또는 해당 그룹의 액세스를 기반으로 검색 결과를 필터링하는 사용자 컨텍스트 필터링에 유용합니다.

주제

- Adobe Experience Manager 데이터 소스에 대한 사용자 컨텍스트 필터링
- Alfresco 데이터 소스의 사용자 컨텍스트 필터링

- Aurora (MySQL) 데이터 원본에 대한 사용자 컨텍스트 필터링
- Aurora (PostgreSQL) 데이터 소스의 사용자 컨텍스트 필터링
- 데이터 소스의 사용자 컨텍스트 필터링 Amazon FSx
- 데이터베이스 데이터 소스의 사용자 컨텍스트 필터링
- Amazon RDS (Microsoft SQL Server) 데이터 소스에 대한 사용자 컨텍스트 필터링
- Amazon RDS (MySQL) 데이터 원본에 대한 사용자 컨텍스트 필터링
- Amazon RDS (Oracle) 데이터 소스에 대한 사용자 컨텍스트 필터링
- Amazon RDS (PostgreSQL) 데이터 소스의 사용자 컨텍스트 필터링
- Amazon S3 데이터 소스의 사용자 컨텍스트 필터링
- Amazon WorkDocs 데이터 소스의 사용자 컨텍스트 필터링
- Box 데이터 소스의 사용자 컨텍스트 필터링
- Confluence 데이터 소스의 사용자 컨텍스트 필터링
- Dropbox 데이터 소스의 사용자 컨텍스트 필터링
- Drupal 데이터 소스의 사용자 컨텍스트 필터링
- 데이터 소스의 사용자 컨텍스트 필터링 GitHub
- Gmail 데이터 소스의 사용자 컨텍스트 필터링
- Google Drive 데이터 소스의 사용자 컨텍스트 필터링
- IBM DB2 데이터 소스의 사용자 컨텍스트 필터링
- Jira 데이터 소스의 사용자 컨텍스트 필터링
- Microsoft Exchange 데이터 소스에 대한 사용자 컨텍스트 필터링
- Microsoft OneDrive 데이터 원본에 대한 사용자 컨텍스트 필터링
- Microsoft OneDrive v2.0 데이터 소스에 대한 사용자 컨텍스트 필터링
- Microsoft SharePoint 데이터 원본에 대한 사용자 컨텍스트 필터링
- Microsoft SQL Server 데이터 소스에 대한 사용자 컨텍스트 필터링
- Microsoft Teams 데이터 소스에 대한 사용자 컨텍스트 필터링
- Microsoft Yammer 데이터 소스에 대한 사용자 컨텍스트 필터링
- MySQL 데이터 소스의 사용자 컨텍스트 필터링
- Oracle Database 데이터 소스에 대한 사용자 컨텍스트 필터링
- PostgreSQL 데이터 소스에 대한 사용자 컨텍스트 필터링
- Quip 데이터 소스의 사용자 컨텍스트 필터링

- Salesforce 데이터 소스의 사용자 컨텍스트 필터링
- ServiceNow 데이터 소스의 사용자 컨텍스트 필터링
- Slack 데이터 소스의 사용자 컨텍스트 필터링
- Zendesk 데이터 소스의 사용자 컨텍스트 필터링

Adobe Experience Manager 데이터 소스에 대한 사용자 컨텍스트 필터링

Adobe Experience Manager 데이터 소스를 사용하는 경우 Adobe Experience Manager 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

그룹 및 사용자 ID는 다음과 같이 매핑됩니다.

- `_group_ids` - 그룹 ID는 액세스 권한이 설정된 Adobe Experience Manager 콘텐츠에 존재합니다. 이들은 Adobe Experience Manager의 그룹 이름을 기반으로 매핑됩니다.
- `_user_id` - 사용자 ID는 액세스 권한이 설정된 Adobe Experience Manager 콘텐츠에 존재합니다. 사용자 이메일에서 Adobe Experience Manager의 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Alfresco 데이터 소스의 사용자 컨텍스트 필터링

Alfresco 데이터 소스를 사용하는 경우 Alfresco 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

그룹 및 사용자 ID는 다음과 같이 매핑됩니다.

- `_group_ids` - Alfresco에는 설정된 액세스 권한이 있는 파일의 그룹 ID가 있습니다. Alfresco에 있는 그룹의 시스템 이름(디스플레이 이름 아님)에서 매핑됩니다.
- `_user_id` - Alfresco에는 설정된 액세스 권한이 있는 파일의 사용자 ID가 있습니다. 사용자 이메일에서 Alfresco의 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Aurora (MySQL) 데이터 원본에 대한 사용자 컨텍스트 필터링

Aurora (MySQL) 데이터 원본을 사용하는 경우 원본 테이블의 열에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 CreateDataSourceAPI의 일부로 사용하여 지정합니다.

Aurora (MySQL) 데이터베이스 데이터 원본에는 다음과 같은 제한이 있습니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Aurora (PostgreSQL) 데이터 소스의 사용자 컨텍스트 필터링

Aurora (PostgreSQL) 데이터 원본을 사용하는 경우 원본 Amazon Kendra 테이블의 열에서 사용자 및 그룹 정보를 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 API의 일부로 사용하여 지정합니다. CreateDataSource

Aurora (PostgreSQL) 데이터베이스 데이터 원본에는 다음과 같은 제한이 있습니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

데이터 소스의 사용자 컨텍스트 필터링 Amazon FSx

Amazon FSx 데이터 원본을 사용하는 경우 Amazon FSx 인스턴스의 디렉터리 서비스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

Amazon FSx 그룹 및 사용자 ID는 다음과 같이 매핑됩니다.

- `_group_ids` - Amazon FSx 에는 설정된 액세스 권한이 있는 파일의 그룹 ID가 있습니다. 이 둘은 의 디렉토리 서비스에 있는 시스템 그룹 이름에서 매핑됩니다. Amazon FSx
- `_user_id`—액세스 권한이 설정된 파일에 사용자 ID가 존재합니다. Amazon FSx 사용자 이름은 의 디렉토리 서비스에 있는 시스템 사용자 이름에서 매핑됩니다. Amazon FSx

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

데이터베이스 데이터 소스의 사용자 컨텍스트 필터링

와 같은 Amazon Aurora PostgreSQL데이터베이스 데이터 원본을 사용하는 경우 원본 테이블의 열에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 개체의 이 열을 CreateDataSourceAPI의 AclConfigurationDatabaseConfiguration개체 일부로 지정합니다.

데이터베이스 데이터 소스에는 다음과 같은 제한 사항이 적용됩니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Amazon RDS (Microsoft SQL Server) 데이터 소스에 대한 사용자 컨텍스트 필터링

Amazon RDS (Microsoft SQL Server) 데이터 원본을 사용하는 경우 원본 테이블의 열에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 CreateDataSourceAPI의 일부로 사용하여 지정합니다.

Amazon RDS (Microsoft SQL Server) 데이터베이스 데이터 원본에는 다음과 같은 제한이 있습니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Amazon RDS (MySQL) 데이터 원본에 대한 사용자 컨텍스트 필터링

Amazon RDS (MySQL) 데이터 원본을 사용하는 경우 원본 테이블의 열에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 CreateDataSourceAPI의 일부로 사용하여 지정합니다.

Amazon RDS (MySQL) 데이터베이스 데이터 원본에는 다음과 같은 제한이 있습니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Amazon RDS (Oracle) 데이터 소스에 대한 사용자 컨텍스트 필터링

Amazon RDS (Oracle) 데이터 원본을 사용하는 경우 원본 테이블의 열에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 CreateDataSourceAPI의 일부로 사용하여 지정합니다.

Amazon RDS (Oracle) 데이터베이스 데이터 원본에는 다음과 같은 제한이 있습니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Amazon RDS (PostgreSQL) 데이터 소스의 사용자 컨텍스트 필터링

Amazon RDS (PostgreSQL) 데이터 원본을 사용하는 경우 원본 Amazon Kendra 테이블의 열에서 사용자 및 그룹 정보를 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 API의 일부로 사용하여 지정합니다. CreateDataSource

Amazon RDS (PostgreSQL) 데이터베이스 데이터 원본에는 다음과 같은 제한이 있습니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Amazon S3 데이터 소스의 사용자 컨텍스트 필터링

문서와 연결된 메타데이터 파일을 사용하여 Amazon S3 데이터 소스의 문서에 사용자 컨텍스트 필터링을 추가합니다. JSON 문서의 `AccessControlList` 필드에 정보를 추가합니다. Amazon S3 데이터 소스에서 인덱싱된 문서에 메타데이터를 추가하는 방법에 대한 자세한 내용은 S3 문서 메타데이터를 참조하세요.

다음과 같은 세 가지 정보를 제공합니다.

- 개체가 가져야 하는 액세스 권한. `ALLOW` 또는 `DENY`라고 말할 수 있습니다.
- 개체의 유형입니다. `USER` 또는 `GROUP`라고 말할 수 있습니다.
- 개체의 이름입니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Amazon WorkDocs 데이터 소스의 사용자 컨텍스트 필터링

Amazon WorkDocs 데이터 원본을 사용하는 경우 Amazon WorkDocs 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

Amazon WorkDocs 그룹 및 사용자 ID는 다음과 같이 매핑됩니다.

- `_group_ids`—그룹 ID는 액세스 권한이 설정된 파일에 존재합니다 Amazon WorkDocs . 에 있는 그룹 이름을 기반으로 매핑됩니다. Amazon WorkDocs
- `_user_id`—액세스 권한이 설정된 파일에 사용자 ID가 있습니다. Amazon WorkDocs 의 사용자 이름에서 매핑됩니다. Amazon WorkDocs

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Box 데이터 소스의 사용자 컨텍스트 필터링

Box 데이터 소스를 사용하는 경우 Box 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

Box 그룹 및 사용자 ID는 다음과 같이 매핑됩니다.

- `_group_ids` - Box에는 설정된 액세스 권한이 있는 파일의 그룹 ID가 있습니다. Box에서 그룹의 이름을 기반으로 매핑됩니다.
- `_user_id` - Box에는 설정된 액세스 권한이 있는 파일의 사용자 ID가 있습니다. 사용자 이메일에서 Box의 사용자 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Confluence 데이터 소스의 사용자 컨텍스트 필터링

Confluence 데이터 소스를 사용하는 경우 Confluence 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

스페이스 권한 페이지를 사용하여 스페이스에 대한 사용자 및 그룹 액세스를 구성합니다. 페이지와 블로그의 경우 제한 페이지를 사용합니다. 스페이스 권한에 대한 자세한 내용은 Confluence 지원 웹 사이트의 스페이스 권한 개요를 참조하세요. 페이지 및 블로그 제한에 대한 자세한 내용은 Confluence 지원 웹 사이트의 페이지 제한을 참조하세요.

Confluence 그룹 및 사용자 이름은 다음과 같이 매핑됩니다.

- `_group_ids` - 그룹 이름은 제한이 있는 스페이스, 페이지 및 블로그에 표시됩니다. Confluence에서 그룹의 이름을 기반으로 매핑됩니다. 그룹 이름은 항상 소문자입니다.

- `_user_id` - 사용자 이름은 제한이 있는 스페이스, 페이지 및 블로그에 표시됩니다. 사용 중인 Confluence 인스턴스 유형에 따라 매핑됩니다.

Confluence 커넥터 v1.0의 경우

- 서버 - `_user_id`는 사용자 이름입니다. 사용자 이름은 항상 소문자입니다.
- 클라우드 - `_user_id`는 사용자의 계정 ID입니다.

Confluence 커넥터 v2.0의 경우

- 서버 - `_user_id`는 사용자 이름입니다. 사용자 이름은 항상 소문자입니다.
- 클라우드 - `_user_id`는 사용자의 이메일 ID입니다.

> **Important**
>
> Confluence 커넥터에서 사용자 컨텍스트 필터링이 제대로 작동하려면 Confluence 페이지에 대한 액세스 권한이 부여된 사용자의 표시 여부를 모두로 설정해야 합니다. 자세한 내용은 Atlassian 개발자 설명서에서 이메일 표시 여부 설정을 참조하세요.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Dropbox 데이터 소스의 사용자 컨텍스트 필터링

Dropbox 데이터 소스를 사용하는 경우 Dropbox 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

그룹 및 사용자 ID는 다음과 같이 매핑됩니다.

- `_group_ids` - Dropbox에는 설정된 액세스 권한이 있는 파일의 그룹 ID가 있습니다. Box에서 그룹의 이름을 기반으로 매핑됩니다.
- `_user_id` - Dropbox에는 설정된 액세스 권한이 있는 파일의 사용자 ID가 있습니다. 사용자 이메일에서 Dropbox의 사용자 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Drupal 데이터 소스의 사용자 컨텍스트 필터링

Drupal 데이터 소스를 사용하는 경우 Drupal 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

그룹 및 사용자 ID는 다음과 같이 매핑됩니다.

- `_group_ids` - Drupal에는 설정된 액세스 권한이 있는 파일의 그룹 ID가 있습니다. Drupal에서 그룹의 이름을 기반으로 매핑됩니다.
- `_user_id` - Drupal에는 설정된 액세스 권한이 있는 파일의 사용자 ID가 있습니다. 사용자 이메일에서 Drupal의 사용자 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

데이터 소스의 사용자 컨텍스트 필터링 GitHub

GitHub 데이터 원본을 사용하는 경우 GitHub 인스턴스에서 사용자 정보를 Amazon Kendra 가져옵니다.

GitHub 사용자 ID는 다음과 같이 매핑됩니다.

- `_user_id`—액세스 권한이 설정된 파일에 사용자 ID가 존재합니다. GitHub 사용자 이메일에서 ID로 매핑됩니다. GitHub

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Gmail 데이터 소스의 사용자 컨텍스트 필터링

Gmail 데이터 소스를 사용하는 경우 Gmail 인스턴스에서 사용자 정보를 Amazon Kendra 가져옵니다.

사용자 ID는 다음과 같이 매핑됩니다.

- `_user_id` - Gmail에는 설정된 액세스 권한이 있는 파일의 사용자 ID가 있습니다. 사용자 이메일에서 Gmail의 사용자 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Google Drive 데이터 소스의 사용자 컨텍스트 필터링

Google Workspace Drive 데이터 소스는 Google Drive 사용자 및 그룹의 사용자 및 그룹 정보를 반환합니다. 그룹 및 도메인 멤버십은 `_group_ids` 인덱스 필드에 매핑됩니다. Google Drive 사용자 이름이 `_user_id` 필드에 매핑됩니다.

`Query` API에 사용자 이메일 주소를 하나 이상 입력하면 해당 이메일 주소와 공유된 문서만 반환됩니다. 다음 `AttributeFilter` 파라미터는 “martha@example.com”과 공유한 문서만 반환합니다.

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_user_id",
                   "Value": {
                       "StringValue": "martha@example.com"
                   }
               }
           }
```

쿼리에 그룹 이메일 주소를 하나 이상 입력하면 그룹과 공유된 문서만 반환됩니다. 다음 `AttributeFilter` 파라미터는 “hr@example.com” 그룹과 공유한 문서만 반환합니다.

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_group_ids",
                   "Value": {
                       "StringListValue": ["hr@example.com"]
                   }
               }
           }
```

쿼리에 도메인을 입력하면 도메인과 공유된 모든 문서가 반환됩니다. 다음 `AttributeFilter` 파라미터는 “example.com” 도메인과 공유된 문서를 반환합니다.

```
"AttributeFilter": {
                "EqualsTo":{
                   "Key": "_group_ids",
                   "Value": {
                       "StringListValue": ["example.com"]
                   }
               }
           }
```

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

IBM DB2 데이터 소스의 사용자 컨텍스트 필터링

IBM DB2 데이터 소스를 사용하는 경우 소스 테이블의 열에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 CreateDataSourceAPI의 일부로 사용하여 지정합니다.

IBM DB2 데이터베이스 데이터 소스에는 다음과 같은 제한 사항이 적용됩니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Jira 데이터 소스의 사용자 컨텍스트 필터링

Jira 데이터 소스를 사용하는 경우 Jira 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

Jira 사용자 ID는 다음과 같이 매핑됩니다.

- `_user_id` - Jira에는 설정된 액세스 권한이 있는 파일의 사용자 ID가 있습니다. 사용자 이메일에서 Jira의 사용자 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Microsoft Exchange 데이터 소스에 대한 사용자 컨텍스트 필터링

Microsoft Exchange 데이터 원본을 사용하는 경우 Microsoft Exchange 인스턴스에서 사용자 정보를 Amazon Kendra 가져옵니다.

Microsoft Exchange 사용자 ID는 다음과 같이 매핑됩니다.

- `_user_id`—Microsoft Exchange 권한에는 사용자가 특정 콘텐츠에 액세스할 수 있는 사용자 ID가 있습니다. 이들은 Microsoft Exchange의 사용자 이름에서 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Microsoft OneDrive 데이터 원본에 대한 사용자 컨텍스트 필터링

Amazon Kendra 사이트에서 문서를 OneDrive 인덱싱할 때 Microsoft로부터 사용자 및 그룹 정보를 검색합니다. 사용자 및 그룹 정보는 호스팅하는 기본 Microsoft SharePoint 사이트에서 가져옵니다 OneDrive.

OneDrive 사용자 또는 그룹을 사용하여 검색 결과를 필터링하는 경우 다음과 같이 ID를 계산하십시오.

1. 사이트 이름을 변경합니다. 예제: `https://host.onmicrosoft.com/sites/siteName.`

2. 사이트 이름의 MD5 해시를 가져오세요. 예를 들어 `430a6b90503eef95c89295c8999c7981`입니다.
3. MD5 해시를 세로 막대(|)와 ID로 연결하여 사용자 이메일 또는 그룹 ID를 생성합니다. 예를 들어, 그룹 이름이 "localGroupName“인 경우 그룹 ID는 다음과 같습니다.

   `"430a6b90503eef95c89295c8999c7981 | localGroupName"`

   > **Note**
   >
   > 세로 막대 앞과 뒤에 공백을 포함하세요. 세로 막대는 MD5 `localGroupName` 해시를 식별하는 데 사용됩니다.

   사용자 이름 “someone@host.onmicrosoft.com”의 경우 사용자 ID는 다음과 같습니다.

   `"430a6b90503eef95c89295c8999c7981 | someone@host.onmicrosoft.com"`

Query API를 호출할 Amazon Kendra 때 사용자 `_user_id` 또는 그룹 ID를 or `_group_id` 속성으로 보내십시오. 예를 들어 그룹을 사용하여 검색 결과를 필터링하는 AWS CLI 명령은 다음과 같습니다.

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "430a6b90503eef95c89295c8999c7981 |
 localGroupName"}
                  }}'
```

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Microsoft OneDrive v2.0 데이터 소스에 대한 사용자 컨텍스트 필터링

Microsoft OneDrive v2.0 데이터 소스는 OneDrive 액세스 제어 목록 (ACL) 엔티티로부터 섹션 및 페이지 정보를 반환합니다. Amazon Kendra OneDrive 테넌트 도메인을 사용하여 OneDrive 인스턴스에 연결한 다음 섹션 및 파일 이름에 대한 사용자 또는 그룹 액세스를 기반으로 검색 결과를 필터링할 수 있습니다.

표준 객체의 경우 `_user_id` 및 `_group_id`는 다음과 같이 사용됩니다.

- _user_id— Microsoft OneDrive 사용자 이메일 ID가 _user_id 필드에 매핑됩니다.
- _group_id— Microsoft OneDrive 그룹 이메일이 _group_id 필드에 매핑됩니다.

AccessControlList 필드에 개체를 200개까지 추가할 수 있습니다.

Microsoft SharePoint 데이터 원본에 대한 사용자 컨텍스트 필터링

Amazon Kendra 사이트 문서를 SharePoint 인덱싱할 때 Microsoft로부터 사용자 및 그룹 정보를 검색합니다. 사용자 또는 그룹 액세스를 기준으로 검색 결과를 필터링하려면 API를 호출할 때 사용자 및 그룹 정보를 제공하십시오. Query

사용자 이름을 사용하여 필터링하려면 사용자의 이메일 주소를 사용합니다. 예: johnstiles@example.com.

SharePoint 그룹을 사용하여 검색 결과를 필터링하는 경우 다음과 같이 그룹 ID를 계산하십시오.

로컬 그룹의 경우

1. 사이트 이름을 변경합니다. 예제: https://host.onmicrosoft.com/sites/siteName.
2. 사이트 이름의 SHA256 해시를 가져옵니다. 예를 들어 430a6b90503eef95c89295c8999c7981입니다.
3. SHA256 해시를 세로 막대(|)와 그룹 이름으로 연결하여 그룹 ID를 생성합니다. 예를 들어, 그룹 이름이 "localGroupName“인 경우 그룹 ID는 다음과 같습니다.

   "430a6b90503eef95c89295c8999c7981 | localGroupName"

> **Note**
>
> 세로 막대 앞과 뒤에 공백을 포함하세요. 세로 막대는 SHA256 해시를 식별하는 localGroupName 데 사용됩니다.

쿼리 API를 호출할 Amazon Kendra 때 그룹 ID를 _group_id 속성으로 전송하십시오. 예를 들어, AWS CLI 명령은 다음과 같습니다.

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
```

```
                  "EqualsTo":{
                    "Key": "_group_id",
                    "Value": {"StringValue": "430a6b90503eef95c89295c8999c7981 |
localGroupName"}
                 }}'
```

AD 그룹의 경우

1. AD 그룹 ID를 사용하여 검색 결과 필터링을 구성할 수 있습니다.

쿼리 API를 호출할 Amazon Kendra 때 그룹 ID를 _group_id 속성으로 전송하십시오. 예를 들어, AWS CLI 명령은 다음과 같습니다.

```
aws kendra  query \
                --index-id index ID
                --query-text "query text"
                --attribute-filter '{
                   "EqualsTo":{
                     "Key": "_group_id",
                     "Value": {"StringValue": "AD group"}
                  }}'
```

AccessControlList 필드에 개체를 200개까지 추가할 수 있습니다.

Microsoft SQL Server 데이터 소스에 대한 사용자 컨텍스트 필터링

Microsoft SQL Server 데이터 원본을 사용하는 경우 원본 테이블의 열에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 CreateDataSourceAPI의 일부로 사용하여 지정합니다.

Microsoft SQL Server 데이터베이스 데이터 소스에는 다음과 같은 제한 사항이 적용됩니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Microsoft Teams 데이터 소스에 대한 사용자 컨텍스트 필터링

Amazon Kendra 문서를 인덱싱할 때 Microsoft Teams에서 사용자 정보를 검색합니다. 사용자 정보는 기본 Microsoft Teams 인스턴스에서 가져옵니다.

AccessControlList 필드에 개체를 200개까지 추가할 수 있습니다.

Microsoft Yammer 데이터 소스에 대한 사용자 컨텍스트 필터링

Amazon Kendra 문서를 인덱싱할 때 Microsoft Yammer에서 사용자 정보를 검색합니다. 사용자 및 그룹 정보는 기본 Microsoft Yammer 인스턴스에서 가져옵니다.

Microsoft Yammer 사용자 ID는 다음과 같이 매핑됩니다.

- _email_id— _user_id 필드에 매핑된 Microsoft 이메일 ID입니다.

AccessControlList 필드에 개체를 200개까지 추가할 수 있습니다.

MySQL 데이터 소스의 사용자 컨텍스트 필터링

MySQL 데이터 원본을 사용하는 경우 원본 테이블의 열에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 CreateDataSourceAPI의 일부로 사용하여 지정합니다.

MySQL 데이터베이스 데이터 소스에는 다음과 같은 제한 사항이 적용됩니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Oracle Database 데이터 소스에 대한 사용자 컨텍스트 필터링

Oracle Database 데이터 원본을 사용하는 경우 원본 테이블의 열에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 CreateDataSourceAPI의 일부로 사용하여 지정합니다.

Oracle Database 데이터 소스에는 다음과 같은 제한 사항이 적용됩니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

PostgreSQL 데이터 소스에 대한 사용자 컨텍스트 필터링

PostgreSQL 데이터 원본을 사용하는 경우 원본 Amazon Kendra 테이블의 열에서 사용자 및 그룹 정보를 가져옵니다. 콘솔에서 이 열을 지정하거나 TemplateConfiguration객체를 API의 일부로 사용하여 지정합니다. CreateDataSource

PostgreSQL 데이터베이스 데이터 소스에는 다음과 같은 제한 사항이 적용됩니다.

- 데이터베이스 데이터 소스의 허용 목록만 지정할 수 있습니다. 거부 목록을 지정할 수 없습니다.
- 그룹만 지정할 수 있습니다. 허용 목록에 개별 사용자를 지정할 수는 없습니다.
- 데이터베이스 열은 세미콜론으로 구분된 그룹 목록을 포함하는 문자열이어야 합니다.

Quip 데이터 소스의 사용자 컨텍스트 필터링

Quip 데이터 소스를 사용하는 경우 Quip 인스턴스에서 사용자 정보를 Amazon Kendra 가져옵니다.

Quip 사용자 ID는 다음과 같이 매핑됩니다.

- `_user_id` - Quip에는 설정된 액세스 권한이 있는 파일의 사용자 ID가 있습니다. 사용자 이메일에서 Quip의 사용자 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Salesforce 데이터 소스의 사용자 컨텍스트 필터링

Salesforce 데이터 소스는 Salesforce 액세스 제어 목록(ACL) 개체로부터 사용자 및 그룹 정보를 반환합니다. 사용자 컨텍스트 필터링을 Salesforce 표준 객체 및 채터 피드에 적용할 수 있습니다. Salesforce 지식 문서에는 사용자 컨텍스트 필터링을 사용할 수 없습니다.

Salesforce 필드를 Amazon Kendra 문서 제목 및 문서 본문 필드에 매핑하는 경우 Amazon Kendra는 검색 응답에서 문서 제목 및 본문 필드의 데이터를 사용합니다.

표준 객체의 경우 `_user_id` 및 `_group_ids`는 다음과 같이 사용됩니다.

- `_user_id` - Salesforce 사용자의 사용자 이름입니다.
- `_group_ids`—
  - Salesforce `Profile`의 이름
  - Salesforce `Group`의 이름

- Salesforce `UserRole`의 이름
- Salesforce `PermissionSet`의 이름

채터 피드의 경우 `_user_id` 및 `_group_ids`는 다음과 같이 사용됩니다.

- `_user_id` - Salesforce 사용자의 사용자 이름입니다. 항목이 사용자 피드에 게시된 경우에만 사용할 수 있습니다.
- `_group_ids` - 그룹 ID는 다음과 같이 사용됩니다. 피드 항목이 채터 또는 공동 작업 그룹에 게시된 경우에만 사용할 수 있습니다.
  - 채터 또는 공동 작업 그룹의 이름.
  - 그룹이 공개 그룹인 경우 `PUBLIC:ALL`.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

ServiceNow 데이터 소스의 사용자 컨텍스트 필터링

사용자 컨텍스트 ServiceNow 필터링은 API 및 커넥터 v2.0에서만 지원됩니다. TemplateConfiguration ServiceNow ServiceNowConfigurationAPI 및 ServiceNow 커넥터 v1.0은 사용자 컨텍스트 필터링을 지원하지 않습니다.

ServiceNow 데이터 소스를 사용하는 경우 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다. ServiceNow

그룹 및 사용자 ID는 다음과 같이 매핑됩니다.

- `_group_ids`—그룹 ID는 액세스 권한이 설정된 파일에 존재합니다 ServiceNow . 이들은 in의 `sys_ids` 역할 이름에서 매핑됩니다. ServiceNow
- `_user_id`—액세스 권한이 설정된 파일에 사용자 ID가 존재합니다. ServiceNow 사용자 이메일에서 ID로 매핑됩니다. ServiceNow

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Slack 데이터 소스의 사용자 컨텍스트 필터링

Slack 데이터 소스를 사용하는 경우 Slack 인스턴스에서 사용자 정보를 Amazon Kendra 가져옵니다.

Slack 사용자 ID는 다음과 같이 매핑됩니다.

- `_user_id` - Slack에서 액세스 권한이 설정된 메시지 및 채널의 사용자 ID가 존재합니다. 사용자 이메일에서 Slack의 사용자 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

Zendesk 데이터 소스의 사용자 컨텍스트 필터링

Zendesk 데이터 소스를 사용하는 경우 Zendesk 인스턴스에서 사용자 및 그룹 정보를 Amazon Kendra 가져옵니다.

그룹 및 사용자 ID는 다음과 같이 매핑됩니다.

- `_group_ids` - 그룹 ID는 액세스 권한이 설정된 Zendesk 티켓 및 문서에 있습니다. Zendesk에서 그룹의 이름을 기반으로 매핑됩니다.
- `_user_id` - 그룹 ID는 액세스 권한이 설정된 Zendesk 티켓 및 문서에 있습니다. 사용자 이메일에서 Zendesk의 사용자 ID로 매핑됩니다.

`AccessControlList` 필드에 개체를 200개까지 추가할 수 있습니다.

쿼리 응답 및 응답 유형

Amazon Kendra 다양한 쿼리 응답 및 응답 유형을 지원합니다.

쿼리 응답

Query API를 호출하면 검색 결과에 대한 정보가 반환됩니다. 결과는 QueryResultItem객체 배열 (`ResultItems`) 으로 표시됩니다. 각 `QueryResultItem` 항목에는 결과 요약이 포함됩니다. 쿼리 결과와 관련된 문서 속성이 포함됩니다.

요약 정보

요약 정보는 결과 유형에 따라 달라집니다. 각 경우에 검색어와 일치하는 문서 텍스트가 포함됩니다. 또한 애플리케이션 출력에서 검색 텍스트를 강조 표시하는 데 사용할 수 있는 강조 표시 정보도 포함되어 있습니다. 예를 들어 검색어가 스페이스 니들 높이는 얼마인가요?인 경우, 요약 정보에는 높이 및 스페이스 니들이라는 단어의 텍스트 위치가 포함됩니다. 응답 유형에 대한 자세한 내용은 쿼리 응답 및 응답 유형 단원을 참조하세요.

문서 속성

각 결과에는 쿼리와 일치하는 문서의 문서 속성이 포함됩니다. DocumentId, DocumentTitle, DocumentUri 등 일부 속성은 미리 정의되어 있습니다. 다른 것들은 사용자가 정의하는 사용자 지정 속성입니다. 문서 속성을 사용하여 Query API의 응답을 필터링할 수 있습니다. 예를 들어, 특정 작성자가 작성한 문서나 문서의 특정 버전만 원할 수 있습니다. 자세한 정보는 검색 필터링 및 패싯을 참조하세요. 인덱스에 문서를 추가할 때 문서 속성을 지정합니다. 자세한 내용은 사용자 지정 필드 또는 속성을 참조하세요.

다음은 쿼리 결과를 위한 샘플 JSON 코드입니다. DocumentAttributes 및 AdditionalAttributes의 문서 속성을 기록해 두세요.

```
{
    "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"
                }, {
                    "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
}
```

응답 유형

Amazon Kendra 세 가지 유형의 쿼리 응답을 반환합니다.

- 답변(테이블 답변 포함)
- 문서
- 질문 및 답변

응답 유형이 QueryResultItem개체의 Type 응답 필드에 반환됩니다.

답변

Amazon Kendra 응답에서 질문 답변을 하나 이상 감지했습니다. 팩토이드는 누가, 무엇을, 언제, 어디를 묻는 질문(예: 가장 가까운 서비스 센터는 어디입니까?)에 대한 응답입니다. Amazon Kendra 는 쿼리와 가장 일치하는 인덱스의 텍스트를 반환합니다. 텍스트는 AnswerText 필드에 있으며 응답 텍스트 내 검색어에 대한 강조 표시 정보를 포함합니다. AnswerText는 전체 문서 발췌문과 강조 표시된 텍스트를 포함하고, DocumentExcerpt는 잘린(290자) 문서 발췌문과 강조 표시된 텍스트를 포함합니다.

Amazon Kendra 문서당 하나의 답변만 반환하며, 이 답변은 신뢰도가 가장 높은 답변입니다. 한 문서에서 여러 개의 답을 반환하려면 문서를 여러 문서로 분할해야 합니다.

```
{
    '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'
}
```

문서

Amazon Kendra 검색어와 일치하는 문서 중 순위가 매겨진 문서를 반환합니다. 순위는 검색 결과의 정확성에 대한 신뢰도를 기반으로 합니다. Amazon Kendra 일치하는 문서에 대한 정보는 에 반환됩니다 QueryResultItem. 문서의 제목이 들어 있습니다. 발췌문에는 검색 텍스트의 강조 표시 정보와 문서 내 일치하는 텍스트 섹션이 포함되어 있습니다. 일치하는 문서의 URI는 SourceURI 문서 속성에 있습니다. 다음 샘플 JSON은 일치하는 문서에 대한 문서 요약을 보여줍니다.

```
{
    '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'
}
```

질문 및 답변

질문이 색인의 자주 묻는 질문 중 Amazon Kendra 하나와 일치하면 질문 및 답변 응답이 반환됩니다. 응답에는 QueryResultItem필드의 일치하는 질문과 답변이 포함됩니다. 또한 쿼리 문자열에서 감지된 쿼리 용어에 대한 강조 표시 정보도 포함됩니다. 다음 JSON은 질문 및 답변 응답을 보여줍니다. 응답에는 질문 텍스트가 포함되어 있다는 점에 유의하세요.

```
{
    '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?'
    }
}
```

인덱스에 질문 및 답변 텍스트를 추가하는 방법에 대한 자세한 내용은 FAQ 생성을 참조하세요.

응답 조정 및 정렬

관련성 조정을 통해 검색 관련성에 대한 필드 또는 속성의 영향을 수정할 수 있습니다. 특정 속성 또는 필드를 기준으로 검색 결과를 정렬할 수도 있습니다.

주제

- 응답 조정
- 응답 정렬

응답 조정

관련성 조정을 통해 검색 관련성에 대한 필드 또는 속성의 영향을 수정할 수 있습니다. 관련성 조정을 빠르게 테스트하려면 Query API를 사용하여 조정 구성을 쿼리에 전달하세요. 그러면 다양한 구성에서 얻은 다양한 검색 결과를 볼 수 있습니다. 콘솔에서는 쿼리 수준의 관련성 조정이 지원되지 않습니다. `StringList` 유형의 필드나 속성도 인덱스 수준에서만 조정할 수 있습니다. 자세한 내용은 검색 관련성 조정을 참조하세요.

기본적으로 쿼리 응답은 응답의 각 결과를 Amazon Kendra 결정하는 관련성 점수를 기준으로 정렬됩니다.

다음 유형의 기본 제공 또는 사용자 지정 속성/필드에 대한 결과를 조정할 수 있습니다.

- 날짜 값
- long 값
- 문자열 값

다음 유형의 속성은 정렬할 수 없습니다.

- 문자열 목록 값

문서 결과 (AWS SDK) 의 순위 지정 및 조정

`Searchable` 파라미터를 true로 설정하면 문서 메타데이터 구성이 향상됩니다.

쿼리의 속성을 조정하려면 `Query` API의 `DocumentRelevanceOverrideConfigurations` 파라미터를 설정하고 조정할 속성의 이름을 지정하세요.

다음 JSON 예제는 인덱스의 “부서”라는 속성에 대한 조정을 재정의하는 `DocumentRelevanceOverrideConfigurations` 객체를 보여줍니다.

```
"DocumentRelevanceOverrideConfigurations" : [
    "Name": "department",
    "Relevance": {
        "Importance": 1,
        "ValueImportanceMap": {
            "IT": 3,
            "HR": 7
        }
    }
]
```

응답 정렬

Amazon Kendra 정렬 속성 또는 필드를 쿼리에서 반환된 문서 기준의 일부로 사용합니다. 예를 들어, “_created_at”로 정렬된 쿼리에서 반환된 결과에는 “_version”으로 정렬된 쿼리와 동일한 결과가 포함되지 않을 수 있습니다.

기본적으로 쿼리 응답은 응답의 각 결과를 Amazon Kendra 결정하는 관련성 점수를 기준으로 정렬됩니다. 정렬 순서를 변경하려면 문서 속성을 정렬 가능하게 만든 다음 해당 속성을 사용하여 응답을 Amazon Kendra 정렬하도록 구성하십시오.

다음 유형의 기본 제공 또는 사용자 지정 속성/필드에 대한 결과를 정렬할 수 있습니다.

- 날짜 값
- long 값
- 문자열 값

다음 유형의 속성은 정렬할 수 없습니다.

- 문자열 목록 값

각 쿼리에서 하나 이상의 문서 속성을 기준으로 정렬할 수 있습니다. 쿼리는 100개의 결과를 반환합니다. 정렬 속성이 설정된 문서가 100개 미만인 경우, 정렬 속성 값이 없는 문서는 쿼리와의 관련성에 따라 정렬되어 결과 끝에 반환됩니다.

문서 결과 정렬하기 (AWS SDK)

1. UpdateIndexAPI를 사용하여 속성을 정렬 가능하게 만들려면 `Sortable` 파라미터를 로 설정합니다. `true` 다음 JSON 예제는 인덱스에 “Department”라는 속성을 추가하고 정렬 가능하게 만드는 데 `DocumentMetadataConfigurationUpdates`를 사용합니다.

```
"DocumentMetadataConfigurationUpdates": [
   {
       "Name": "Department",
       "Type": "STRING_VALUE",
       "Search": {
           "Sortable": "true"
       }
   }
]
```

2. 쿼리에 정렬 가능한 속성 하나를 사용하려면 Query API의 `SortingConfiguration` 파라미터를 설정합니다. 정렬할 속성의 이름과 응답을 오름차순 또는 내림차순으로 정렬할지 여부를 지정합니다.

   다음 JSON 예제는 “Department” 속성을 기준으로 쿼리 결과를 오름차순으로 정렬하는 데 사용하는 `SortingConfiguration` 파라미터를 보여줍니다.

```
   "SortingConfiguration": {
      "DocumentAttributeKey": "Department",
      "SortOrder": "ASC"
   }
```

3. 쿼리에 정렬 가능한 속성을 두 개 이상 사용하려면 Query API의 SortingConfigurations 파라미터를 설정합니다. Amazon Kendra 가 결과를 정렬해야 하는 필드를 3개까지 설정할 수 있습니다. 결과를 오름차순 또는 내림차순으로 정렬할지 지정할 수도 있습니다. 정렬 필드 할당량을 늘릴 수 있습니다.

   정렬 구성을 제공하지 않으면 결과를 Amazon Kendra 결정하는 관련성에 따라 결과가 정렬됩니다. 결과 정렬이 같을 경우 관련성에 따라 결과가 정렬됩니다.

   다음 JSON 예제는 "Name" 및 "Price" 속성을 기준으로 쿼리 결과를 오름차순으로 정렬하는 데 사용하는 SortingConfigurations 파라미터를 보여줍니다.

```
"CollapseConfiguration" : {
    "DocumentAttributeKey": "Name",
    "SortingConfigurations": [
        {
            "DocumentAttributeKey": "Price",
            "SortOrder": "ASC"
        }
    ],
    "MissingAttributeKeyStrategy": "IGNORE"
}
```

문서 결과 정렬 (콘솔)

> **Note**
>
> 다중 속성 정렬은 현재 AWS Management Console에서 지원되지 않습니다.

1. 콘솔에서 속성을 정렬 가능하게 만들려면 속성 정의에서 정렬 가능을 선택합니다. 속성을 생성할 때 속성을 정렬 가능하게 만들거나 나중에 수정할 수 있습니다.
2. 콘솔에서 쿼리 응답을 정렬하려면 정렬 메뉴에서 응답을 정렬할 속성을 선택합니다. 데이터소스 구성 중에 정렬 가능으로 표시된 속성만 목록에 표시됩니다.

쿼리 결과 축소/확대

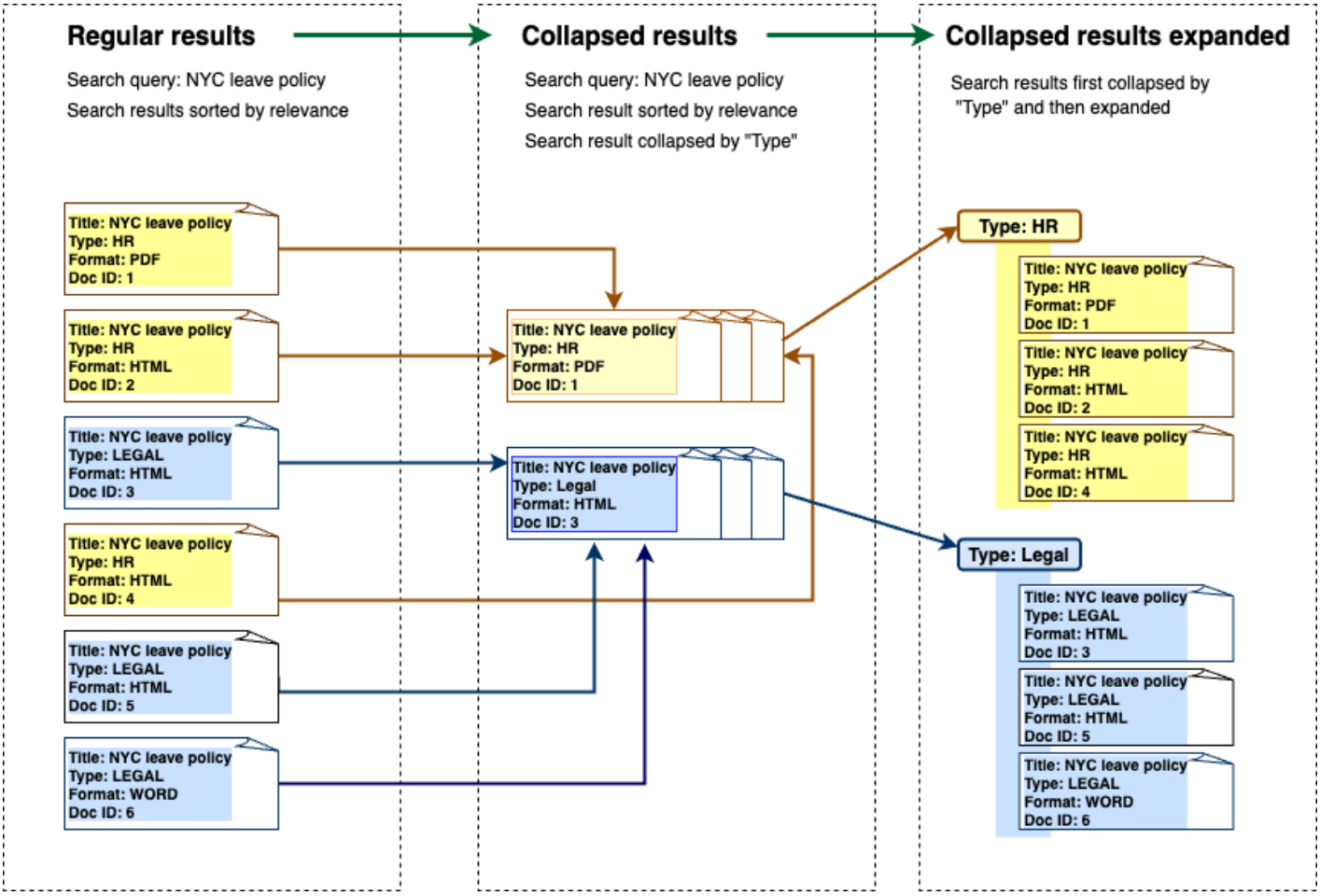
데이터에 Amazon Kendra 연결하면, 및 같은 문서 메타데이터 _document_title 속성을 크롤링하고 _document_id 이러한 속성이나 필드를 사용하여 쿼리 중에 고급 검색 기능을 제공합니다. _created_at

Amazon Kendra의 쿼리 결과 축소 및 확장 기능을 사용하면 공통 문서 속성을 사용하여 검색 결과를 그룹화하고 지정된 기본 문서 아래에 축소되거나 부분적으로 확장된 검색 결과를 표시할 수 있습니다.

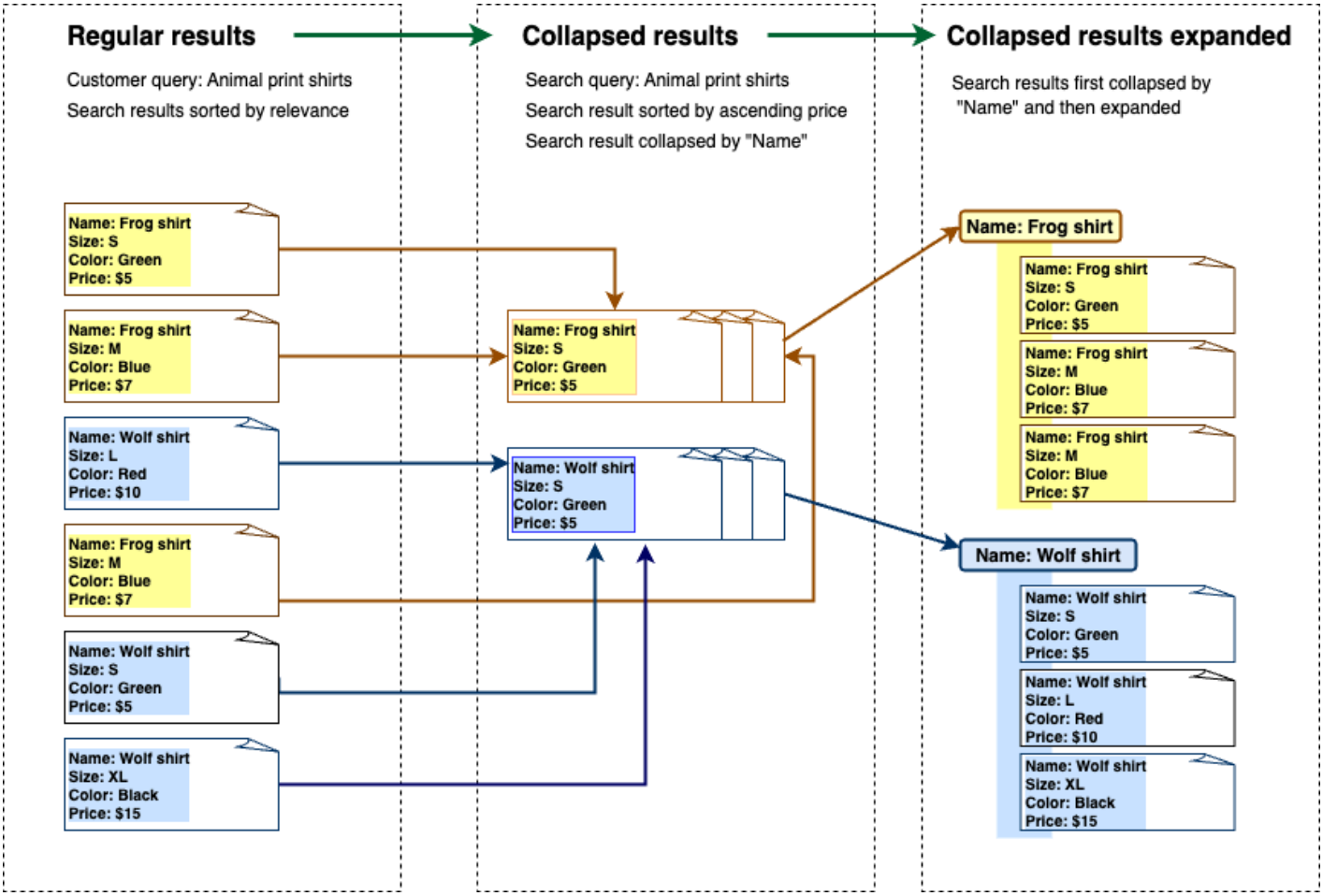
> **Note**
>
> 쿼리 결과 축소 및 확장 기능은 현재 Amazon Kendra API를 통해서만 사용할 수 있습니다.

이는 다음과 같은 검색 상황에서 유용합니다.

- 인덱스 내 문서에는 여러 버전의 콘텐츠가 있습니다. 최종 사용자가 인덱스를 쿼리할 때는 중복 항목이 숨겨지거나 접힌 상태로 가장 관련성이 높은 문서 버전을 보길 원할 것입니다. 예를 들어 색인에 “NYC leave policy”라는 문서의 여러 버전이 포함되어 있는 경우 “유형” 속성/필드를 사용하여 특정 그룹 “HR”과 “Legal”에 맞게 문서를 축소하도록 선택할 수 있습니다.



- 인덱스에는 제품 인벤토리와 같이 한 종류의 품목 또는 객체에 대한 고유 정보로 구성된 여러 문서가 포함되어 있습니다. 항목 정보를 편리하게 캡처하고 정렬하려면 최종 사용자가 항목 또는 객체로 연결된 모든 문서를 하나의 검색 결과로 액세스할 수 있도록 해야 합니다. 아래 예에서 고객이 “애니멀 프린트 셔츠”를 검색하면 이름별로 그룹화되고 가격 오름차순으로 정렬된 결과가 나타납니다.



결과 축소

유사하거나 관련된 문서를 그룹화하려면 축소할 속성을 지정해야 합니다 (예: 문서 축소/그룹화 기준). `_category` 이렇게 하려면 Query API를 호출하고 CollapseConfiguration객체를 사용하여 축소할 대상을 지정하십시오. `DocumentAttributeKey` `DocumentAttributeKey`는 검색 결과를 축소할 필드를 제어합니다. 지원되는 속성 키 필드에는 `String` 및 `Number`이 포함됩니다. `String list` 및 `Date` 유형은 지원되지 않습니다.

정렬 순서를 사용하여 기본 문서 선택

축소된 그룹에 표시되도록 기본 문서를 구성하려면 의 `SortingConfigurations` 매개 변수를 사용합니다. CollapseConfiguration 예를 들어, 문서의 최신 버전을 가져오려면 축소된 각 그룹을 기준으로 정렬합니다. `_version` `SortingConfigurations`를 사용하여 정렬 기준으로 삼을 속성/필드를 최대 3개까지 지정하고 각 속성/필드의 정렬 순서를 지정할 수 있습니다. 정렬 속성 수에 대한 할당량 증가를 요청할 수 있습니다.

기본적으로 는 응답의 각 결과에 대해 결정된 관련성 점수를 Amazon Kendra 기준으로 쿼리 응답을 정렬합니다. 기본 정렬 순서를 변경하려면 문서 속성을 정렬 가능하게 만든 다음 이러한 속성을 사용하여 응답을 Amazon Kendra 정렬하도록 구성하십시오. 자세한 내용은 응답 정렬을 참조하세요.

누락된 문서 키 전략

문서에 축소 속성 값이 없는 경우 Amazon Kendra 는 다음 세 가지 사용자 지정 옵션을 제공합니다.

- 한 그룹에서 값이 null이거나 누락된 모든 문서를 COLLAPSE하기로 선택합니다. 기본 구성입니다.
- 값이 null이거나 누락된 문서를 IGNORE하기로 선택합니다. 무시된 문서는 쿼리 결과에 표시되지 않습니다.
- 값이 null이거나 누락된 각 문서를 하나의 고유한 그룹으로 EXPAND하기로 선택합니다.

결과 확대

개체의 Expand 매개 변수를 사용하여 축소된 검색 결과 그룹을 확장할지 여부를 선택할 수 있습니다. CollapseConfiguration 확장된 결과에는 그룹의 기본 문서를 선택할 때 사용한 것과 동일한 정렬 순서가 유지됩니다.

축소된 검색 결과 그룹의 수를 확장하도록 구성하려면 개체의 MaxResultItemstoExpand 매개 변수를 사용합니다. ExpandConfiguration 예를 들어 이 값을 10으로 설정하면 100개의 결과 그룹 중 처음 10개만 확장 기능을 사용할 수 있습니다.

축소된 기본 문서당 표시할 확장된 결과 수를 구성하려면 MaxExpandResultsPerItem 파라미터를 사용하세요. 예를 들어 이 값을 3으로 설정하면 축소된 그룹당 최대 3개의 결과가 표시됩니다.

다른 Amazon Kendra 기능과의 상호 작용

- 결과를 축소 및 확장해도 패싯 수가 변경되거나 표시된 총 결과 수에는 영향을 주지 않습니다.
- Amazon Kendra 추천 검색 결과는 사용자가 구성한 축소 필드와 필드 값이 같더라도 축소되지 않습니다.
- 결과 축소 및 확장은 유형의 DOCUMENT 결과에만 적용됩니다.

검색 관련성 조정

Amazon Kendra 쿼리는 관련성을 기준으로 순위가 매겨진 검색 결과를 생성합니다. 인덱스의 검색 가능한 필드 또는 속성이 모두 이 순위에 영향을 미칩니다.

관련성 조정을 통해 검색 관련성에 대한 필드 또는 속성의 영향을 수정할 수 있습니다. 검색 관련성 조정은 인덱스에 대한 조정 구성을 설정하는 인덱스 수준에서 수동으로 수행하거나 인덱스 수준에서 설정된 구성을 재정의하여 쿼리 수준에서 수행할 수 있습니다.

관련성 조정 기능을 사용하면 쿼리에 필드 또는 속성과 일치하는 용어가 포함되어 있는 경우 응답에서 결과가 부스트됩니다. 또한 일치하는 항목이 있을 때 문서가 받는 부스트 정도를 지정할 수 있습니다. 관련성 Amazon Kendra 조정으로 인해 쿼리 응답에 문서가 포함되는 것은 아니며 문서의 관련성을 결정하는 데 Amazon Kendra 사용되는 요소 중 하나일 뿐입니다.

인덱스의 특정 필드 또는 속성을 높여 특정 응답에 더 높은 중요도를 할당할 수 있습니다. 예를 들어 누군가 “re:Invent는 언제인가요?” 라고 검색하는 경우를 예로 들 수 있습니다. 현장에서 문서 최신성의 관련성을 높일 수 있습니다. `_last_update_at` 또는 연구 보고서 인덱스에서 '소스' 필드의 특정 데이터 소스를 부스트할 수 있습니다.

포럼 및 기타 지원 지식 기반에서 흔히 볼 수 있는 방법인 투표 또는 조회수를 기반으로 문서를 부스트할 수도 있습니다. 예를 들어 여러 부스트를 조합하여 최근에 본 문서뿐만 아니라 더 많이 본 문서도 부스트할 수 있습니다.

`Importance` 파라미터를 사용하여 문서가 받는 부스트 양을 설정합니다. `Importance` 값이 높을수록 필드 또는 속성이 문서의 관련성을 더 높게 부스트합니다. 인덱스를 조정하거나 쿼리 수준에서 조정할 때는 원하는 효과를 얻을 때까지 `Importance` 파라미터 값을 조금씩 늘리세요. 검색 결과가 개선되고 있는지 확인하려면 검색을 수행하고 결과를 이전 쿼리와 비교하세요.

날짜, 숫자 또는 문자열 속성을 지정하여 인덱스를 조정하거나 쿼리 수준에서 조정할 수 있습니다. `StringList` 유형의 필드나 속성은 인덱스 수준에서만 조정할 수 있습니다. 각 필드 또는 속성에는 결과를 부스트하는 시기에 대한 특정 기준이 있습니다.

- 날짜 필드 또는 속성 - 날짜 필드, `Duration`, `Freshness`, `RankOrder`에는 세 가지 특정 기준이 있습니다.
  - `Duration`는 부스트가 적용되는 기간을 지정합니다. 예를 들어 기간을 86400초(즉, 하루)로 설정하면 하루가 지난 후 부스트 효과가 줄어들기 시작합니다. 중요도가 높을수록 부스트 효과가 더 빨리 감소합니다.

  - Freshness는 필드 또는 속성에 적용될 때 문서의 최신성을 결정합니다. 만든 날짜 또는 마지막 업데이트 날짜 필드에 Freshness를 적용하면 최근에 만든 문서나 마지막으로 업데이트된 문서가 이전 문서보다 "최신" 상태로 간주됩니다. 예를 들어, 문서 1이 11월 14일에 작성되었고 문서 2가 11월 5일에 작성된 경우 문서 1은 문서 2보다 "최신"입니다. 문서 1이 11월 14일에 마지막으로 업데이트되었고 문서 2가 11월 20일에 마지막으로 업데이트되었다면 문서 2가 문서 1보다 "최신"입니다. 문서가 최신 버전일수록 이 부스트가 더 많이 적용됩니다. 인덱스에는 Freshness 필드가 한 개만 있을 수 있습니다.
  - RankOrder는 부스트를 오름차순 또는 내림차순으로 적용합니다. ASCENDING를 지정하는 경우 이후 날짜가 우선합니다. DESCENDING를 지정하는 경우 이전 날짜가 우선합니다.
- 숫자 필드 또는 속성 - 숫자 필드 또는 속성의 경우 필드 또는 속성의 관련성을 결정할 때 Amazon Kendra 사용해야 하는 순위 순서를 지정할 수 있습니다. ASCENDING를 지정하는 경우 더 높은 숫자가 우선합니다. DESCENDING를 지정하는 경우 더 낮은 숫자가 우선합니다.
- 문자열 필드 또는 속성 - 문자열 필드 또는 속성의 경우 필드의 범주를 만들어 각 범주에 다르게 적용할 수 있습니다. 예를 들어, "부서"라는 필드 또는 속성을 부스트하면 "HR"의 문서와 "법률"의 문서를 다르게 부스트할 수 있습니다. String 유형의 필드 또는 속성을 부스트할 수 있습니다. 인덱스 수준에서만 StringList 필드를 부스트할 수 있습니다.

인덱스 수준에서의 관련성 조정

콘솔을 사용하여 인덱스 세부 정보에서 조정을 설정하거나 API를 사용하여 인덱스 수준에서 필드 또는 속성의 관련성을 조정합니다. UpdateIndex

다음 예제에서는 _last_updated_at 필드를 문서의 Freshness 필드로 설정합니다.

```
"DocumentMetadataConfigurationUpdates" : [
    {
        "Name": "_last_updated_at",
        "Type": "DATE_VALUE",
        "Relevance": {
            "Freshness": TRUE,
            "Importance": 2
        }
    }
]
```

다음 예제에서는 "부서" 필드의 범주별로 서로 다른 중요도를 적용합니다.

```
"DocumentMetadataConfigurationUpdates" : [
```

```
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 3,
                "Legal": 1
            }
        }
    }
]
```

쿼리 수준에서의 관련성 조정

Query API를 사용하여 쿼리 수준에서 필드 또는 속성의 관련성을 조정합니다.

콘솔에서는 쿼리 수준의 관련성 조정이 지원되지 않습니다.

쿼리 수준에서 조정하면 각 테스트에 대해 인덱스의 조정 구성을 수동으로 업데이트할 필요가 없으므로 관련성 조정 테스트 프로세스의 속도를 높일 수 있습니다. 쿼리에 조정 구성을 전달하여 문서의 관련성을 조정할 수 있습니다. 그러면 다양한 구성에서 얻은 다양한 결과를 볼 수 있습니다. 쿼리에 전달된 구성은 인덱스 수준에서 설정된 구성을 재정의합니다.

다음 예제는 위 예와 같이 인덱스 수준에서 설정된 "부서" 필드 및 각 부서 범주에 적용되는 중요도를 재정의합니다. 사용자가 검색어를 입력하면 "부서" 필드의 중요도가 상당히 높으며 법무 부서가 HR 부서보다 중요도가 높습니다.

```
"DocumentRelevanceOverrideConfigurations" : [
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 2,
                "Legal": 8
            }
        }
    }
]
```

검색 분석을 통한 인사이트 확보

Amazon Kendra 검색 분석을 사용하여 검색 애플리케이션이 사용자가 정보를 찾도록 지원하는 데 성공했는지 또는 실패했는지에 대한 통찰력을 얻을 수 있습니다.

Amazon Kendra 애널리틱스는 사용자가 검색 애플리케이션과 상호 작용하는 방식과 검색 애플리케이션 구성이 얼마나 효과적인지에 대한 스냅샷을 제공합니다. GetSnapshotsAPI를 사용하거나 콘솔의 탐색 패널에서 분석을 선택하여 지표 데이터를 볼 수 있습니다.

GetSnapshots에서 생성된 데이터를 사용자 지정 대시보드에서 렌더링할 수 있습니다. 또는 콘솔에 제공된 지표 대시보드(시각적 그래프 포함)를 사용할 수도 있습니다. 시각적 대시보드를 사용하면 시간 경과에 따른 사용자 행동의 추세나 패턴을 찾아보거나 검색 애플리케이션 구성의 문제를 찾아낼 수 있습니다. 예를 들어 일일 쿼리 수가 일정하고 꾸준히 증가하는 선 그래프는 채택률과 사용량이 증가했음을 의미할 수 있습니다. 반면, 급격한 하락은 조사해야 할 문제가 있음을 의미할 수 있습니다.

지표를 사용하여 서로 다른 데이터 지점을 연결하면 사용자가 정보를 쿼리하는 방식과 관련된 문제를 해결하거나 비즈니스 기회를 발견할 수 있습니다. 예를 들어 'AI는 어떻게 작동합니까?'라는 문서는 검색 결과에서 가장 많이 클릭된 문서이며, 가장 많이 검색된 쿼리는 '기계 학습은 어떻게 작동합니까?'입니다. 이를 통해 사용자가 선호하는 용어와 언어를 알 수 있습니다. 이러한 용어를 문서에 통합하거나 이러한 용어에 대한 사용자 지정 동의어를 사용하면 사용자가 문서를 더 쉽게 검색할 수 있도록 할 수 있습니다.

검색 지표

검색 애플리케이션의 성능이나 사용자가 검색하는 정보를 분석하기 위한 10가지 지표가 있습니다. 지표 데이터를 검색하려면 GetSnapshots 호출 시 검색하려는 지표 데이터의 문자열 이름을 지정합니다.

또한 지표 데이터를 보려면 시간 간격이나 기간을 제공해야 합니다. 시간 간격은 색인의 시간대를 사용합니다. 다음 시간 창에서 데이터를 볼 수 있습니다.

- THIS_WEEK: 일요일에 시작하여 현재 날짜의 전날에 끝나는 이번 주.
- ONE_WEEK_AGO: 이전 주, 일요일에 시작하여 다음 토요일에 끝납니다.
- TWO_WEEKS_AGO: 이전 주의 전 주, 일요일에 시작하여 다음 토요일에 끝납니다.
- THIS_MONTH: 이번 달, 그 달의 첫째 날에 시작하여 현재 날짜 전날에 끝납니다.
- ONE_MONTH_AGO: 이전 달, 그 달의 첫째 날에 시작하여 마지막 날에 끝납니다.
- TWO_MONTHS_AGO: 이전 달의 전 달, 그 달의 첫째 날에 시작하여 마지막 날에 끝납니다.

콘솔에서 지원되는 기간은 이번 주, 이전 주, 이번 달, 이전 달입니다.

클릭률

검색 결과에서 문서로 클릭을 유도한 쿼리의 비율입니다. 이를 통해 검색 애플리케이션 구성이 사용자가 쿼리와 관련된 정보를 찾는 데 도움이 되는지 파악할 수 있습니다. 즉각적인 답변을 반환하는 쿼리의 경우 사용자는 문서를 클릭하여 자세한 내용을 확인할 필요가 없을 수도 있습니다. 자세한 설명은 the section called “즉각적인 응답률” 섹션을 참조하세요. 클릭스루 피드백이 수집되도록 SubmitFeedback하려면 전화를 걸어야 합니다.

GetSnapshots API를 사용하여 클릭률에 대한 데이터를 검색하려면 metricType를 AGG_QUERY_DOC_METRICS로 지정합니다. 탐색 패널에서 분석을 선택하여 콘솔에서 이 지표를 볼 수도 있습니다.

제로 클릭률

검색 결과에서 클릭으로 이어지지 않은 쿼리의 비율입니다. 이를 통해 관련 없는 검색 결과를 제공하는 콘텐츠 격차를 파악할 수 있습니다. 즉각적인 답변을 반환하는 쿼리의 경우 사용자는 문서를 클릭하여 자세한 내용을 확인할 필요가 없을 수도 있습니다. 자세한 설명은 the section called “즉각적인 응답률” 섹션을 참조하세요. 또한 조정 구성과 같은 검색 설정이 검색 결과에 문서가 반환되는 방식에 영향을 미칠 수 있습니다.

GetSnapshots API를 사용하여 제로 클릭률에 대한 데이터를 검색하려면 metricType를 AGG_QUERY_DOC_METRICS로 지정합니다. 탐색 패널에서 분석을 선택하여 콘솔에서 이 지표를 볼 수도 있습니다.

제로 검색 결과 비율

검색 결과가 없는 쿼리의 비율입니다. 이를 통해 관련된 검색 결과가 없는 콘텐츠 격차를 파악할 수 있습니다.

GetSnapshots API를 사용하여 제로 검색 결과 비율에 대한 데이터를 검색하려면 metricType를 AGG_QUERY_DOC_METRICS로 지정합니다. 탐색 패널에서 분석을 선택하여 콘솔에서 이 지표를 볼 수도 있습니다.

즉각적인 응답률

즉각적인 답변 또는 FAQ가 반환된 쿼리의 비율입니다. 이를 통해 정보 제공에서 즉각적 답변이 어떤 역할을 하는지 이해할 수 있습니다.

GetSnapshots API를 사용하여 즉각적인 응답률에 대한 데이터를 검색하려면 metricType를 AGG_QUERY_DOC_METRICS로 지정합니다. 탐색 패널에서 분석을 선택하여 콘솔에서 이 지표를 볼 수도 있습니다.

상위 쿼리

사용자가 검색한 상위 100개 쿼리. 이를 통해 인기 있는 검색어와 사용자가 가장 관심을 갖는 정보의 종류를 파악할 수 있습니다.

지표에는 쿼리 검색 횟수, 문서에 대한 클릭률, 문서로 클릭되지 않은 건의 비율, 쿼리 검색 결과의 평균 클릭 깊이, 쿼리에 대한 즉각적인 답변 비율, 쿼리의 처음 10개 검색 결과에 대한 평균 신뢰도 등이 포함됩니다.

GetSnapshots API를 사용하여 상위 쿼리에 대한 데이터를 검색하려면 metricType를 QUERIES_BY_COUNT로 지정합니다. 콘솔의 탐색 패널에서 분석을 선택한 다음 쿼리 목록에서 상위 쿼리를 선택하면 콘솔에서 이 지표도 볼 수 있습니다.

클릭 없는 상위 쿼리

검색 결과에서 클릭으로 이어지지 않은 상위 100개 쿼리입니다. 이를 통해 일부 검색어와 관련된 문서가 부족하거나 검색 애플리케이션 구성에서 관련 없는 검색 결과가 반환되는 등의 콘텐츠 격차를 파악할 수 있습니다. 즉각적인 답변을 반환하는 쿼리의 경우 사용자는 문서를 클릭하여 자세한 내용을 확인할 필요가 없을 수도 있습니다. 자세한 설명은 the section called “즉각적인 응답률” 섹션을 참조하세요.

지표에는 쿼리가 클릭으로 이어지지 않은 횟수, 쿼리에 대한 제로 클릭의 비율, 쿼리에 대한 즉각적인 답변 비율, 쿼리의 처음 10개 검색 결과에 대한 평균 신뢰도가 포함됩니다.

GetSnapshots API를 사용하여 클릭 없는 상위 쿼리에 대한 데이터를 검색하려면 metricType를 QUERIES_BY_ZERO_CLICK_RATE로 지정합니다. 콘솔의 탐색 패널에서 분석을 선택한 다음 쿼리 목록에서 상위 제로 클릭 쿼리를 선택하면 콘솔에서 이 지표도 볼 수 있습니다.

검색 결과가 없는 상위 쿼리

검색 결과가 없는 상위 100개 쿼리입니다. 이를 통해 일부 쿼리와 관련된 문서가 없는 콘텐츠의 격차를 파악할 수 있습니다. 또는 사용자가 특수 용어로 쿼리하여 검색 결과가 나오지 않을 수 있으며, 이를 처리하기 위해 사용자 지정 동의어를 만들라는 메시지가 표시될 수 있습니다.

지표에는 쿼리의 검색 결과가 없었던 횟수, 검색 결과가 없는 쿼리의 비율, 모든 쿼리 수와 비교한 쿼리 검색 횟수의 비율 등이 포함됩니다.

GetSnapshots API를 사용하여 검색 결과가 없는 상위 쿼리에 대한 데이터를 검색하려면 metricType를 QUERIES_BY_ZERO_RESULT_RATE로 지정합니다. 콘솔의 탐색 패널에서 분석을 선택한 다음 쿼리 목록에서 상위 제로 결과 쿼리를 선택하면 콘솔에서 이 지표도 볼 수 있습니다.

가장 많이 클릭된 문서

검색 결과에서 가장 많이 클릭된 상위 100개 문서입니다. 이를 통해 사용자가 정보를 쿼리할 때 가장 관련성이 높은 문서 또는 검색 결과를 파악할 수 있습니다.

지표에는 문서를 클릭한 횟수, 문서가 사용자로부터 받은 좋아요 수(엄지 손가락 위로), 문서가 사용자로부터 받은 싫어요 수(엄지 손가락 아래로)가 포함됩니다.

GetSnapshots API를 사용하여 가장 많이 클릭된 문서에 대한 데이터를 검색하려면 metricType를 DOCS_BY_CLICK_COUNT로 지정합니다. 콘솔의 탐색 패널에서 분석을 선택한 다음 쿼리 목록에서 가장 많이 클릭된 문서를 선택하면 콘솔에서 이 지표도 볼 수 있습니다.

총 쿼리 수

사용자가 검색한 총 쿼리 수입니다. 이를 통해 사용자가 검색 애플리케이션에 얼마나 참여하고 있는지 파악할 수 있습니다.

GetSnapshots API를 사용하여 총 쿼리 수에 대한 데이터를 검색하려면 metricType를 AGG_QUERY_DOC_METRICS로 지정합니다. 탐색 패널에서 분석을 선택하여 콘솔에서 이 지표를 볼 수도 있습니다.

총 문서

인덱스의 총 문서 수입니다. 이를 통해 인덱스 크기를 총 쿼리 수와 비교하여 쿼리 볼륨에 적합한 문서 수가 있는지 확인할 수 있습니다.

GetSnapshots API를 사용하여 총 문서 수에 대한 데이터를 검색하려면 metricType를 AGG_QUERY_DOC_METRICS로 지정합니다. 탐색 패널에서 분석을 선택하여 콘솔에서 이 지표를 볼 수도 있습니다.

지표 데이터 검색 예제

다음 코드는 지난 달의 상위 쿼리에 대한 데이터를 검색하는 예제입니다.

Console

지난 달의 상위 쿼리를 검색하려면

1. 왼쪽 탐색 창의 인덱스에서 인덱스를 선택한 다음 분석을 선택합니다.
2. 분석 페이지에서 이번 주 버튼을 선택하여 데이터 검색 기간을 지난 달로 변경합니다.
3. 분석 페이지의 쿼리 목록에서 상위 쿼리를 선택합니다.

CLI

지난 달의 상위 쿼리를 검색하려면

```
aws kendra get-snapshots \
--index-id index-id \
--interval "ONE_MONTH_AGO" \
--metric-type "QUERIES_BY_COUNT"
```

Python

지난 달의 상위 쿼리를 검색하려면

```
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

지난 달의 상위 쿼리를 검색하려면

```
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()))
```

지표에서 실행 가능한 인사이트로

실행 가능한 인사이트는 원시 데이터에서 추출한 의미 있는 정보로, 행동이나 결정을 안내하는 지침으로 사용됩니다. 지표에서 의미를 추출하고 이를 사용하여 실행 가능한 인사이트를 이끌어내려면 지표를 분리해서 보는 것뿐만 아니라 여러 지표를 연결하여 보는 것도 중요합니다.

예를 들어 클릭 없는 상위 쿼리는 '현재 사용 가능한 지역은 어디입니까?'입니다. 하지만 즉각적인 응답률도 100퍼센트입니다. 따라서 사용자는 검색 결과나 사용 가능한 리전에 대한 정보를 제공하는 문서를 클릭하지 않고도 이 질문에 대한 답을 얻을 수 있습니다. 제로 클릭만 고려하면 전체 스토리를 이해할 수 없으며 검색 애플리케이션 구성이 이 쿼리를 성공적으로 처리했는지에 대해 잘못된 결론을 내릴 수 있습니다.

실행 가능한 인사이트의 또 한 가지 예시는 비즈니스 기회를 발견하는 것입니다. 기업은 종종 검색 지표를 분석하여 고객을 유치할 기회를 찾습니다. 가장 많이 클릭된 문서는 '사용 가능한 지역'입니다. 이 외에도 가장 많이 검색된 쿼리 대부분은 오세아니아 지역의 제품 가용성에 대한 질문과 관련이 있습니

다. 100%의 즉각적인 응답률과 함께, 답변의 일부로 사용 가능한 리전에 대한 추가 정보가 높은 클릭률을 기록했습니다. 이는 해당 리전에서 귀사의 제품 또는 서비스에 대한 관심과 수요가 있음을 시사합니다.

검색 분석 시각화 및 보고

시간 경과에 따른 추세나 패턴을 시각화하고 살펴볼 수 있는 트렌드 데이터가 포함된 다섯 가지 지표가 있습니다. 콘솔을 사용하는 경우 트렌드 데이터의 그래프가 제공됩니다. API를 사용하는 경우 트렌드 데이터를 검색하여 자체 그래프나 시각화를 만들 수 있습니다. 콘솔에 있는 대부분의 그래프는 선택한 기간 동안의 일일 데이터 포인트를 표시합니다.

콘솔은 보고 싶은 그래프와 상위 목록을 선택할 수 있는 지표 대시보드를 제공합니다. 분석 홈 페이지에서 내보내기를 선택하면 대시보드에 표시된 지표를 CSV 형식으로 내보낼 수 있습니다. 비즈니스 문서나 프레젠테이션에 이러한 보고서를 포함할 수 있습니다.

다음 지표를 볼 수 있습니다.

총 쿼리 그래프

하루에 실행된 쿼리 수를 나타내는 선 그래프입니다. 그래프를 통해 일일 사용자 참여 패턴을 시각화할 수 있습니다. 몇 가지 예로는 사용자 참여가 꾸준히 증가 또는 감소하거나, 검색 애플리케이션의 다운 또는 웹사이트 관련 문제로 인해 쿼리가 0으로 급감하는 경우를 들 수 있습니다.

API를 사용하는 경우 TREND_QUERY_DOC_METRICS 지정을 통해 이러한 데이터를 검색할 수 있습니다. 데이터를 사용하여 그래프를 직접 만들거나 콘솔에 제공된 그래프를 사용할 수 있습니다.

클릭률 그래프

일일 클릭률의 비율을 나타낸 선 그래프입니다. 그래프를 통해 일일 클릭률의 패턴을 시각화할 수 있습니다. 클릭률의 꾸준한 증가 또는 감소, 클릭률 증가에 영향을 미칠 수 있는 즉각적인 답변의 감소 등을 예로 들 수 있습니다.

API를 사용하는 경우 TREND_QUERY_DOC_METRICS 지정을 통해 이러한 데이터를 검색할 수 있습니다. 데이터를 사용하여 그래프를 직접 만들거나 콘솔에 제공된 그래프를 사용할 수 있습니다.

제로 클릭률 그래프

일일 제로 클릭의 비율을 나타낸 선 그래프입니다. 그래프를 통해 일일 제로 클릭률의 패턴을 시각화할 수 있습니다. 제로 클릭률의 꾸준한 증가 또는 감소, 제로 클릭 증가에 영향을 미칠 수 있는 즉각적인 답변의 증가 등을 예로 들 수 있습니다.

API를 사용하는 경우 TREND_QUERY_DOC_METRICS 지정을 통해 이러한 데이터를 검색할 수 있습니다. 데이터를 사용하여 그래프를 직접 만들거나 콘솔에 제공된 그래프를 사용할 수 있습니다.

제로 검색 결과 비율 그래프

일일 제로 검색 결과의 비율을 나타낸 선 그래프입니다. 그래프를 통해 일일 제로 검색 결과 비율의 패턴을 시각화할 수 있습니다. 검색 결과가 없는 비율이 꾸준히 증가하거나 감소하는 경우, 검색 결과 제로의 증가에 영향을 미칠 수 있는 인덱스의 문서 수 급감 등을 예로 들 수 있습니다.

API를 사용하는 경우 TREND_QUERY_DOC_METRICS 지정을 통해 이러한 데이터를 검색할 수 있습니다. 데이터를 사용하여 그래프를 직접 만들거나 콘솔에 제공된 그래프를 사용할 수 있습니다.

즉각적인 응답률 그래프

즉각적인 답변 또는 FAQ가 반환된 쿼리의 비율을 나타낸 선 그래프입니다. 그래프를 통해 일일 즉각적인 응답률의 패턴을 시각화할 수 있습니다. 몇 가지 예로는 질문-답변 유형 쿼리의 꾸준한 증가 또는 감소, 즉각적인 답변 증가에 영향을 미칠 수 있는 클릭률 감소 등이 있습니다.

API를 사용하는 경우 TREND_QUERY_DOC_METRICS 지정을 통해 이러한 데이터를 검색할 수 있습니다. 데이터를 사용하여 그래프를 직접 만들거나 콘솔에 제공된 그래프를 사용할 수 있습니다.

점진적 학습을 위한 피드백 제출

Amazon Kendra 증분 학습을 사용하여 검색 결과를 개선합니다. 쿼리의 피드백을 활용하는 점진적 학습은 순위 결정 알고리즘을 개선하고 검색 결과를 최적화하여 정확도를 높입니다.

예를 들어 사용자가 "의료 혜택"이라는 문구를 검색하였다고 가정하겠습니다. 사용자가 지속적으로 목록에서 두 번째 결과를 선택하면 시간이 지날수록 Amazon Kendra 는 해당 결과를 첫 번째 검색 결과로 올립니다. 부스트는 시간이 지남에 따라 감소하므로 사용자가 결과 선택을 중단하면 Amazon Kendra 결국에는 해당 결과가 제거되고 대신 더 인기 있는 다른 결과가 표시됩니다. 이렇게 하면 관련성, 연령, 콘텐츠를 기반으로 결과의 Amazon Kendra 우선 순위를 정하는 데 도움이 됩니다.

모든 인덱스와 지원되는 모든 문서 유형에 대해 점진적 학습이 활성화됩니다.

Amazon Kendra 피드백을 제공하는 즉시 학습을 시작하지만 피드백 결과를 확인하는 데 24시간 이상 걸릴 수 있습니다. Amazon Kendra 에서는 피드백을 제출할 수 있는 세 가지 방법, 즉 AWS 콘솔, 검색 결과 페이지에 포함할 수 있는 JavaScript 라이브러리, 사용할 수 있는 API를 제공합니다.

Amazon Kendra 두 가지 유형의 사용자 피드백을 수락합니다.

- 클릭 - 사용자가 선택한 쿼리 결과에 대한 정보. 피드백에는 검색 결과를 선택한 날짜 및 시간의 결과 ID와 Unix 타임스탬프가 포함됩니다.

  클릭 피드백을 제출하려면 애플리케이션이 사용자의 활동으로부터 클릭 정보를 수집한 다음 해당 정보를 Amazon Kendra에 제출해야 합니다. 콘솔, JavaScript 라이브러리 및 Amazon Kendra API를 사용하여 클릭 정보를 수집할 수 있습니다.
- 관련성 - 사용자가 일반적으로 제공하는 검색 결과의 관련성에 대한 정보입니다. 피드백에는 결과 ID와 관련성 지표(`RELEVANT` 또는 `NOT_RELEVANT`)가 포함됩니다. 사용자가 관련성 정보를 결정합니다.

  관련성 피드백을 제출하려면 사용자가 쿼리 결과에 적합한 관련성을 선택한 다음 해당 정보를 Amazon Kendra에 제출할 수 있는 피드백 메커니즘을 애플리케이션이 제공해야 합니다. 콘솔과 Amazon Kendra API로만 관련성 정보를 수집할 수 있습니다.

피드백은 인덱스가 활성화된 동안 사용됩니다. 피드백은 제출된 인덱스에만 영향을 미치며, 인덱스 전체나 다른 계정에는 사용할 수 없습니다.

Amazon Kendra 색인을 쿼리할 때 추가 사용자 컨텍스트를 제공해야 합니다. 사용자 컨텍스트를 제공하면 Amazon Kendra 피드백이 단일 사용자에 의해 제공되는지 여러 사용자가 제공하는지 구분하고 그에 따라 검색 결과를 조정할 수 있습니다.

사용자 컨텍스트를 제공하면 쿼리에 대한 피드백이 컨텍스트에 제공된 특정 사용자와 연결됩니다. 사용자 컨텍스트를 지정하지 않는 경우 쿼리를 그룹화하고 집계하는 데 사용되는 방문자 ID를 제공할 수 있습니다.

사용자 컨텍스트 또는 방문자 ID를 제공하지 않는 경우 피드백은 익명으로 처리되며 다른 익명 피드백과 집계됩니다.

다음 코드는 사용자 컨텍스트를 토큰 또는 방문자 ID로 포함하는 방법을 보여줍니다.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    UserToken = {
        Token = "token"
    })

    OR

    response = kendra.query(
    QueryText = query,
    IndexId = index,
    VisitorId = "visitor-id")
```

웹 애플리케이션의 경우 쿠키, 위치 또는 브라우저 사용자를 사용하여 각 사용자에 대한 방문자 ID를 생성할 수 있습니다.

쿼리 볼륨이 가장 큰 헤드 쿼리의 경우, 클릭률 피드백을 제공하면 전반적인 정확도를 높이는 데 충분한 정보를 얻을 수 있습니다. 드문 테일 쿼리의 경우, 주제 전문가가 관련 있는 피드백과 관련 없는 피드백을 모두 제출하여 해당 쿼리의 정확도를 높여야 합니다.

콘솔 외에도 JavaScript 라이브러리 또는 SubmitFeedbackAPI라는 두 가지 방법 중 하나를 사용할 수 있습니다. 피드백을 수집하는 방법은 한 가지만 사용해야 합니다. 최상의 결과를 얻으려면 쿼리 후 24시간 이내에 피드백을 제출해야 합니다.

주제

- Amazon Kendra JavaScript 라이브러리를 사용하여 피드백 제출
- Amazon Kendra API를 사용하여 피드백 제출

Amazon Kendra JavaScript 라이브러리를 사용하여 피드백 제출

Amazon Kendra 검색 결과 페이지에 클릭 피드백을 추가하는 데 사용할 수 있는 JavaScript 라이브러리를 제공합니다. 라이브러리를 사용하려면 검색 결과를 표시하는 클라이언트 코드에 스크립트 태그를 삽입한 다음 결과 목록의 각 문서 링크에 정보를 추가합니다. 사용자가 문서를 보기 위한 링크를 선택하면 클릭 정보가 Amazon Kendra로 전송됩니다.

라이브러리는 JavaScript 버전 ES6/ES2015를 지원하는 브라우저에서 작동합니다.

1단계: 검색 애플리케이션에 스크립트 태그 삽입 Amazon Kendra

Amazon Kendra 검색 결과를 렌더링하는 클라이언트 코드에 <script>태그를 삽입하고 JavaScript 라이브러리에 대한 참조를 추가합니다.

```
<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>
```

스크립트는 Amazon Kendra 호스팅된 CDN에서 JavaScript 라이브러리를 비동기적으로 다운로드하고 선택적 매개 변수를 설정할 수 있는 kendraFeedback 라는 글로벌 변수를 초기화합니다.

URL# ### ###### ### 호스팅하는 지역에 따라 다음 표의 식별자로 바꾸십시오.
Amazon Kendra

리전	다운로드 URL	피드백 엔드포인트
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
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

리전	다운로드 URL	피드백 엔드포인트
eu-west-2	https://d1tsrujswld1d1.cloudfront.net/ksf-v1.js	https://qi7mct3x7f.execute-api.eu-west-2.amazonaws.com/prod/submit

예를 들어 인덱스가 미국 동부(버지니아 북부)에 있는 경우 `##### #### URL`은 `https://d2zm0lpns956f8.cloudfront.net/ksf-v1.js`이고 `### #####`는`https://ujxwp5s92h.execute-api.us-east-1.amazonaws.com/prod/submit`입니다.

Amazon Kendra JavaScript 라이브러리에 적용할 수 있는 두 가지 선택적 설정이 있습니다.

- `disableCookies`— 기본적으로 사용자를 고유하게 식별하는 쿠키를 Amazon Kendra 설정합니다. 쿠키를 비활성화하려면 이 옵션을 `true`로 설정합니다.

  ```
  kendraFeedback('disableCookie', 'true | false');
  ```

  `searchDivClassName` - 기본적으로 Amazon Kendra 는 검색 결과 페이지의 모든 링크에서 클릭을 모니터링합니다. 지정된 클래스의 링크만 모니터링하려면 이 이름을 클래스 이름으로 설정합니다.

  ```
  kendraFeedback('searchDivClassName', 'class name');
  ```

2단계: 검색 결과에 피드백 토큰 추가

결과 페이지에서 쿼리 응답의 문서에 대한 링크가 포함된 앵커 태그 또는 바로 위 상위 div 태그에 `data-kendra-token`이라는 HTML 속성을 추가합니다. 예:

```
<a href="document location" data-kendra-token="feedback token value"></a>
OR
<div data-url="document location" data-kendra-token="feedback token value"></div>
```

쿼리 응답에는 `feedbackToken` 필드에 토큰이 포함되어 있습니다. 토큰은 사용자가 선택한 응답을 고유하게 식별합니다. 토큰 값을 `data-kendra-token` 속성에 할당합니다. Amazon Kendra JavaScript 라이브러리는 사용자가 결과를 선택할 때 이 토큰을 찾아 엔드포인트에 피드백으로 제출합니다. Amazon Kendra

Amazon Kendra JavaScript 라이브러리는 피드백 토큰과 기타 메타데이터 (예: 결과 선택 시간, 고유한 방문자 ID) 만 제출합니다.

3단계: 피드백 스크립트 테스트

JavaScript 라이브러리가 올바르게 구성되고 올바른 엔드포인트로 피드백을 보내려면 다음을 수행하십시오. 이 예시에서는 Chrome 브라우저를 사용합니다.

1. 브라우저에서 웹 개발자 도구를 엽니다. Chrome에서는 브라우저 오른쪽 상단의 Chrome 메뉴를 열고 추가 도구 선택한 다음 개발자 도구를 선택합니다.
2. 콘솔 탭에 Amazon Kendra JavaScript 라이브러리와 관련된 오류가 없는지 확인하십시오.
3. 검색하고 원하는 결과를 선택하세요. 개발자 도구의 네트워크 탭에서. 피드백 엔드포인트로 전송된 요청, 결과 토큰, 200 OK 상태가 표시되어야 합니다.

Amazon Kendra API를 사용하여 피드백 제출

Amazon Kendra API를 사용하여 쿼리 피드백을 제출하려면 SubmitFeedbackAPI를 사용하십시오. 쿼리를 식별하려면 쿼리가 적용되는 인덱스의 인덱스 ID와 쿼리 API의 응답으로 반환된 쿼리 ID를 제공합니다.

다음 예에서는 Amazon Kendra API를 사용하여 클릭 및 관련성 피드백을 제출하는 방법을 보여줍니다. `ClickFeedbackItems` 및 `RelevanceFeedbackItems` 배열을 통해 여러 피드백 세트를 제출할 수 있습니다. 이 예시는 단일 클릭과 단일 관련성 피드백 항목을 제출합니다. 피드백 제출에는 현재 시간이 사용됩니다.

검색에 대한 피드백 제출하기 (AWS SDK)

1. 다음 예제 코드를 필수 값과 함께 사용할 수 있습니다.

   a. `index id`—쿼리가 적용되는 인덱스의 ID.

   b. `query id`—피드백을 제공하려는 쿼리.

   c. `result id`—피드백을 제공하려는 쿼리 결과의 ID. 쿼리 응답에는 결과 ID가 포함됩니다.

   d. `relevance value`RELEVANT—(쿼리 결과가 관련이 있음) 또는 NOT_RELEVANT (쿼리 결과가 관련이 없음) 중 하나입니다.

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. 코드를 실행합니다. 피드백이 제출되면 코드에 메시지가 표시됩니다.

인덱스에 사용자 지정 동의어 추가

인덱스에 사용자 지정 동의어를 추가하려면 동의어를 사전 파일에 지정합니다. 동의어 Amazon Kendra 사용에 비즈니스별 또는 특수 용어를 포함할 수 있습니다. Amazon Kendra 와 같은 일반 영어 동의어 (예: `leader, head` 하이픈을 사용하는 일반 동의어 포함) 는 사전 파일에 내장되어 있으므로 포함해서는 안 됩니다. Amazon Kendra 응답 유형 및/또는 응답 유형을 비롯한 모든 응답 유형의 동의어를 지원합니다. DOCUMENT QUESTION_ANSWER ANSWER Amazon Kendra 현재는 중지 단어로 플래그가 지정된 동의어 추가를 지원하지 않습니다. 향후 릴리스에 포함될 예정입니다.

Amazon Kendra 동의어 간에 상관 관계를 지정합니다. 예를 들어, 동의어 `Dynamo, Amazon DynamoDB` 쌍을 사용하면 Dynamo와 상호 연관됩니다. Amazon Kendra Amazon DynamoDB“dynamo란 무엇인가?”라는 쿼리 그러면 “What is?” 와 같은 문서를 반환합니다. Amazon DynamoDB동의어를 사용하면 상관 관계를 더 쉽게 파악할 Amazon Kendra 수 있습니다.

사전 파일은 버킷에 저장된 텍스트 파일입니다. Amazon S3 인덱스에 사전 추가 섹션을 참조하십시오.

사전 파일은 Solr 동의어 형식을 사용합니다. Amazon Kendra 색인당 사전 수에는 제한이 있습니다. 할당량을 참조하세요.

다음과 같은 상황에서 동의어가 유용할 수 있습니다.

- `NLP, Natural Language Processing` 같이 전통적인 영어 동의어가 아닌 특수 용어.
- 복잡한 의미적 연관성을 지닌 고유 명사. 이러한 명사는 일반 대중이 이해하기 어려운 명사입니다. 예를 들어 기계 학습에서는 이러한 명사들을 이해할 수 없습니다. `cost, loss, model performance`
- 다양한 형태의 제품 이름, 예: `Elastic Compute Cloud, EC2`.
- 도메인별 또는 비즈니스별 용어, 예: 제품 이름. 예를 들어 `Route53, DNS`입니다.

다음과 같은 시나리오에서는 동의어를 사용하지 마세요.

- 일반적인 영어 동의어, 예: `leader, head`. 이러한 동의어는 특정 도메인에만 국한되지 않으므로 이러한 시나리오에서 동의어를 사용하면 의도하지 않은 결과가 발생할 수 있습니다.
- 다음과 같은 서체 오류, `teh => the`.
- 명사의 복수형과 소유격, 형용사의 비교형과 최상위형, 동사의 과거형, 과거분사, 진행형 동사와 같은 형태학적 변형. 비교 형용사와 최상위 형용사의 한 예는 다음과 같습니다. `good, better, best`.

- 유니그램(한 단어)은 WHO 같은 단어를 뜻합니다. 유니그램 불용어는 사전에서 허용되지 않으며 검색에서 제외됩니다. 예를 들어, `WHO => World Health Organization`은 거부됩니다. 그러나 `W.H.O.`는 동의어 용어로 사용할 수 있으며, 불용어는 여러 단어로 된 동의어의 일부로 사용할 수 있습니다. 예를 들어, `of`는 허용되지 않지만 `United States of America`는 허용됩니다.

사용자 지정 동의어를 사용하면 비즈니스별 동의어를 포함하도록 쿼리를 확장하여 비즈니스별 용어에 대한 이해도를 쉽게 높일 Amazon Kendra수 있습니다. 동의어를 사용하면 검색 정확도가 향상될 수 있지만 이에 맞게 최적화하려면 동의어가 대기 시간에 어떤 영향을 미치는지 이해하는 것이 중요합니다.

동의어에 대한 일반적인 규칙은 쿼리에서 동의어와 일치하고 확장되는 용어가 많을수록 지연 시간에 미치는 영향이 커진다는 것입니다. 지연 시간에 영향을 미치는 다른 요인으로는 색인되는 문서의 평균 크기, 색인 크기, 검색 결과에 대한 필터링, 색인에 가해지는 전체 부하 등이 있습니다. Amazon Kendra 어떤 동의어와도 일치하지 않는 쿼리는 영향을 받지 않습니다.

동의어가 지연 시간에 미치는 영향에 대한 일반 지침:

사용 사례	지연 시간 증가*
각각 3~5단어의 일반적인 자연어 또는 키워드 쿼리	15% 미만
쿼리 용어 1개가 동의어 3개로 확장	
약 500,000 건의 문서(문서당 평균 10.48KB의 텍스트 추출) 또는 30,000개의 FAQ/질문 쌍으로 구성된 인덱스	

*성능은 인덱스의 특정 동의어 사용 및 구성에 따라 달라집니다. 특정 사용 사례에 맞는 더 정확한 벤치마크를 얻으려면 검색 성능을 테스트하는 것이 가장 좋습니다.

사전이 크고 용어 확장률이 높으며 지연 시간 증가가 허용 범위 내에 있지 않은 경우 다음 중 하나 또는 두 가지를 모두 시도해 볼 수 있습니다.

- 사전을 다듬어 확장률(용어당 동의어 수)을 줄이세요.
- 용어의 전체 적용 범위(사전의 줄 수)를 줄이세요.

또는 프로비저닝 용량(가상 스토리지 단위)을 늘려 지연 시간 증가를 상쇄할 수 있습니다.

주제

사전 파일 생성

Amazon Kendra 사전 파일은 Solr 동의어 목록 형식의 동의어 목록을 포함하는 UTF-8로 인코딩된 파일입니다. 사전 파일은 5MB 미만이어야 합니다.

동의어 매핑을 지정하는 방법에는 두 가지가 있습니다.

- 양방향 동의어는 쉼표로 구분된 용어 목록으로 지정됩니다. 사용자가 용어를 쿼리하면 목록의 모든 용어가 원래 쿼리된 용어가 포함된 문서를 검색하는 데 사용됩니다.
- 단방향 동의어는 용어를 해당 동의어에 매핑하기 위해 기호 “=>”로 구분된 용어로 지정됩니다. 사용자가 기호 “=>” 왼쪽에 있는 용어를 쿼리하면 오른쪽에 있는 용어에 매핑되어 동의어를 사용하여 문서를 검색합니다. 반대로 매핑되지 않으므로 단방향입니다.

동의어 자체는 대소문자를 구분하지만 매핑되는 용어는 대소문자를 구분하지 않습니다. 예를 들어, `ML => Machine Learning`은 사용자가 “ML” 또는 “ml”을 쿼리하거나 다른 대/소문자를 사용하는 경우 “Machine Learning”에 매핑된다는 의미입니다. 반대로 `Machine Learning => ML`과 같이 매핑하면 “Machine Learning” 또는 “machine learning” 또는 다른 대소문자는 “ML”에 매핑됩니다.

동의어는 특수 문자와 정확히 일치하는 항목을 검색하지 않습니다. 예를 들어 ""를 검색하면 dead-letter-queue “데드레터 큐" (하이픈 없음) 와 일치하는 문서를 Amazon Kendra 반환할 수 있습니다. 문서에 하이픈 (예: "dead-letter-queue“) 이 포함된 경우 검색 중에 문서를 Amazon Kendra 처리하여 하이픈을 제거합니다. 사전 파일에 포함되지만 포함해서는 안 되는 일반적인 영어 동의어 용어의 경우 용어의 하이픈 Amazon Kendra 버전과 하이픈이 아닌 버전을 모두 검색할 Amazon Kendra 수 있습니다. 예를 들어, “제3자”와 “제3자”를 검색하면 해당 용어의 두 버전 중 하나와 일치하는 문서가 Amazon Kendra 반환됩니다.

중지 단어 또는 일반적으로 사용되는 단어가 포함된 동의어의 경우 중지 단어를 포함하여 용어와 일치하는 문서를 Amazon Kendra 반환합니다. 예를 들어, “온보딩”과 “온보딩”을 매핑하는 동의어 규칙을

만들 수 있습니다. 동의어에는 스톱워드만 사용할 수 없습니다. 예를 들어, “on”을 검색하면 “on”이 포함된 모든 문서를 Amazon Kendra 반환할 수 없습니다.

일부 동의어 규칙은 무시됩니다. 예를 들어 a => b 는 규칙이지만 a => a 무시되며 규칙으로 간주되지 않습니다.

용어 수는 사전 파일에 있는 고유한 용어 수입니다. 아래 예제 파일에는 용어 AWS CodeStarML,Machine Learning, autoscaling groupASG, 등이 포함되어 있습니다.

사전에는 최대 동의어 규칙의 수와 용어당 최대 동의어 수가 있습니다. 자세한 정보는 에 대한 할당량 Amazon Kendra을 참조하세요.

다음 예제는 동의어 규칙이 있는 사전 파일을 보여줍니다. 각 줄에는 단일 동의어 규칙이 포함되어 있습니다. 빈 줄과 주석은 무시됩니다.

```
# 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.
```

인덱스에 사전 추가

다음 절차는 인덱스에 동의어가 포함된 사전 파일을 추가하는 방법을 보여줍니다. 업데이트된 사전 파일의 효과를 확인하는 데 최대 30분이 소요될 수 있습니다. 사전 파일에 대한 자세한 내용은 사전 파일 생성를 참조하세요.

Console

사전을 추가하려면

1. 왼쪽 탐색 창의 동의어 목록을 추가할 인덱스 아래에서, 사전, 동의어를 선택합니다.
2. 동의어 페이지에서 사전 추가를 선택합니다.
3. 사전 정의에서 사전의 이름과 선택적 설명을 입력합니다.
4. 사전 설정에서 사전 파일의 경로를 입력합니다. Amazon S3 파일은 5MB보다 작아야 합니다.
5. IAM 역할의 경우 역할을 선택하거나 새 역할 생성을 선택하고 역할 이름을 지정하여 새 역할을 생성합니다. Amazon Kendra 이 역할을 사용하여 사용자를 대신하여 Amazon S3 리소스에 액세스합니다. IAM 역할의 접두사는 "AmazonKendra- “입니다.
6. 저장을 선택하여 구성을 저장하고 사전을 추가합니다. 사전이 수집되면 해당 사전이 활성화되고 동의어가 결과에 강조 표시됩니다. 사전 파일의 효과를 확인하는 데 최대 30분이 소요될 수 있습니다.

CLI

를 사용하여 색인에 동의어 사전을 추가하려면 다음을 호출하십시오. AWS CLIcreate-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
```

list-thesauri를 호출하면 사전 목록을 볼 수 있습니다.

```
aws kendra list-thesauri \
--index-id index-id
```

사전의 세부 정보를 보려면 다음을 호출합니다. describe-thesaurus

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

사전 파일의 효과를 확인하는 데 최대 30분이 소요될 수 있습니다.

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.");
  }
}
```

사전 업데이트

사전이 생성된 이후 구성을 변경할 수 있습니다. 사전 이름 및 IAM 정보와 같은 세부 정보를 변경할 수 있습니다. 또한 사전 파일 Amazon S3 경로의 위치를 변경할 수 있습니다. 사전 파일의 경로를 변경하면 Amazon Kendra 가 기존 사전을 업데이트된 경로에 지정된 사전으로 대체합니다.

업데이트된 사전 파일의 효과를 확인하는 데 최대 30분이 소요될 수 있습니다.

> **Note**
>
> 사전 파일에 검증 또는 구문 오류가 있는 경우 이전에 업로드한 사전 파일이 보존됩니다.

다음 절차는 사전 세부 정보를 수정하는 방법을 보여줍니다.

Console

사전 세부 정보를 수정하는 방법

1. 왼쪽 탐색 창의 수정할 인덱스에서 동의어를 선택합니다.
2. 동의어 페이지에서 수정하려는 사전을 선택한 다음 편집을 선택합니다.
3. 사전 업데이트 페이지에서 사전 세부 정보를 업데이트합니다.
4. (선택 사항) 사전 파일 경로 변경을 선택한 다음 새 동의어 사전 파일의 Amazon S3 경로를 지정합니다. 기존 사전 파일은 지정한 파일로 대체됩니다. 경로를 변경하지 않는 경우 기존 경로에서 Amazon Kendra 사전을 다시 로드합니다.

   현재 사전 파일 유지를 선택하면 사전 파일이 다시 로드되지 Amazon Kendra 않습니다.
5. 저장을 선택하여 구성을 저장합니다.

기존 사전 경로에서 사전을 다시 로드할 수도 있습니다.

기존 경로에서 사전을 다시 로드하려면

1. 왼쪽 탐색 창의 수정할 인덱스에서 동의어를 선택합니다.
2. 동의어 페이지에서 다시 로드하려는 사전을 선택한 다음 새로 고침을 선택합니다.
3. 사전 파일 다시 로드 페이지에서 사전 파일을 새로 고칠지 확인합니다.

CLI

사전을 업데이트하려면 다음을 호출하세요. `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.");
  }
}
```

사전 삭제

다음 절차는 사전을 삭제하는 방법을 보여줍니다.

Console

1. 왼쪽 탐색 창의 수정할 인덱스에서 동의어를 선택합니다.
2. 동의어 페이지에서 삭제할 사전을 선택합니다.
3. 사전 세부 정보 페이지에서 삭제를 선택한 다음, 삭제할 것인지 확인합니다.

CLI

를 사용하여 색인의 동의어 사전을 삭제하려면 다음을 호출하십시오. AWS CLIdelete-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);
  }
}
```

검색 결과에 강조 표시하기

동의어 강조 표시는 기본적으로 켜져 있습니다. 하이라이트 정보는 Amazon Kendra SDK 및 CLI 쿼리 결과에 포함됩니다. SDK 또는 CLI를 Amazon Kendra 사용하여 상호 작용하는 경우 결과를 표시하는 방법을 결정합니다.

동의어 강조 표시에는 `THESAURUS_SYNONYM` 강조 표시 유형이 있습니다. 강조 표시에 대한 자세한 내용은 강조 표시 객체를 참조하세요.

자습서: Amazon Kendra를 사용하여 메타데이터가 풍부한 지능형 검색 솔루션 구축

이 자습서에서는 Amazon Kendra, Amazon Comprehend, Amazon Simple Storage Service (S3), AWS CloudShell를 사용하여 엔터프라이즈 데이터를 위한 메타데이터가 풍부한 자연어 기반의 지능형 검색 솔루션을 구축하는 방법을 보여줍니다.

Amazon Kendra는 비정형 자연어 데이터 리포지토리에 대한 검색 인덱스를 구축할 수 있는 지능형 검색 서비스입니다. 고객이 관련 답변을 더 쉽게 찾고 필터링할 수 있도록 Amazon Comprehend를 사용하여 데이터에서 메타데이터를 추출하고 이를 Amazon Kendra 검색 인덱스로 수집할 수 있습니다.

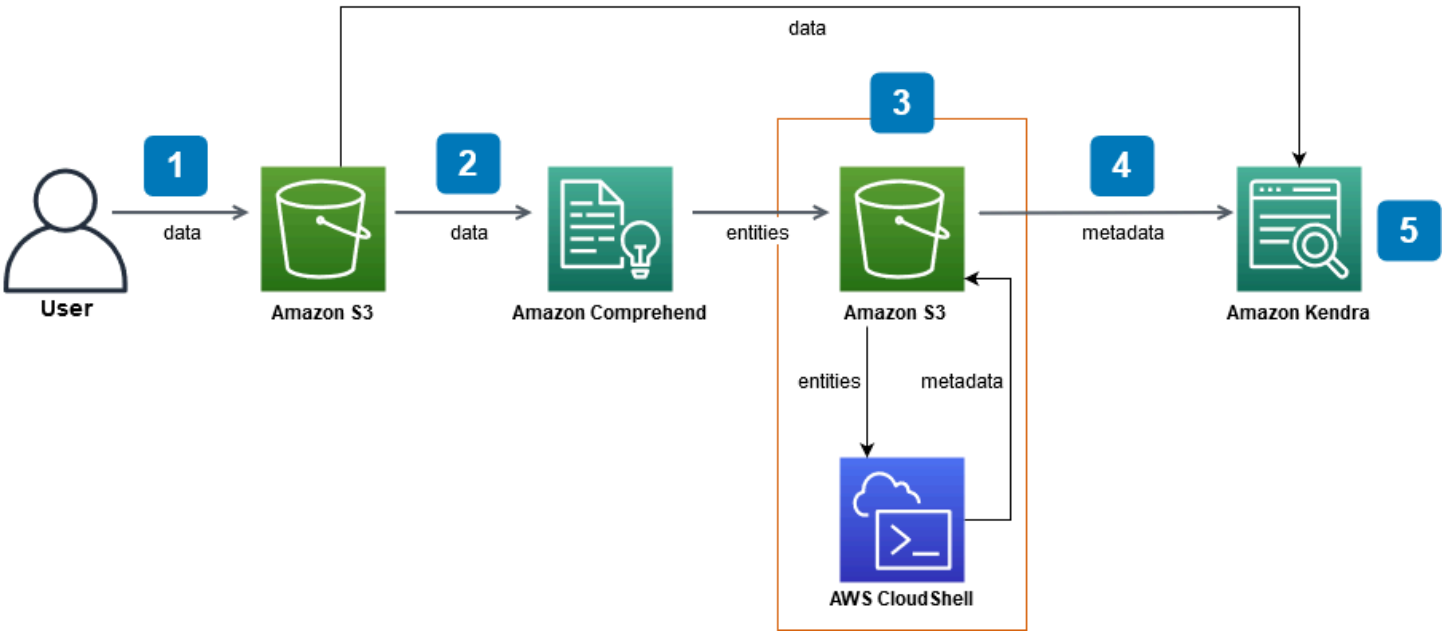
Amazon Comprehend는 개체에 대한 식별이 가능한 자연어 처리(NLP) 서비스입니다. 개체는 데이터에 있는 사람, 장소, 위치, 조직 및 객체에 대한 참조입니다.

이 자습서에서는 뉴스 기사의 샘플 데이터 세트를 사용하여 항목을 추출하고 메타데이터로 변환한 다음 Amazon Kendra 인덱스로 수집하여 검색을 실행합니다. 추가된 메타데이터를 사용하면 이러한 개체의 하위 집합을 사용하여 검색 결과를 필터링하여 검색 정확도를 높일 수 있습니다. 이 자습서를 따라하면 특별한 기계 학습 지식 없이도 엔터프라이즈 데이터를 위한 검색 솔루션을 만드는 방법을 배울 수 있습니다.

이 자습서에서는 다음 단계를 사용하여 검색 솔루션을 구축하는 방법을 보여줍니다.

1. Amazon S3에 뉴스 기사의 샘플 데이터 세트를 저장합니다.
2. Amazon Comprehend를 사용하여 데이터에서 개체를 추출합니다.
3. Python 3 스크립트를 실행하여 개체를 Amazon Kendra 인덱스 메타데이터 형식으로 변환하고 이 메타데이터를 S3에 저장합니다.
4. Amazon Kendra 검색 인덱스를 생성하고 데이터와 메타데이터를 수집합니다.
5. 검색 인덱스를 쿼리합니다.

다음 다이어그램은 워크플로를 보여줍니다.



이 자습서를 완료하는 데 걸리는 예상 시간: 1시간

예상 비용: 이 자습서의 일부 작업은 AWS 계정에 요금이 부과됩니다. 각 서비스의 비용에 대한 자세한 내용은 Amazon S3, Amazon Comprehend, AWS CloudShell, Amazon Kendra의 요금 페이지를 참조하세요.

주제

필수 조건

이 자습서를 완료하려면 다음 리소스가 필요합니다.

- AWS 계정. AWS 계정이 없는 경우 Amazon Kendra 설정의 단계에 따라 계정을 설정하십시오. AWS

- Windows, macOS 또는 Linux를 실행하며 AWS Management Console에 액세스할 수 있는 개발 컴퓨터. 자세한 내용은 AWS 관리 콘솔 구성을 참조하십시오.
- AWS Identity and Access Management (IAM) 사용자. 계정에 IAM 사용자 및 그룹을 설정하는 방법을 알아보려면 IAM 사용 설명서의 시작하기 섹션을 참조하세요.

  를 사용하는 경우 IAM 사용자에게 다음 정책을 추가하여 이 자습서를 완료하는 데 필요한 기본 권한을 부여해야 합니다. AWS Command Line Interface

  자세한 정보는 IAM 정책 생성 및 IAM 자격 증명 권한 추가 및 제거 섹션을 참조하세요.
- AWS 리전 서비스 목록. 지연 시간을 줄이려면 Amazon Comprehend와 Amazon Kendra에서 모두 지원하는 지리적 위치와 가장 가까운 AWS 리전을 선택해야 합니다.
- (선택 사항) AWS Key Management Service. 이 자습서에서는 암호화를 사용하지 않지만 특정 사용 사례에 맞는 암호화 모범 사례를 사용하는 것이 좋습니다.
- (선택 사항) Amazon Virtual Private Cloud. 이 자습서에서는 VPC를 사용하지 않지만 VPC 모범 사례를 사용하여 특정 사용 사례의 데이터 보안을 보장하는 것이 좋습니다.

1단계: Amazon S3에 문서 추가

데이터 세트에 대해 Amazon Comprehend 개체 분석 작업을 실행하기 전에 데이터, 메타데이터 및 Amazon Comprehend 개체 분석 출력을 호스팅할 Amazon S3 버킷을 생성합니다.

주제

- 샘플 데이터 세트 다운로드
- Amazon S3 버킷 생성
- S3 버킷에 데이터 및 메타데이터 폴더 생성
- 입력 데이터 로드

샘플 데이터 세트 다운로드

Amazon Comprehend가 데이터에 대한 개체 분석 작업을 실행하려면 먼저 데이터 세트를 다운로드하고 추출한 다음 S3 버킷에 업로드해야 합니다.

데이터 세트를 다운로드하고 추출하려면 (콘솔)

1. 디바이스에 tutorial-dataset.zip 폴더를 다운로드합니다.

2. tutorial-dataset 폴더를 추출하여 data 폴더에 접근합니다.

데이터 세트를 다운로드하고 추출하려면 (터미널)

1. tutorial-dataset를 다운로드하려면 터미널 창에서 다음 명령을 실행합니다.

   Linux

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   위치:

   - *path/*는 zip 폴더를 저장하려는 위치의 로컬 파일 경로입니다.

   macOS

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   위치:

   - *path/*는 zip 폴더를 저장하려는 위치의 로컬 파일 경로입니다.

   Windows

   ```
   curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/tutorial-dataset.zip
   ```

   위치:

   - *path/*는 zip 폴더를 저장하려는 위치의 로컬 파일 경로입니다.

2. zip 폴더에서 데이터를 추출하려면 터미널 창에서 다음 명령을 실행합니다.

   Linux

   ```
   unzip path/tutorial-dataset.zip -d path/
   ```

위치:

- *path/*는 저장된 zip 폴더의 로컬 파일 경로입니다.

macOS

```
unzip path/tutorial-dataset.zip -d path/
```

위치:

- *path/*는 저장된 zip 폴더의 로컬 파일 경로입니다.

Windows

```
tar -xf path/tutorial-dataset.zip -C path/
```

위치:

- *path/*는 저장된 zip 폴더의 로컬 파일 경로입니다.

이 단계가 끝나면 압축을 푼 파일을 `tutorial-dataset`라는 이름의 압축 해제 폴더에 저장해야 합니다. 이 폴더에는 Apache 2.0 오픈 소스 속성이 있는 `README` 파일과 이 자습서의 데이터 세트가 들어 있는 `data` 폴더가 포함되어 있습니다. 데이터 세트는 `.story` 확장자를 가진 100개의 파일로 구성되어 있습니다.

Amazon S3 버킷 생성

샘플 데이터 폴더를 다운로드하고 추출한 후 Amazon S3 버킷에 저장합니다.

> **Important**
>
> Amazon S3 버킷의 이름은 모든 AWS에 대해 고유해야 합니다.

Amazon S3 버킷 생성(콘솔)

1. AWS Management Console 로그인하고 https://console.aws.amazon.com/s3/ 에서 Amazon S3 콘솔을 엽니다.
2. 버킷에서 버킷 생성을 선택합니다.
3. [버킷 이름(Bucket name)]에 고유한 이름을 입력합니다.
4. 리전에서 버킷을 생성할 AWS 리전을 선택합니다.

> **Note**
>
> Amazon Comprehend와 Amazon Kendra를 모두 지원하는 리전을 선택해야 합니다. 버킷을 생성한 후에는 버킷의 리전을 변경할 수 없습니다.

5. 이 버킷의 퍼블릭 액세스 차단 설정, 버킷 버전 관리 및 태그의 기본 설정을 유지하세요.
6. 기본 암호화의 경우, 비활성화를 선택합니다.
7. 고급 설정의 기본 설정을 유지합니다.
8. 버킷 구성을 검토한 다음 버킷 생성을 선택합니다.

S3 버킷을 생성하려면 (AWS CLI)

1. S3 버킷을 생성하려면 AWS CLI에서 create-bucket 명령을 사용합니다.

   Linux

   ```
   aws s3api create-bucket \
           --bucket DOC-EXAMPLE-BUCKET \
           --region aws-region \
           --create-bucket-configuration LocationConstraint=aws-region
   ```

   위치:

   - DOC-EXAMPLE-BUCKET은 버킷 이름입니다.
   - *aws-region*은 버킷을 생성하려는 리전입니다.

   macOS

   ```
   aws s3api create-bucket \
   ```

```
        --bucket DOC-EXAMPLE-BUCKET \
        --region aws-region \
        --create-bucket-configuration LocationConstraint=aws-region
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.
- *aws-region*은 버킷을 생성하려는 리전입니다.

Windows

```
aws s3api create-bucket ^
        --bucket DOC-EXAMPLE-BUCKET ^
        --region aws-region ^
        --create-bucket-configuration LocationConstraint=aws-region
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.
- *aws-region*은 버킷을 생성하려는 리전입니다.

> **Note**
>
> Amazon Comprehend와 Amazon Kendra를 모두 지원하는 리전을 선택해야 합니다. 버킷을 생성한 후에는 버킷의 리전을 변경할 수 없습니다.

2. 버킷이 성공적으로 생성되었는지 확인하려면 list 명령을 사용하세요.

Linux

```
aws s3 ls
```

macOS

```
aws s3 ls
```

Windows

```
aws s3 ls
```

S3 버킷에 데이터 및 메타데이터 폴더 생성

S3 버킷을 생성한 후 그 안에 데이터 및 메타데이터 폴더를 생성합니다.

S3 버킷에서 폴더 생성(콘솔)

1. https://console.aws.amazon.com/s3/에서 S3 콘솔을 엽니다.
2. 버킷에서 버킷 목록에서 해당하는 버킷의 이름을 클릭합니다.
3. 객체 탭에서 폴더 생성을 선택합니다.
4. 새 폴더 이름으로 **data**을 입력합니다.
5. 암호화 설정에서 비활성화를 선택합니다.
6. 폴더 생성을 선택합니다.
7. 3~6단계를 반복하여 Amazon Kendra 메타데이터를 저장할 다른 폴더를 생성하고 4단계에서 생성한 폴더의 이름을 **metadata**로 지정합니다.

S3 버킷에서 폴더 생성(AWS CLI)

1. S3 버킷에 `data` 폴더를 생성하려면 AWS CLI에서 put-object 명령을 사용합니다.

   Linux

   ```
   aws s3api put-object \
           --bucket DOC-EXAMPLE-BUCKET \
           --key data/
   ```

   위치:

   - DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

   macOS

   ```
   aws s3api put-object \
   ```

```
        --bucket DOC-EXAMPLE-BUCKET \
        --key data/
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

Windows

```
aws s3api put-object ^
        --bucket DOC-EXAMPLE-BUCKET ^
        --key data/
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

2. S3 버킷에 `metadata` 폴더를 생성하려면 AWS CLI에서 put-object 명령을 사용합니다.

Linux

```
aws s3api put-object \
        --bucket DOC-EXAMPLE-BUCKET \
        --key metadata/
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

macOS

```
aws s3api put-object \
        --bucket DOC-EXAMPLE-BUCKET \
        --key metadata/
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

Windows

```
aws s3api put-object ^
        --bucket DOC-EXAMPLE-BUCKET ^
        --key metadata/
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

3. 폴더가 성공적으로 생성되었는지 확인하려면 list 명령어를 사용하여 버킷의 콘텐츠를 확인하세요.

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

위치:

- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

입력 데이터 로드

데이터 및 메타데이터 폴더를 만든 후 샘플 데이터 세트를 data 폴더에 업로드합니다.

샘플 데이터 세트를 데이터 폴더에 업로드하려면 (콘솔)

1. https://console.aws.amazon.com/s3/에서 S3 콘솔을 엽니다.
2. 버킷에서 버킷 목록에서 버킷의 이름을 클릭한 다음 data를 클릭합니다.
3. 업로드를 선택한 후 파일 추가를 선택합니다.
4. 대화 상자에서 로컬 디바이스의 tutorial-dataset 폴더 내 data 폴더로 이동하여 모든 파일을 선택한 다음 열기를 선택합니다.
5. 대상, 권한 및 속성에 대한 기본 설정을 유지합니다.
6. 업로드를 선택합니다.

샘플 데이터 세트를 데이터 폴더에 업로드하려면 (AWS CLI)

1. 샘플 데이터를 data 폴더에 업로드하려면 AWS CLI에서 copy 명령어를 사용하세요.

   Linux

   ```
   aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
   ```

   위치:

   - *path/*는 디바이스에 있는 tutorial-dataset 폴더의 파일 경로입니다.
   - DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

   macOS

   ```
   aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
   ```

   위치:

   - *path/*는 디바이스에 있는 tutorial-dataset 폴더의 파일 경로입니다.
   - DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

Windows

```
aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
```

위치:

- *path/*는 디바이스에 있는 tutorial-dataset 폴더의 파일 경로입니다.
- DOC-EXAMPLE-BUCKET은 버킷 이름입니다.

2. 데이터 세트 파일이 data 폴더에 성공적으로 업로드되었는지 확인하려면 AWS CLI에서 list 명령어를 사용하세요.

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

이 단계가 끝나면 `data` 폴더 내에 데이터 세트가 저장된 S3 버킷과 Amazon Kendra 메타데이터를 저장할 빈 `metadata` 폴더가 생깁니다.

2단계: Amazon Comprehend에서 개체 분석 작업 실행

샘플 데이터 세트를 S3 버킷에 저장한 후 Amazon Comprehend 개체 분석 작업을 실행하여 문서에서 개체를 추출합니다. 이러한 개체는 Amazon Kendra 사용자 지정 속성을 형성하고 인덱스에서 검색 결과를 필터링하는 데 도움이 됩니다. 자세한 내용은 개체 감지를 참조하세요.

주제

- Amazon Comprehend 개체 분석 작업 실행

Amazon Comprehend 개체 분석 작업 실행

데이터 세트에서 개체를 추출하려면 Amazon Comprehend 개체 분석 작업을 실행합니다.

이 단계에서 AWS CLI를 사용하는 경우, 먼저 Amazon Comprehend에 대한 AWS IAM 역할 및 정책을 생성하여 연결한 다음 항목 분석 작업을 실행합니다. 샘플 데이터에서 개체 분석 작업을 실행하려면 Amazon Comprehend에 다음이 필요합니다.

- 신뢰할 수 있는 개체로 인식하는 AWS Identity and Access Management (IAM) 역할
- AWS S3 버킷에 액세스할 수 있는 권한을 부여하는 IAM 역할에 연결된 IAM 정책

자세한 내용은 Amazon Comprehend에서 IAM을 사용하는 방법 및 Amazon Comprehend의 자격 증명 기반 정책을 참조하세요.

Amazon Comprehend 개체 분석 작업을 실행하려면 (콘솔)

1. https://console.aws.amazon.com/comprehend/에서 Amazon Comprehend 콘솔을 엽니다.

   > **Important**
   >
   > Amazon S3 버킷을 생성한 리전과 동일한 리전에 있는지 확인합니다. 다른 지역에 있는 경우 상단 탐색 표시줄의 AWS 지역 선택기에서 S3 버킷을 생성한 지역을 선택합니다.

2. Amazon Comprehend 시작을 선택합니다.
3. 왼쪽 탐색 창에서 분석 작업을 선택합니다.

4. 작업 생성을 선택합니다.
5. 작업 설정 섹션에서 다음을 수행합니다.
   a. 이름에 **data-entities-analysis**를 입력합니다.
   b. 분석 유형에서 개체를 선택합니다.
   c. 언어에서 영어를 선택합니다.
   d. 작업 암호화를 끈 상태로 둡니다.
6. 입력 데이터 섹션에서 다음을 수행합니다.
   a. 데이터 소스로 내 문서를 선택합니다.
   b. S3 위치의 경우 S3 찾아보기를 선택합니다.
   c. 리소스 선택의 경우, 버킷 목록에서 해당하는 버킷의 이름을 클릭합니다.
   d. 객체의 경우 `data` 옵션 버튼을 선택하고 선택을 선택합니다.
   e. 입력 형식으로 파일 하나에 문서 하나 선택합니다.
7. 출력 데이터 섹션에서 다음을 수행합니다.
   a. S3 위치에서 S3 찾아보기를 선택한 다음 버킷 목록에서 버킷에 대한 옵션 상자를 선택하고 선택을 선택합니다.
   b. 암호화를 끈 상태로 둡니다.
8. 권한 연결 섹션에서 다음을 수행합니다.
   a. IAM 역할에서 IAM 역할 생성을 선택합니다.
   b. 액세스 권한은 입력 및 출력 S3 버킷을 선택합니다.
   c. 이름 접미사에 **comprehend-role**을 입력합니다. 이 역할은 Amazon S3 버킷에 대한 액세스를 제공합니다.
9. 기본 VPC 설정을 유지합니다.
10. 작업 생성을 선택합니다.

Amazon Comprehend 개체 분석 작업을 실행하려면 (AWS CLI)

1. 신뢰할 수 있는 개체로 인식되는 Amazon Comprehend용 IAM 역할을 생성하고 연결하려면 다음을 수행하세요.
   a. 다음 신뢰 정책을 로컬 디바이스의 텍스트 편집기에서 `comprehend-trust-policy.json`이라는 JSON 파일로 저장합니다.

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

b. comprehend-role라는 IAM 역할을 생성하고 저장된 comprehend-trust-policy.json 파일을 첨부하려면 create-role 명령을 사용합니다.

Linux

```
aws iam create-role \
          --role-name comprehend-role \
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 comprehend-trust-policy.json의 파일 경로입니다.

macOS

```
aws iam create-role \
          --role-name comprehend-role \
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 comprehend-trust-policy.json의 파일 경로입니다.

Windows

```
aws iam create-role ^
          --role-name comprehend-role ^
          --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 comprehend-trust-policy.json의 파일 경로입니다.

c. Amazon 리소스 이름(ARN)을 텍스트 편집기에 복사하고 로컬에 comprehend-role-arn로 저장합니다.

> **Note**
>
> ARN 형식은 *arn:aws:iam::123456789012:role/comprehend-role*과 비슷합니다. Amazon Comprehend 분석 작업을 실행하려면 comprehend-role-arn로 저장한 ARN이 필요합니다.

2. S3 버킷에 액세스할 권한을 부여하는 IAM 정책을 생성하여 IAM 역할에 연결하려면 다음을 수행하세요.

a. 다음 신뢰 정책을 로컬 디바이스의 텍스트 편집기에서 comprehend-S3-access-policy.json이라는 JSON 파일로 저장합니다.

```
{
    "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. S3 버킷에 액세스하기 위해 comprehend-S3-access-policy라는 IAM 정책을 생성하려면 create-policy 명령을 사용하세요.

Linux

```
aws iam create-policy \
          --policy-name comprehend-S3-access-policy \
          --policy-document file://path/comprehend-S3-access-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 comprehend-S3-access-policy.json의 파일 경로입니다.

macOS

```
aws iam create-policy \
          --policy-name comprehend-S3-access-policy \
          --policy-document file://path/comprehend-S3-access-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 comprehend-S3-access-policy.json의 파일 경로입니다.

Windows

```
aws iam create-policy ^
          --policy-name comprehend-S3-access-policy ^
          --policy-document file://path/comprehend-S3-access-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 comprehend-S3-access-policy.json의 파일 경로입니다.

c. Amazon 리소스 이름(ARN)을 텍스트 편집기에 복사하고 로컬에 comprehend-S3-access-arn로 저장합니다.

> **Note**
>
> ARN 형식은 *arn:aws:iam::123456789012:role/comprehend-S3-access-policy*와 비슷합니다. IAM 역할에 comprehend-S3-access-policy를 연결하려면 comprehend-S3-access-arn로 저장한 ARN이 필요합니다.

d. IAM 역할에 comprehend-S3-access-policy 연결하려면 다음 명령을 사용하십시오. attach-role-policy

Linux

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name comprehend-role
```

위치:

- *policy-arn*은 comprehend-S3-access-arn로 저장한 ARN입니다.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name comprehend-role
```

위치:

- *policy-arn*은 comprehend-S3-access-arn로 저장한 ARN입니다.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name comprehend-role
```

위치:

- *policy-arn*은 comprehend-S3-access-arn로 저장한 ARN입니다.

3. Amazon Comprehend 엔티티 분석 작업을 실행하려면 다음 명령을 사용하십시오. 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
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.
- *role-arn*은 comprehend-role-arn로 저장한 ARN입니다.

- *AWS-###* 사용자 지역입니다. 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
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.
- *role-arn*은 comprehend-role-arn로 저장한 ARN입니다.
- *AWS-###* 사용자 지역입니다. 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
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.
- *role-arn*은 comprehend-role-arn로 저장한 ARN입니다.
- *AWS-###* 사용자 지역입니다. AWS

4. 개체 분석 JobId를 복사하여 텍스트 편집기에 comprehend-job-id로 저장합니다. JobId는 개체 분석 작업의 상태를 추적하는 데 도움이 됩니다.

5. 엔티티 분석 작업의 진행 상황을 추적하려면 다음 명령을 사용하십시오. describe-entities-detection-job

Linux

```
aws comprehend describe-entities-detection-job \
        --job-id entities-job-id \
        --region aws-region
```

위치:

- *entities-job-id*comprehend-job-id저장되었나요?
- *AWS-###* 사용자 지역입니다. AWS

macOS

```
aws comprehend describe-entities-detection-job \
        --job-id entities-job-id \
        --region aws-region
```

위치:

- *entities-job-id*저장하셨나요? comprehend-job-id
- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws comprehend describe-entities-detection-job ^
        --job-id entities-job-id ^
        --region aws-region
```

위치:

- *entities-job-id*저장하셨나요? comprehend-job-id
- *AWS-###* 사용자 지역입니다. AWS

JobStatus가 COMPLETED로 변경되는 데 몇 분 정도 걸릴 수 있습니다.

이 단계가 끝나면 Amazon Comprehend는 개체 분석 결과를 S3 버킷의 자동 생성 폴더 내 output폴더에 압축된 output.tar.gz 파일로 저장합니다. 다음 단계로 넘어가기 전에 분석 작업 상태가 완료되었는지 확인합니다.

3단계: 개체 분석 출력을 Amazon Kendra 메타데이터로 형식 지정

Amazon Comprehend에서 추출한 개체를 Amazon Kendra 인덱스에 필요한 메타데이터 형식으로 변환하려면 Python 3 스크립트를 실행합니다. 변환 결과는 Amazon S3 버킷의 metadata 폴더에 저장됩니다.

Amazon Kendra 메타데이터 형식 및 구조에 대한 자세한 내용은 S3 문서 메타데이터를 참조하세요.

주제

- Amazon Comprehend 출력 다운로드 및 추출
- 출력을 S3 버킷에 업로드
- 출력을 Amazon Kendra 메타데이터 형식으로 변환
- Amazon S3 버킷 정리

Amazon Comprehend 출력 다운로드 및 추출

Amazon Comprehend 개체 분석 출력의 형식을 지정하려면 먼저 Amazon Comprehend 개체 분석 output.tar.gz 아카이브를 다운로드하고 개체 분석 파일을 추출해야 합니다.

출력 파일을 다운로드하려면(콘솔)

1. Amazon Comprehend 콘솔의 탐색 창에서 분석 작업으로 이동합니다.
2. 개체 분석 작업 data-entities-analysis를 선택합니다.
3. 출력에서 출력 데이터 위치 옆에 표시된 링크를 선택합니다. 그러면 S3 버킷의 output.tar.gz 아카이브로 리디렉션됩니다.
4. 개요 탭에서 다운로드를 선택합니다.

> **Tip**
>
> 모든 Amazon Comprehend 분석 작업의 출력 이름은 동일합니다. 아카이브 이름을 변경하면 아카이브를 더 쉽게 추적할 수 있습니다.

5. 디바이스에 다운로드한 Amazon Comprehend 파일의 압축을 풀고 추출합니다.

출력 파일을 다운로드하려면(AWS CLI)

1. 항목 분석 작업 결과가 포함된 S3 버킷의 Amazon Comprehend 자동 생성 폴더 이름에 액세스하려면 다음 명령을 사용하십시오. describe-entities-detection-job

   Linux

   ```
   aws comprehend describe-entities-detection-job \
             --job-id entities-job-id \
             --region aws-region
   ```

   위치:

   - *entities-job-id*저장한 원본은 다음과 같습니다. comprehend-job-id the section called “2단계: 개체 감지”
   - *AWS-###* 사용자 지역입니다. AWS

   macOS

   ```
   aws comprehend describe-entities-detection-job \
             --job-id entities-job-id \
             --region aws-region
   ```

   위치:

   - *entities-job-id*comprehend-job-id저장한 출처입니다. the section called “2단계: 개체 감지”
   - *AWS-###* 사용자 지역입니다. AWS

   Windows

   ```
   aws comprehend describe-entities-detection-job ^
             --job-id entities-job-id ^
             --region aws-region
   ```

   위치:

- *entities-job-id*comprehend-job-id저장한 출처입니다. the section called “2단계: 개체 감지”
- *AWS-###* 사용자 지역입니다. AWS

2. 개체 작업 설명의 OutputDataConfig 객체에서 S3Uri 값을 복사하고 텍스트 편집기에 comprehend-S3uri로 저장합니다.

> **Note**
>
> S3Uri*## ### s3://DOC-EXAMPLE-BUCKET/... # #####.* /output/output.tar.gz.

3. 개체 출력 아카이브를 다운로드하려면 copy 명령을 사용합니다.

Linux

```
aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
```

위치:

- *s3://DOC-EXAMPLE-BUCKET/...* /output/output.tar.gz 는 저장한 S3Uri 값으로comprehend-S3uri,
- *path/*는 출력을 저장하려는 로컬 디렉터리입니다.

macOS

```
aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
```

위치:

- *s3://DOC-EXAMPLE-BUCKET/...* /output/output.tar.gz 는 저장한 S3Uri 값으로comprehend-S3uri,
- *path/*는 출력을 저장하려는 로컬 디렉터리입니다.

Windows

```
aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
```

위치:

- *s3://DOC-EXAMPLE-BUCKET/...* /output/output.tar.gz 는 저장한 S3Uri 값으로comprehend-S3uri,
- *path/*는 출력을 저장하려는 로컬 디렉터리입니다.

4. 개체 출력을 추출하려면 터미널 창에서 다음 명령을 실행합니다.

Linux

```
tar -xf path/output.tar.gz -C path/
```

위치:

- *path/*는 로컬 디바이스에 다운로드한 output.tar.gz 아카이브의 파일 경로입니다.

macOS

```
tar -xf path/output.tar.gz -C path/
```

위치:

- *path/*는 로컬 디바이스에 다운로드한 output.tar.gz 아카이브의 파일 경로입니다.

Windows

```
tar -xf path/output.tar.gz -C path/
```

위치:

- *path/*는 로컬 디바이스에 다운로드한 output.tar.gz 아카이브의 파일 경로입니다.

이 단계가 끝나면 Amazon Comprehend에서 식별한 개체 목록이 포함된 output라는 파일이 디바이스에 생성되어야 합니다.

출력을 S3 버킷에 업로드

Amazon Comprehend 개체 분석 파일을 다운로드하고 추출한 후 추출한 output 파일을 Amazon S3 버킷에 업로드합니다.

추출한 Amazon Comprehend 출력 파일 업로드(콘솔)

1. https://console.aws.amazon.com/s3/에서 S3 콘솔을 엽니다.
2. 버킷에서 버킷 이름을 선택한 다음 업로드를 선택합니다.
3. 파일 및 폴더에서 파일 추가를 선택합니다.
4. 대화 상자에서 디바이스의 추출된 output 파일을 찾아 선택한 다음 열기를 선택합니다.
5. 대상, 권한 및 속성에 대한 기본 설정을 유지합니다.
6. 업로드를 선택합니다.

추출한 Amazon Comprehend 출력 파일 업로드(AWS CLI)

1. 추출된 output 파일을 버킷에 업로드하려면 copy 명령을 사용합니다.

   Linux

   ```
   aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
   ```

   위치:

   - *path/*는 추출된 output 파일의 로컬 파일 경로입니다.
   - DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

   macOS

   ```
   aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
   ```

   위치:

   - *path/*는 추출된 output 파일의 로컬 파일 경로입니다.
   - DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

Windows

```
aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
```

위치:

- *path/*는 추출된 output 파일의 로컬 파일 경로입니다.
- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

2. output 파일이 S3 버킷에 성공적으로 업로드되었는지 확인하려면 list 명령을 사용하여 파일의 내용을 확인하세요.

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

출력을 Amazon Kendra 메타데이터 형식으로 변환

Amazon Comprehend 출력을 Amazon Kendra 메타데이터로 변환하려면 Python 3 스크립트를 실행합니다. 콘솔을 사용하는 AWS CloudShell 경우 이 단계에서 사용합니다.

Python 3 스크립트를 실행하려면 (콘솔)

1. converter.py.zip 파일을 디바이스에 다운로드합니다.
2. Python 3 파일 `converter.py`를 추출합니다.
3. AWS 관리 콘솔에 로그인하고 AWS 지역이 S3 버킷 및 Amazon Comprehend 분석 작업과 동일한 지역으로 설정되어 있는지 확인하십시오.
4. AWS CloudShell 아이콘을 선택하거나 상단 탐색 표시줄의 검색 상자에 입력하여 환경을 시작하십시오. AWS CloudShell

   > **Note**
   >
   > 새 브라우저 창에서 처음 실행하면 AWS CloudShell 시작 패널이 표시되고 주요 기능이 나열됩니다. 이 패널을 닫고 명령 프롬프트가 표시되면 쉘이 상호 작용할 준비가 된 것입니다.

5. 터미널이 준비되면 탐색 창에서 작업을 선택한 다음 메뉴에서 파일 업로드를 선택합니다.
6. 이때 열리는 대화 상자에서 파일 선택 을 선택한 후 디바이스에서 다운로드한 Python 3 파일 `converter.py`를 선택합니다. 업로드를 선택합니다.
7. AWS CloudShell 환경에서 다음 명령을 입력합니다.

   ```
   python3 converter.py
   ```

8. 쉘 인터페이스에서 S3 버킷 이름을 입력하라는 메시지가 표시되면 S3 버킷의 이름을 입력하고 Enter 키를 누릅니다.
9. 쉘 인터페이스에서 Comprehend 출력 파일의 전체 파일 경로를 입력하라는 메시지가 표시되면 **`output`**를 입력하고 Enter 키를 누릅니다.
10. 쉘 인터페이스에서 메타데이터 폴더의 전체 파일 경로를 입력하라는 메시지가 표시되면 **`metadata/`**를 입력하고 Enter 키를 누릅니다.

> **Important**
> 메타데이터의 형식을 올바르게 지정하려면 8-10단계의 입력 값이 정확해야 합니다.

Python 3 스크립트를 실행하려면 (AWS CLI)

1. Python 3 파일 `converter.py`를 다운로드하려면 터미널 창에서 다음 명령을 실행합니다.

   Linux

   ```
   curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/converter.py.zip
   ```

   위치:

   - *path/*는 압축된 파일을 저장하려는 위치의 파일 경로입니다.

   macOS

   ```
   curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/converter.py.zip
   ```

   위치:

   - *path/*는 압축된 파일을 저장하려는 위치의 파일 경로입니다.

   Windows

   ```
   curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/
   samples/converter.py.zip
   ```

   위치:

   - *path/*는 압축된 파일을 저장하려는 위치의 파일 경로입니다.

2. Python 3 파일을 추출하려면 터미널 창에서 다음 명령을 실행합니다.

Linux

```
unzip path/converter.py.zip -d path/
```

위치:

- *path/*는 저장한 converter.py.zip의 파일 경로입니다.

macOS

```
unzip path/converter.py.zip -d path/
```

위치:

- *path/*는 저장한 converter.py.zip의 파일 경로입니다.

Windows

```
tar -xf path/converter.py.zip -C path/
```

위치:

- *path/*는 저장한 converter.py.zip의 파일 경로입니다.

3. 다음 명령을 실행하여 Boto3가 디바이스에 설치되어 있는지 확인합니다.

Linux

```
pip3 show boto3
```

macOS

```
pip3 show boto3
```

Windows

```
pip3 show boto3
```

> **Note**
>
> Boto3가 설치되어 있지 않은 경우 `pip3 install boto3`를 실행하여 설치합니다.

4. Python 3 스크립트를 실행하여 `output` 파일을 변환하려면 다음 명령을 실행합니다.

   Linux

   ```
   python path/converter.py
   ```

   위치:

   - *path/*는 저장한 `converter.py.zip`의 파일 경로입니다.

   macOS

   ```
   python path/converter.py
   ```

   위치:

   - *path/*는 저장한 `converter.py.zip`의 파일 경로입니다.

   Windows

   ```
   python path/converter.py
   ```

   위치:

   - *path/*는 저장한 `converter.py.zip`의 파일 경로입니다.

5. AWS CLI 메시지가 표시되면 S3 버킷의 이름을 입력하고 Enter 키를 누릅니다. `Enter the name of your S3 bucket`
6. AWS CLI 메시지가 표시되면 를 **`output`** 입력하고 Enter 키를 누릅니다. `Enter the full filepath to your Comprehend output file`
7. AWS CLI 메시지가 표시되면 Enter를 **`metadata/`** 입력하고 Enter 키를 누릅니다. `Enter the full filepath to your metadata folder`

> **Important**
>
> 메타데이터의 형식을 올바르게 지정하려면 5-7단계의 입력 값이 정확해야 합니다.

이 단계가 끝나면 형식이 지정된 메타데이터가 S3 버킷의 `metadata` 폴더 내에 보관됩니다.

Amazon S3 버킷 정리

Amazon Kendra 인덱스는 버킷에 저장된 모든 파일을 동기화하므로 검색 결과가 중복되지 않도록 Amazon S3 버킷을 정리하는 것이 좋습니다.

Amazon S3 버킷을 정리하려면 (콘솔)

1. https://console.aws.amazon.com/s3/에서 S3 콘솔을 엽니다.
2. 버킷에서 버킷을 선택한 다음 Amazon Comprehend 개체 분석 출력 폴더, Amazon Comprehend 개체 분석 `.temp` 파일 및 추출된 Amazon Comprehend `output` 파일을 선택합니다.
3. 개요 탭에서 삭제를 선택합니다.
4. 객체 삭제에서 객체를 영구적으로 삭제하시겠습니까?를 선택하고 텍스트 입력 필드에 **`permanently delete`**를 입력합니다.
5. 객체 삭제를 선택합니다.

Amazon S3 버킷 정리(AWS CLI)

1. `data` 및 `metadata` 폴더를 제외한 S3 버킷의 모든 파일 및 폴더를 삭제하려면 AWS CLI에서 remove 명령을 사용합니다.

   Linux

   ```
   aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude
    "metadata/*"
   ```

   위치:

   - DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

macOS

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude
 "metadata/*"
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

Windows

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude
 "metadata/*"
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

2. 객체가 S3 버킷에서 성공적으로 삭제되었는지 확인하려면 list 명령을 사용하여 해당 콘텐츠를 확인하세요.

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

위치:

- DOC-EXAMPLE-BUCKET은 S3 버킷 이름입니다.

이 단계가 끝나면 Amazon Comprehend 개체 분석 출력을 Amazon Kendra 메타데이터로 변환하게 됩니다. 이제 Amazon Kendra 인덱스를 생성할 준비가 완료되었습니다.

4단계: Amazon Kendra 인덱스를 생성하고 메타데이터를 수집

지능형 검색 솔루션을 구현하려면 Amazon Kendra 인덱스를 생성하고 이 인덱스에 S3 데이터 및 메타데이터를 수집해야 합니다.

Amazon Kendra 인덱스에 메타데이터를 추가하기 전에 사용자 지정 문서 속성에 해당하는 사용자 지정 인덱스 필드를 생성합니다. 그러면 Amazon Comprehend 개체 유형과 맞게 됩니다. Amazon Kendra는 사용자가 생성한 인덱스 필드와 사용자 지정 문서 속성을 사용하여 문서를 검색하고 필터링합니다.

자세한 내용은 인덱싱 및 사용자 지정 문서 속성 생성을 참조하세요.

주제

- Amazon Kendra 인덱스를 생성합니다.
- Amazon S3 액세스를 위한 IAM 역할 업데이트
- Amazon Kendra 사용자 지정 검색 인덱스 필드 생성
- Amazon S3 버킷을 인덱스의 데이터 소스로 추가
- Amazon Kendra 인덱스 동기화

Amazon Kendra 인덱스를 생성합니다.

소스 문서를 쿼리하려면 Amazon Kendra 인덱스를 생성해야 합니다.

이 단계에서 를 사용하는 경우 인덱스를 생성하기 전에 Amazon Kendra가 로그에 CloudWatch 액세스할 수 있도록 허용하는 AWS IAM 역할 및 정책을 생성하여 연결합니다. AWS CLI 자세한 내용은 사전 조건을 참조하세요.

Amazon Kendra 인덱스를 생성하려면 (콘솔)

1. https://console.aws.amazon.com/kendra/에서 Amazon Kendra 콘솔을 엽니다.

   > **Important**
   >
   > Amazon Comprehend 개체 분석 작업, Amazon S3 버킷을 생성한 리전과 동일한 리전에 있는지 확인하세요. 다른 지역에 있는 경우 상단 탐색 표시줄의 AWS 지역 선택기에서 Amazon S3 버킷을 생성한 지역을 선택합니다.

2. 인덱스 생성을 선택합니다.
3. 인덱스 세부 정보 지정 페이지의 인덱스 세부 정보에 대해 다음을 수행하세요.

   a. 인덱스 이름에 **`kendra-index`**를 입력합니다.

   b. 설명 필드는 비워 두세요.

   c. IAM 역할에서 새 역할 생성을 선택합니다. 이 역할은 Amazon S3 버킷에 대한 액세스를 제공합니다.

   d. 역할 이름에 **`kendra-role`**를 입력합니다. IAM 역할에는 `AmazonKendra-` 접두사가 붙습니다.

   e. 암호화 및 태그의 기본 설정을 유지하고 다음을 선택합니다.
4. 사용자 액세스 제어 구성 페이지의 액세스 제어 설정에서 아니오를 선택한 후 다음을 선택합니다.
5. 프로비저닝 세부 정보 페이지에 있는 프로비저닝 에디션의 경우 Developer Edition을 선택하고 생성을 선택합니다.

Amazon Kendra 인덱스를 생성하려면 (AWS CLI)

1. 신뢰할 수 있는 개체로 인식되는 Amazon Kendra용 IAM 역할을 생성하고 연결하려면 다음을 수행하세요.

   a. 다음 신뢰 정책을 로컬 디바이스의 텍스트 편집기에서 `kendra-trust-policy.json`이라는 JSON 파일로 저장합니다.

   ```
   {
   ```

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

b. kendra-role라는 IAM 역할을 생성하고 저장된 kendra-trust-policy.json 파일을 첨부하려면 create-role 명령을 사용합니다.

Linux

```
aws iam create-role \
          --role-name kendra-role \
          --assume-role-policy-document file://path/kendra-trust-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 kendra-trust-policy.json의 파일 경로입니다.

macOS

```
aws iam create-role \
          --role-name kendra-role \
          --assume-role-policy-document file://path/kendra-trust-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 kendra-trust-policy.json의 파일 경로입니다.

Windows

```
aws iam create-role ^
          --role-name kendra-role ^
          --assume-role-policy-document file://path/kendra-trust-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 kendra-trust-policy.json의 파일 경로입니다.

c. Amazon 리소스 이름(ARN)을 텍스트 편집기에 복사하고 로컬에 kendra-role-arn로 저장합니다.

> **Note**
>
> ARN 형식은 *arn:aws:iam::123456789012:role/kendra-role*과 비슷합니다. Amazon Kendra 작업을 실행하려면 kendra-role-arn로 저장한 ARN이 필요합니다.

2. 색인을 생성하기 전에 kendra-role CloudWatch Logs에 쓸 수 있는 권한을 제공해야 합니다. 이렇게 하려면 다음 단계를 완료하세요.

a. 다음 신뢰 정책을 로컬 디바이스의 텍스트 편집기에서 kendra-cloudwatch-policy.json이라는 JSON 파일로 저장합니다.

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

aws-region# 해당 지역과 AWS 12자리 계정 *aws-account-id*ID로 바꾸십시오. AWS

b. CloudWatch 로그에 액세스할 수 있는 IAM 정책을 생성하려면 create-policy 명령을 사용하십시오.

Linux

```
aws iam create-policy \
          --policy-name kendra-cloudwatch-policy \
          --policy-document file://path/kendra-cloudwatch-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 kendra-cloudwatch-policy.json의 파일 경로입니다.

macOS

```
aws iam create-policy \
          --policy-name kendra-cloudwatch-policy \
          --policy-document file://path/kendra-cloudwatch-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 kendra-cloudwatch-policy.json의 파일 경로입니다.

Windows

```
aws iam create-policy ^
          --policy-name kendra-cloudwatch-policy ^
          --policy-document file://path/kendra-cloudwatch-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 kendra-cloudwatch-policy.json의 파일 경로입니다.

c. Amazon 리소스 이름(ARN)을 텍스트 편집기에 복사하고 로컬에 kendra-cloudwatch-arn로 저장합니다.

> **Note**
>
> ARN의 형식은 arn:aws:iam: *:123456789012:role/#* 비슷합니다. kendra-cloudwatch-policy IAM 역할에 kendra-cloudwatch-policy를 연결하려면 kendra-cloudwatch-arn로 저장한 ARN이 필요합니다.

d. IAM 역할에 연결하려면 다음 명령을 사용하십시오. kendra-cloudwatch-policy attach-role-policy

Linux

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

위치:

- *policy-arn*은 저장한 kendra-cloudwatch-arn입니다.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

위치:

- *policy-arn*은 저장한 kendra-cloudwatch-arn입니다.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name kendra-role
```

위치:

- *policy-arn*은 저장한 kendra-cloudwatch-arn입니다.

3. 인덱스를 만들려면 create-index 명령을 사용하세요.

Linux

```
aws kendra create-index \
        --name kendra-index \
        --edition DEVELOPER_EDITION \
        --role-arn role-arn \
        --region aws-region
```

위치:

- *role-arn*은 저장한 kendra-role-arn입니다.
- *AWS-###* 사용자 지역입니다. AWS

macOS

```
aws kendra create-index \
        --name kendra-index \
        --edition DEVELOPER_EDITION \
        --role-arn role-arn \
        --region aws-region
```

위치:

- *role-arn*은 저장한 kendra-role-arn입니다.
- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra create-index ^
        --name kendra-index ^
        --edition DEVELOPER_EDITION ^
        --role-arn role-arn ^
        --region aws-region
```

위치:

- *role-arn*은 저장한 kendra-role-arn입니다.
- *AWS-###* 사용자 지역입니다. AWS

4. 인덱스 Id를 복사하여 텍스트 편집기에 kendra-index-id로 저장합니다. Id는 인덱스 생성 상태를 추적하는 데 도움이 됩니다.
5. 인덱스 생성 작업의 진행률을 추적하려면 describe-index 명령을 사용하세요.

Linux

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

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

macOS

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

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra describe-index ^
        --id kendra-index-id ^
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

인덱스 생성 프로세스는 평균 15분이 소요되지만 더 오래 걸릴 수 있습니다. 인덱스 상태가 활성 상태이면 인덱스를 사용할 준비가 된 것입니다. 인덱스가 생성되는 동안 다음 단계를 시작할 수 있습니다.

이 단계에서 를 사용하는 경우 S3 AWS CLI 버킷에 액세스할 수 있는 인덱스 권한을 부여하는 IAM 정책을 생성하여 Amazon Kendra IAM 역할에 연결합니다.

Amazon S3 액세스를 위한 IAM 역할 업데이트

인덱스를 생성하는 동안 Amazon Kendra IAM 역할을 업데이트하여 생성한 인덱스가 Amazon S3 버킷에서 데이터를 읽을 수 있도록 합니다. 자세한 내용을 알아보려면 Amazon Kendra의 IAM 액세스 역할을 참조하세요.

IAM 역할을 업데이트하려면 (콘솔)

1. https://console.aws.amazon.com/iam/에서 IAM 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 역할을 선택하고 역할 이름 위의 검색 상자에 **kendra-role**를 입력합니다.
3. 제안된 옵션 중에서 kendra-role를 클릭합니다.
4. 요약에서 정책 연결을 선택합니다.
5. 권한 연결의 검색 상자에 제안 옵션 중 ReadOnlyAccessAmazonS3 정책 옆의 확인란을 **S3** 입력하고 선택합니다.

6. 정책 연결을 선택합니다. 이제 요약 페이지에서 IAM 역할에 연결된 두 개의 정책을 볼 수 있습니다.
7. https://console.aws.amazon.com/kendra/에서 Amazon Kendra 콘솔로 돌아가 인덱스 상태가 생성 중에서 활성으로 변경될 때까지 기다린 후 다음 단계로 진행하세요.

IAM 역할을 업데이트하려면 (AWS CLI)

1. 다음 텍스트를 로컬 디바이스의 텍스트 편집기에서 `kendra-S3-access-policy.json`이라는 JSON 파일로 저장합니다.

```
{
   "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"
         ]
      }
```

```
    ]
}
```

DOC-EXAMPLE-BUCKET을 S3 버킷 이름으로, aws-region은 해당 지역, *12##* 계정 ID, AWS 그리고 저장된 이름으로 *aws-account-id*바꾸십시오. AWS *kendra-index-id*kendra-index-id

2. S3 버킷에 액세스하기 위한 IAM 정책을 생성하려면 create-policy 명령을 사용하세요.

Linux

```
aws iam create-policy \
          --policy-name kendra-S3-access-policy \
          --policy-document file://path/kendra-S3-access-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 kendra-S3-access-policy.json의 파일 경로입니다.

macOS

```
aws iam create-policy \
          --policy-name kendra-S3-access-policy \
          --policy-document file://path/kendra-S3-access-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 kendra-S3-access-policy.json의 파일 경로입니다.

Windows

```
aws iam create-policy ^
          --policy-name kendra-S3-access-policy ^
          --policy-document file://path/kendra-S3-access-policy.json
```

위치:

- *path/*는 로컬 디바이스에 있는 kendra-S3-access-policy.json의 파일 경로입니다.

3. Amazon 리소스 이름(ARN)을 텍스트 편집기에 복사하고 로컬에 kendra-S3-access-arn로 저장합니다.

> **Note**
>
> ARN의 형식은 *arn:aws:iam::123456789012:role/kendra-S3-access-policy*와 비슷한 형식입니다. IAM 역할에 kendra-S3-access-policy를 연결하려면 kendra-S3-access-arn로 저장한 ARN이 필요합니다.

4. 를 Amazon Kendra IAM 역할에 kendra-S3-access-policy 연결하려면 다음 명령을 사용하십시오. attach-role-policy

Linux

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

위치:

- *policy-arn*은 저장한 kendra-S3-access-arn입니다.

macOS

```
aws iam attach-role-policy \
          --policy-arn policy-arn \
          --role-name kendra-role
```

위치:

- *policy-arn*은 저장한 kendra-S3-access-arn입니다.

Windows

```
aws iam attach-role-policy ^
          --policy-arn policy-arn ^
          --role-name kendra-role
```

위치:

- *policy-arn*은 저장한 kendra-S3-access-arn입니다.

Amazon Kendra 사용자 지정 검색 인덱스 필드 생성

Amazon Kendra가 메타데이터를 사용자 지정 문서 속성으로 인식하도록 준비하려면 Amazon Comprehend 개체 유형에 해당하는 사용자 지정 필드를 생성합니다. 다음 9개의 Amazon Comprehend 개체 유형을 사용자 지정 필드로 입력합니다.

- 상업_품목
- 날짜
- 이벤트
- 위치
- 조직
- 기타
- 개인
- 수량
- 제목

> **Important**
> 철자가 틀린 개체 유형은 인덱스에서 인식되지 않습니다.

Amazon Kendra 인덱스를 위한 사용자 지정 필드를 만들려면 (콘솔)

1. https://console.aws.amazon.com/kendra/에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 목록에서 kendra-index를 클릭합니다.
3. 왼쪽 탐색 패널의 데이터 관리에서 패싯 정의를 선택합니다.
4. 인덱스 필드 메뉴에서 필드 추가를 선택합니다.
5. 인덱스 필드 추가 대화 상자에서 다음을 수행합니다.
   a. 필드 이름에 **COMMERCIAL_ITEM**을 입력합니다.

b. 데이터 유형에서 문자열 목록을 선택합니다.

c. 사용 유형에서 패싯 가능, 검색 가능, 표시 가능을 선택한 다음 추가를 선택합니다.

d. 각 Amazon Comprehend 개체 유형(상업_품목, 날짜, 이벤트, 위치, 조직, 기타, 개인, 수량, 제목)에 대해 a단계부터 c단계까지 반복합니다.

콘솔에는 필드 추가 성공 메시지가 표시됩니다. 다음 단계를 진행하기 전에 닫도록 선택할 수 있습니다.

Amazon Kendra 인덱스를 위한 사용자 지정 필드를 만들려면 (AWS CLI)

1. 다음 텍스트를 로컬 디바이스의 텍스트 편집기에서 `custom-attributes.json`이라는 JSON 파일로 저장합니다.

```
[
   {
       "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. 인덱스에 사용자 지정 필드를 만들려면 update-index 명령어를 사용하세요.

Linux

```
aws kendra update-index \
        --id kendra-index-id \
        --document-metadata-configuration-updates file://path/custom-
attributes.json \
        --region aws-region
```

위치:

- *kendra-index-id*저장되었나요? kendra-index-id
- *path/*는 로컬 디바이스에 있는 custom-attributes.json의 파일 경로입니다.
- *AWS-###* 사용자 지역입니다. AWS

macOS

```
aws kendra update-index \
        --id kendra-index-id \
        --document-metadata-configuration-updates file://path/custom-
attributes.json \
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *path/*는 로컬 디바이스에 있는 custom-attributes.json의 파일 경로입니다.

- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra update-index ^
        --id kendra-index-id ^
        --document-metadata-configuration-updates file://path/custom-
attributes.json ^
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *path/*는 로컬 디바이스에 있는 custom-attributes.json의 파일 경로입니다.
- *AWS-###* 사용자 지역입니다. AWS

3. 사용자 지정 속성이 인덱스에 추가되었는지 확인하려면 describe-index 명령을 사용하세요.

Linux

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

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

macOS

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

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id

- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra describe-index ^
        --id kendra-index-id ^
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

Amazon S3 버킷을 인덱스의 데이터 소스로 추가

인덱스를 동기화하려면 먼저 S3 데이터 소스를 인덱스에 연결해야 합니다.

S3 버킷을 Amazon Kendra 인덱스에 연결하는 방법 (콘솔)

1. https://console.aws.amazon.com/kendra/에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 목록에서 kendra-index를 클릭합니다.
3. 왼쪽 탐색 메뉴의 데이터 관리에서 데이터 소스를 선택합니다.
4. 데이터 소스 커넥터 유형 선택 섹션에서 Amazon S3로 이동한 다음 커넥터 추가를 선택합니다.
5. 데이터 소스 세부 정보 지정 페이지에서 다음 작업을 수행합니다.
   a. 이름 및 설명 아래 데이터 소스 이름에 **S3-data-source**를 입력합니다.
   b. 설명 섹션은 비워 두세요.
   c. 태그의 기본 설정을 유지하세요.
   d. 다음을 선택합니다.
6. 동기화 설정 구성 페이지의 동기화 범위 섹션에서 다음을 수행합니다.
   a. 데이터 소스 위치 입력에서 S3 찾아보기를 선택합니다.
   b. 리소스 선택에서 S3 버킷을 선택한 다음 선택을 선택합니다.
   c. 메타데이터 파일 접두사 폴더 위치에서 S3 찾아보기를 선택합니다.
   d. 리소스 선택의 경우, 버킷 목록에서 해당하는 버킷의 이름을 클릭합니다.

e. 객체의 경우 metadata 옵션 상자를 선택하고 선택을 선택합니다. 이제 위치 필드에 metadata/라고 표시됩니다.

f. 액세스 제어 목록 구성 파일 위치, 해독 키 선택 및 추가 구성에 대한 기본 설정을 유지합니다.

7. IAM 역할의 경우 동기화 설정 구성 페이지에서 kendra-role를 선택합니다.
8. 동기화 설정 구성 페이지의 동기화 실행 일정에서 빈도에 대해 요청 시 실행을 선택한 후 다음을 선택합니다.
9. 검토 및 생성 페이지에서 데이터 소스 세부 정보 선택 결과를 검토한 다음 데이터 소스 추가를 선택합니다.

S3 버킷을 Amazon Kendra 인덱스에 연결하는 방법 (AWS CLI)

1. 다음 텍스트를 로컬 디바이스의 텍스트 편집기에서 S3-data-connector.json이라는 JSON 파일로 저장합니다.

```
{
   "S3Configuration":{
      "BucketName":"DOC-EXAMPLE-BUCKET",
      "DocumentsMetadataConfiguration":{
         "S3Prefix":"metadata"
      }
   }
}
```

DOC-EXAMPLE-BUCKET을 S3 버킷의 이름으로 대체합니다.

2. S3 버킷을 인덱스에 연결하려면 다음 명령을 사용합니다. create-data-source

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
```

위치:

- *kendra-index-id*kendra-index-id저장되었습니까?
- *path/*는 로컬 디바이스에 있는 S3-data-connector.json의 파일 경로입니다.
- *role-arn*은 저장한 kendra-role-arn입니다.
- *AWS-###* 사용자 지역입니다. 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
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *path/*는 로컬 디바이스에 있는 S3-data-connector.json의 파일 경로입니다.
- *role-arn*은 저장한 kendra-role-arn입니다.
- *AWS-###* 사용자 지역입니다. 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
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *path/*는 로컬 디바이스에 있는 S3-data-connector.json의 파일 경로입니다.

- *role-arn*은 저장한 kendra-role-arn입니다.
- *AWS-###* 사용자 지역입니다. AWS

3. 커넥터 Id를 복사하여 텍스트 편집기에 S3-connector-id로 저장합니다. Id는 데이터 연결 프로세스의 상태를 추적하는 데 도움이 됩니다.
4. S3 데이터 소스가 성공적으로 연결되었는지 확인하려면 다음 명령을 사용하십시오. describe-data-source

Linux

```
aws kendra describe-data-source \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

위치:

- *S3-connector-id*는 저장한 S3-connector-id입니다.
- *kendra-index-id*kendra-index-id저장되었습니까?
- *AWS-###* 사용자 지역입니다. AWS

macOS

```
aws kendra describe-data-source \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

위치:

- *S3-connector-id*는 저장한 S3-connector-id입니다.
- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra describe-data-source ^
        --id S3-connector-id ^
```

```
        --index-id kendra-index-id ^
        --region aws-region
```

위치:

- *S3-connector-id*는 저장한 S3-connector-id입니다.
- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

이 단계가 끝나면 Amazon S3 데이터 소스가 인덱스에 연결됩니다.

Amazon Kendra 인덱스 동기화

Amazon S3 데이터 소스가 추가되었으므로 이제 Amazon Kendra 인덱스를 여기에 동기화합니다.

Amazon Kendra 인덱스를 동기화하려면 (콘솔)

1. https://console.aws.amazon.com/kendra/에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 목록에서 kendra-index를 클릭합니다.
3. 왼쪽 탐색 메뉴에서 데이터 소스를 선택합니다.
4. 데이터 소스에서S3-data-source를 선택합니다.
5. 상단 탐색 모음에서 지금 동기화를 선택합니다.

Amazon Kendra 인덱스를 동기화하려면 (AWS CLI)

1. 인덱스를 동기화하려면 start-data-source-sync-job 명령을 사용하십시오.

   Linux

   ```
   aws kendra start-data-source-sync-job \
           --id S3-connector-id \
           --index-id kendra-index-id \
           --region aws-region
   ```

   위치:

   - *S3-connector-id*는 저장한 S3-connector-id입니다.
   - *kendra-index-id*kendra-index-id저장되었나요?

- *AWS-###* 사용자 지역입니다. AWS

macOS

```
aws kendra start-data-source-sync-job \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

위치:

- *S3-connector-id*는 저장한 S3-connector-id입니다.
- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra start-data-source-sync-job ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

위치:

- *S3-connector-id*는 저장한 S3-connector-id입니다.
- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

2. 인덱스 동기화 상태를 확인하려면 list-data-source-sync-jobs 명령을 사용합니다.

Linux

```
aws kendra list-data-source-sync-jobs \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

위치:

- *S3-connector-id*는 저장한 S3-connector-id입니다.
- *kendra-index-id*kendra-index-id저장되었습니까?
- *AWS-###* 사용자 지역입니다. AWS

macOS

```
aws kendra list-data-source-sync-jobs \
        --id S3-connector-id \
        --index-id kendra-index-id \
        --region aws-region
```

위치:

- *S3-connector-id*는 저장한 S3-connector-id입니다.
- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra list-data-source-sync-jobs ^
        --id S3-connector-id ^
        --index-id kendra-index-id ^
        --region aws-region
```

위치:

- *S3-connector-id*는 저장한 S3-connector-id입니다.
- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

이 단계가 끝나면 데이터 세트에 대해 검색 및 필터링 가능한 Amazon Kendra 인덱스 생성이 완료됩니다.

5단계: Amazon Kendra 인덱스 쿼리

이제 Amazon Kendra 인덱스를 자연어 쿼리에 사용할 준비가 되었습니다. 인덱스를 검색할 때 Amazon Kendra는 사용자가 제공한 모든 데이터와 메타데이터를 사용하여 검색 쿼리에 가장 정확한 답변을 제공합니다.

Amazon Kendra가 응답할 수 있는 쿼리에는 세 가지 종류가 있습니다.

- 팩토이드 쿼리("누가", "무엇을", "언제", "어디에서" 질문)
- 서술형 쿼리("어떻게" 질문)
- 키워드 검색(의도와 범위가 명확하지 않은 질문)

주제

Amazon Kendra 인덱스 쿼리

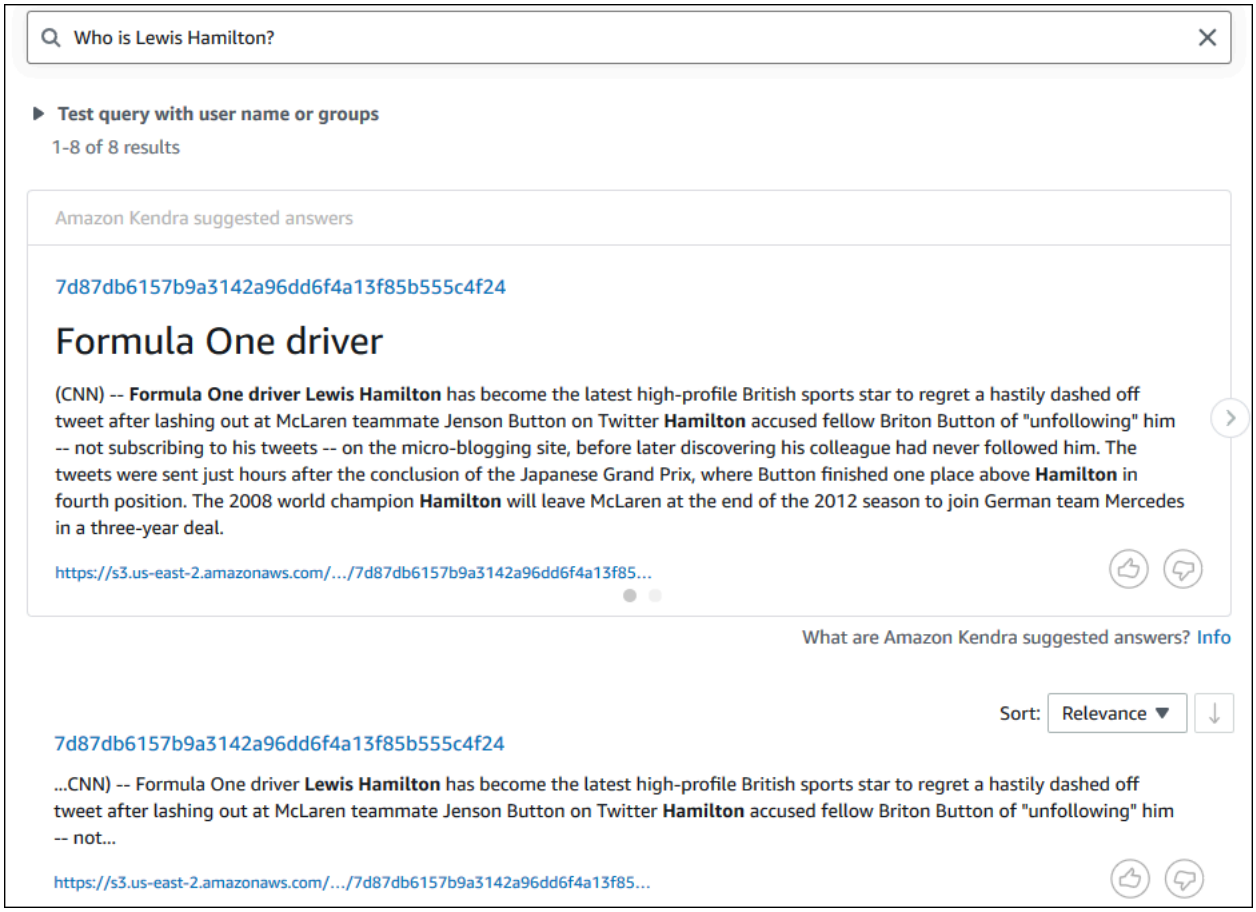
Amazon Kendra가 지원하는 세 가지 쿼리 유형에 해당하는 질문을 사용하여 Amazon Kendra 인덱스를 쿼리할 수 있습니다. 자세한 내용은 쿼리를 참조하세요.

이 섹션의 예제 질문은 샘플 데이터 세트를 기반으로 선택되었습니다.

Amazon Kendra 인덱스를 쿼리하려면 (콘솔)

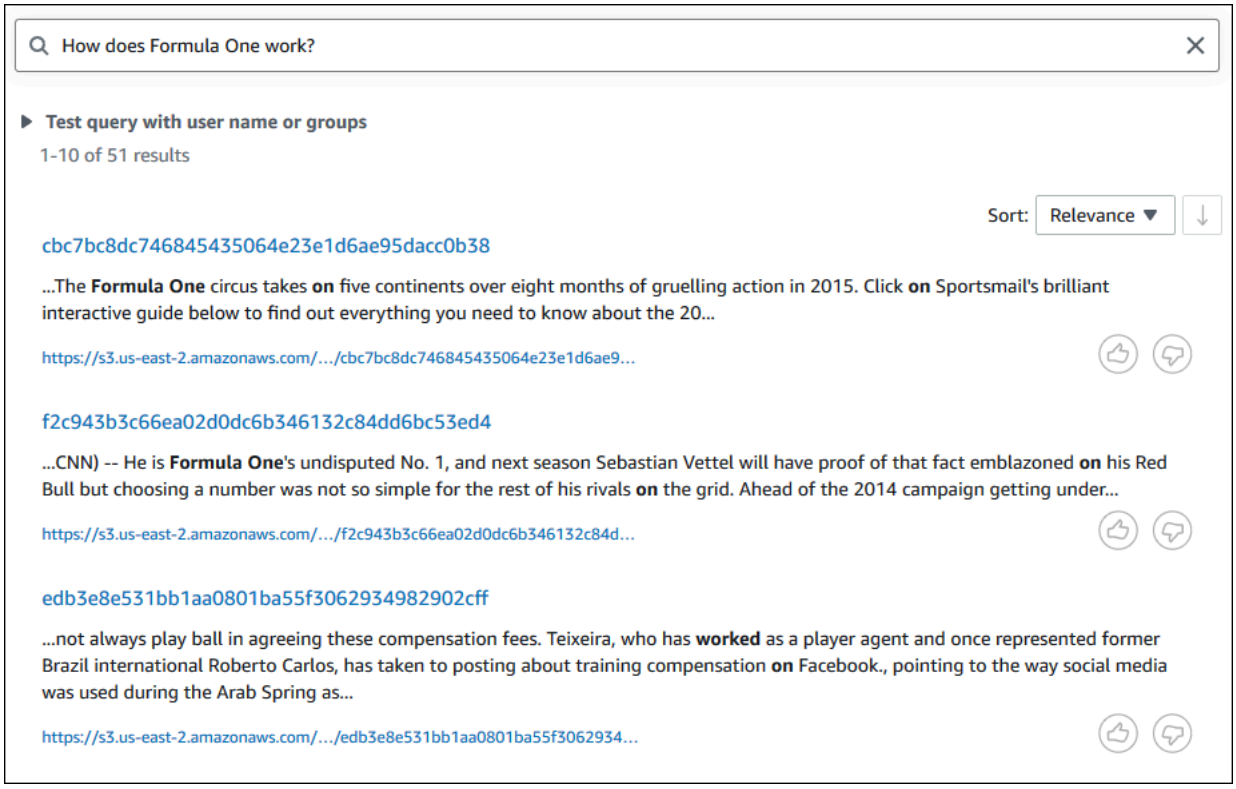
1. https://console.aws.amazon.com/kendra/에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 목록에서 `kendra-index`를 클릭합니다.
3. 왼쪽 탐색 메뉴에서 인덱스 검색 옵션을 선택합니다.
4. 샘플 팩토이드 쿼리를 실행하려면 검색 상자에 **`Who is Lewis Hamilton?`**을 입력하고 Enter 키를 누릅니다.

   반환된 첫 번째 결과는 Amazon Kendra 제안 답변이며, 답변이 포함된 데이터 파일이 포함됩니다. 나머지 결과는 권장 문서 세트를 구성합니다.



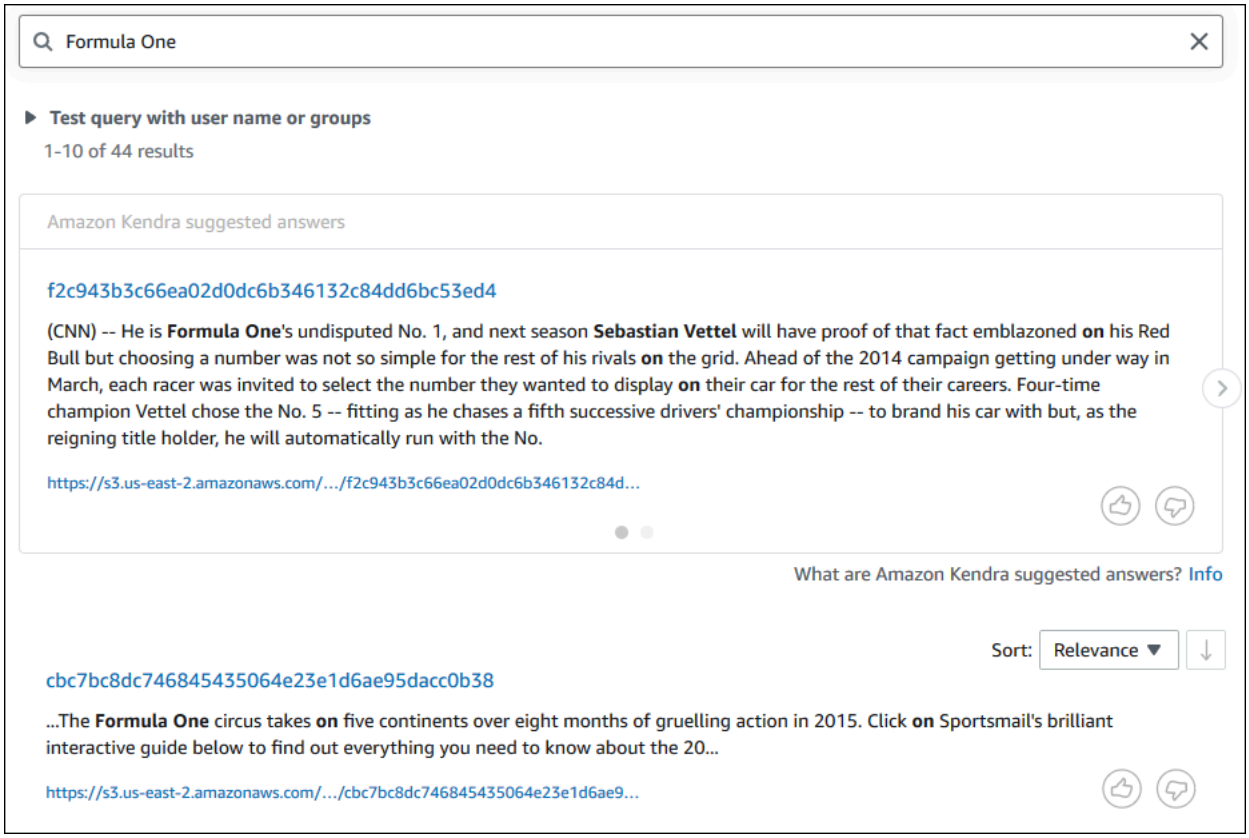
5. 서술적 쿼리를 실행하려면 검색 상자에 **`How does Formula One work?`**을 입력하고 Enter 키를 누릅니다.

   Amazon Kendra 콘솔에서 반환된 또 다른 결과를 확인할 수 있습니다. 이번에는 관련 문구가 강조 표시되어 있습니다.



6. 키워드 검색을 실행하려면 검색 상자에 **`Formula One`**을 입력하고 Enter 키를 누릅니다.

   Amazon Kendra 콘솔에서 반환한 또 다른 결과와 데이터 세트에 있는 해당 구문에 대한 다른 모든 언급의 결과가 표시됩니다.



Amazon Kendra 인덱스를 쿼리하려면 (AWS CLI)

1. 샘플 팩토이드 쿼리를 실행하려면 query 명령을 사용하세요.

   Linux

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Who is Lewis Hamilton?" \
           --region aws-region
   ```

   위치:

   - *kendra-index-id*저장하셨나요? kendra-index-id
   - *AWS-###* 사용자 지역입니다. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Who is Lewis Hamilton?" \
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Who is Lewis Hamilton?" ^
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

쿼리 결과가 AWS CLI 표시됩니다.

2. 샘플 서술적 쿼리를 실행하려면 query 명령을 사용하세요.

Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "How does Formula One work?" \
        --region aws-region
```

위치:

- *kendra-index-id*저장되었습니다kendra-index-id.

- *AWS-###* 사용자 지역입니다. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "How does Formula One work?" \
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "How does Formula One work?" ^
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

쿼리 결과가 AWS CLI 표시됩니다.

3. 샘플 키워드 검색을 실행하려면 query 명령을 사용하세요.

Linux

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Formula One" \
        --region aws-region
```

위치:

- *kendra-index-id*저장되었습니다kendra-index-id.
- *AWS-###* 사용자 지역입니다. AWS

macOS

```
aws kendra query \
        --index-id kendra-index-id \
        --query-text "Formula One" \
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

Windows

```
aws kendra query ^
        --index-id kendra-index-id ^
        --query-text "Formula One" ^
        --region aws-region
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

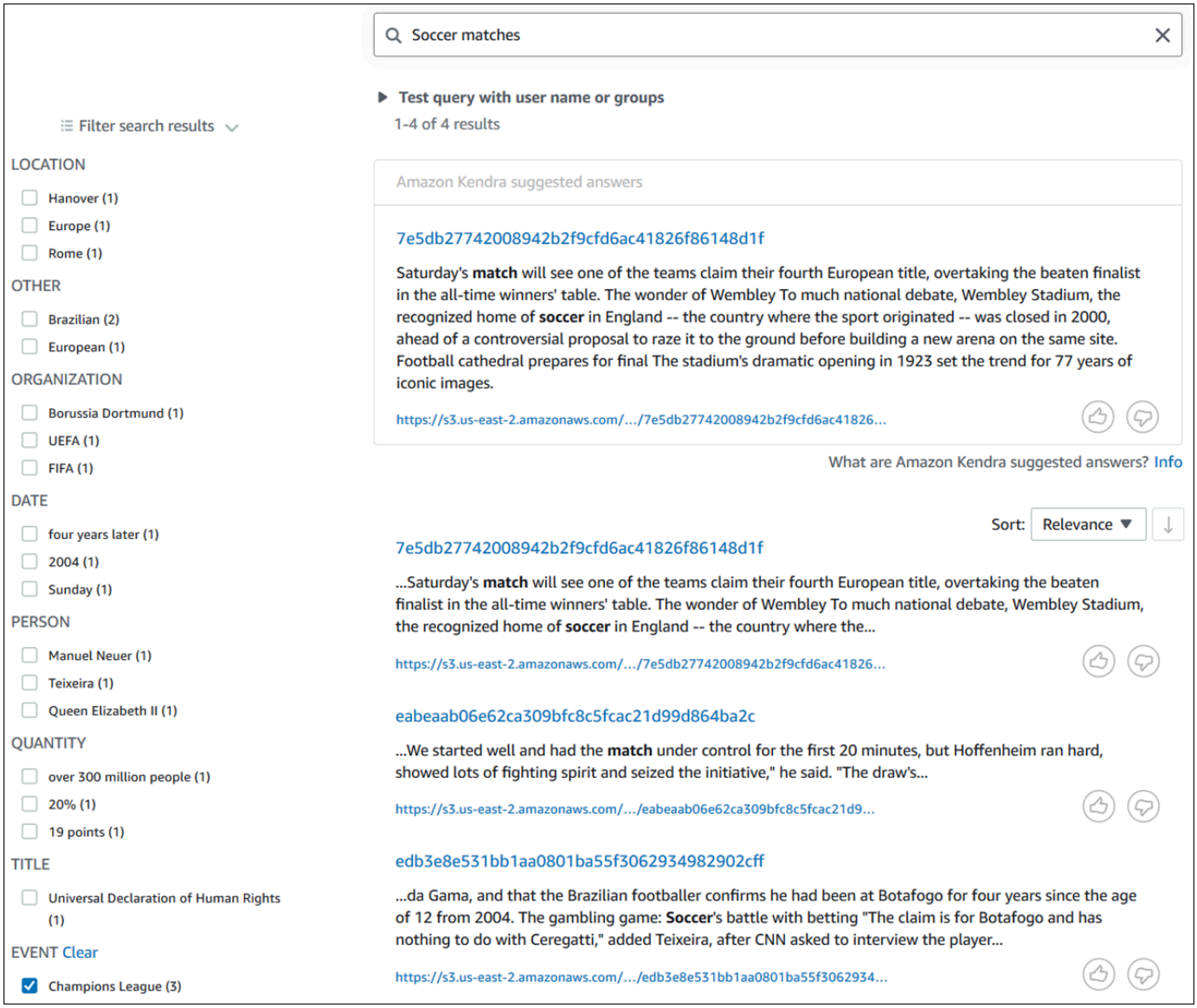
AWS CLI 에는 쿼리에 대해 반환된 답변이 표시됩니다.

검색 결과 필터링

Amazon Kendra 콘솔에서 사용자 지정 문서 속성을 사용하여 검색 결과를 필터링하고 정렬할 수 있습니다. Amazon Kendra가 쿼리를 처리하는 방법에 대한 자세한 내용은 쿼리 필터링을 참조하세요.

검색 결과를 필터링하려면 (콘솔)

1. https://console.aws.amazon.com/kendra/에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 목록에서 `kendra-index`를 클릭합니다.
3. 왼쪽 탐색 메뉴에서 인덱스 검색 옵션을 선택합니다.
4. 검색 상자에 **`Soccer matches`**를 쿼리로 입력하고 Enter 키를 누릅니다.
5. 왼쪽 탐색 메뉴에서 검색 결과 필터링을 선택하면 검색을 필터링하는 데 사용할 수 있는 패싯 목록이 표시됩니다.
6. 이벤트 부제목 아래의 “Champions League” 확인란을 선택하면 “Champions League”가 포함된 결과로만 필터링된 검색 결과를 볼 수 있습니다.



검색 결과를 필터링하려면 (AWS CLI)

1. 검색에 사용할 수 있는 특정 유형(예:EVENT)의 개체를 보려면 query 명령어를 사용하세요.

   Linux

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Soccer matches" \
           --facets '[{"DocumentAttributeKey":"EVENT"}]' \
           --region aws-region
   ```

   위치:

   - *kendra-index-id*저장하셨나요kendra-index-id?
   - *AWS-###* 사용자 지역입니다. AWS

   macOS

   ```
   aws kendra query \
           --index-id kendra-index-id \
           --query-text "Soccer matches" \
           --facets '[{"DocumentAttributeKey":"EVENT"}]' \
           --region aws-region
   ```

   위치:

   - *kendra-index-id*저장하셨나요? kendra-index-id
   - *AWS-###* 사용자 지역입니다. AWS

   Windows

   ```
   aws kendra query ^
           --index-id kendra-index-id ^
           --query-text "Soccer matches" ^
           --facets '[{"DocumentAttributeKey":"EVENT"}]' ^
           --region aws-region
   ```

   위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

검색 결과가 AWS CLI 표시됩니다. 유형의 EVENT 패싯 목록을 가져오려면 AWS CLI 출력의 "FacetResults" 섹션으로 이동하여 필터링 가능한 패싯 목록과 해당 개수를 확인하십시오. 예를 들어, 패싯 중 하나는 “Champions League”입니다.

> **Note**
>
> EVENT 대신, DocumentAttributeKey 값에 대해 the section called “Amazon Kendra 인덱스를 생성합니다.”에서 생성한 인덱스 필드 중 하나를 선택할 수 있습니다.

2. 동일한 검색을 실행하되 “Champions League”가 포함된 결과로만 필터링하려면 query 명령어를 사용하세요.

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
```

위치:

- *kendra-index-id*저장하셨습니까? kendra-index-id
- *AWS-###* 사용자 지역입니다. 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
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. 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
```

위치:

- *kendra-index-id*저장하셨나요? kendra-index-id
- *AWS-###* 사용자 지역입니다. AWS

필터링된 검색 결과를 AWS CLI 표시합니다.

6단계: 정리

파일 정리

이 자습서를 완료한 후 AWS 계정에서 요금이 발생하지 않도록 하려면 다음 단계를 수행하면 됩니다.

1. Amazon S3 버킷 삭제

   버킷 삭제에 대한 자세한 내용은 버킷 삭제를 참조하세요.

2. Amazon Kendra 인덱스 삭제

   Amazon Kendra 인덱스 삭제에 대한 자세한 내용은 인덱스 삭제를 참조하세요.

3. **converter.py** 삭제

   - 콘솔의 경우: 로 AWS CloudShell이동하여 지역이 해당 지역으로 설정되어 있는지 확인하십시오. AWS bash 쉘이 로드되면 환경에 다음 명령을 입력하고 Enter 키를 누릅니다.

```
rm converter.py
```

- 대상 AWS CLI: 터미널 창에서 다음 명령을 실행합니다.

  Linux

  ```
  rm file/converter.py
  ```

  위치:

  - *file/*는 로컬 디바이스에 있는 converter.py의 파일 경로입니다.

  macOS

  ```
  rm file/converter.py
  ```

  위치:

  - *file/*는 로컬 디바이스에 있는 converter.py의 파일 경로입니다.

  Windows

  ```
  rm file/converter.py
  ```

  위치:

  - *file/*는 로컬 디바이스에 있는 converter.py의 파일 경로입니다.

자세히 알아보기

Amazon Kendra를 워크플로에 통합하는 방법에 대해 자세히 알아보려면 다음 블로그 게시물을 참조하세요.

- 향상된 검색을 위한 콘텐츠 메타데이터 태깅
- 자동화된 콘텐츠 보강 기능을 갖춘 지능형 검색 솔루션 구축

Amazon Comprehend에 대한 자세한 내용은 Amazon Comprehend 개발자 가이드를 참조하세요.

Amazon Kendra에 대한 모니터링 및 로깅

주제

인덱스 모니터링 (콘솔)

Amazon Kendra 콘솔을 사용하여 인덱스 및 데이터 소스의 상태를 모니터링할 수 있습니다. 이 정보를 사용하여 인덱스의 크기 및 스토리지 요구 사항을 추적하고 인덱스와 데이터 소스 간의 동기화 진행 상황 및 성공 여부를 모니터링할 수 있습니다.

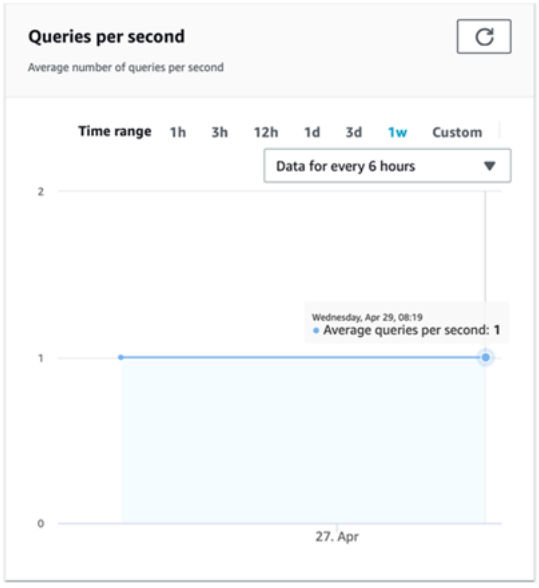
인덱스 지표를 보려면 (콘솔)

1. AWS Management Console에 로그인하고 https://console.aws.amazon.com/kendra/home에서 Amazon Kendra 콘솔을 엽니다.
2. 인덱스 목록에서 확인할 인덱스를 선택합니다.
3. 화면을 스크롤하여 인덱스 지표를 확인합니다.

인덱스에 대한 다음 지표를 확인할 수 있습니다.

- 문서 수 - 인덱싱된 총 문서 수입니다. 여기에는 모든 데이터 소스의 모든 문서가 포함됩니다. 이 지표를 사용하여 인덱스에 사용할 스토리지 단위를 더 많이 구매해야 하는지 아니면 더 적게 구매해야 하는지 결정하세요.

- 초당 쿼리 - 초당 요청되는 인덱스 쿼리 수입니다. 이 지표를 사용하여 인덱스에 사용할 쿼리 단위를 더 많이 구매해야 하는지 아니면 더 적게 구매해야 하는지 결정하세요.



인덱스와 데이터 소스 간의 동기화 진행 상황과 성공 여부를 모니터링하려면 Amazon Kendra 콘솔을 사용하세요. 이 정보를 사용하면 데이터 소스의 상태를 확인하는 데 도움이 됩니다.

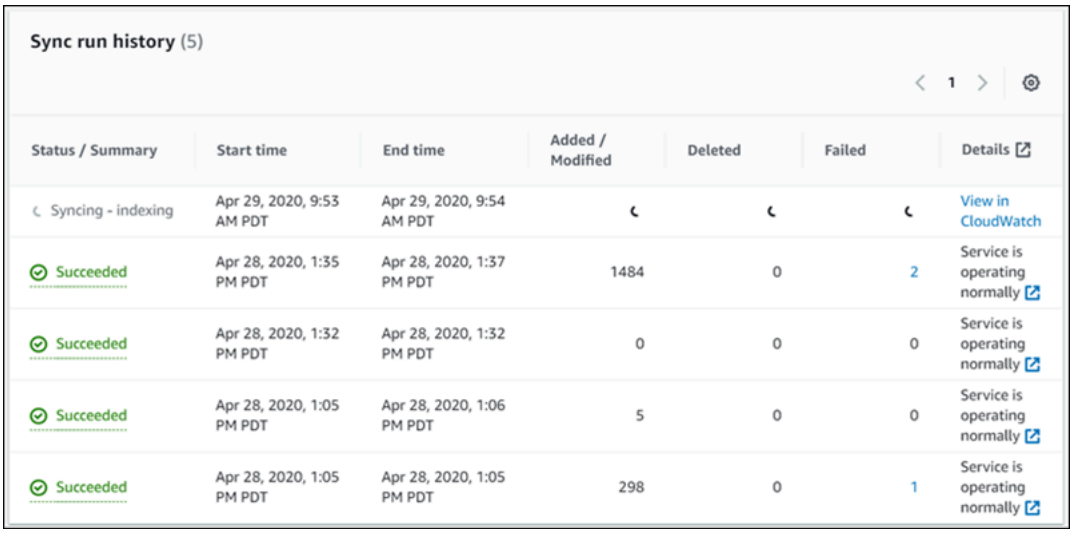
동기화 지표를 보려면 (콘솔)

1. AWS Management Console에 로그인하고 https://console.aws.amazon.com/kendra/home에서 Amazon Kendra 콘솔을 엽니다.

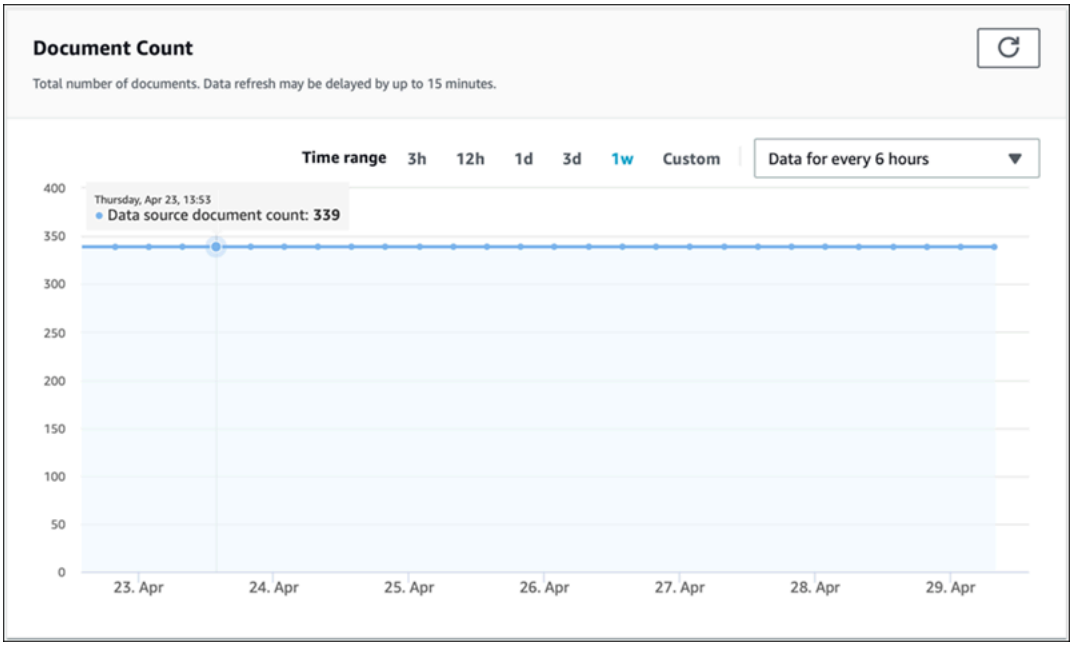
2. 인덱스 목록에서 동기화 지표를 확인할 인덱스를 선택합니다.
3. 왼쪽 메뉴에서 데이터 소스를 선택합니다.
4. 데이터 소스 목록에서 확인할 데이터 소스를 선택합니다.
5. 화면을 스크롤하여 동기화 실행 지표를 확인합니다.

다음 정보를 볼 수 있습니다.

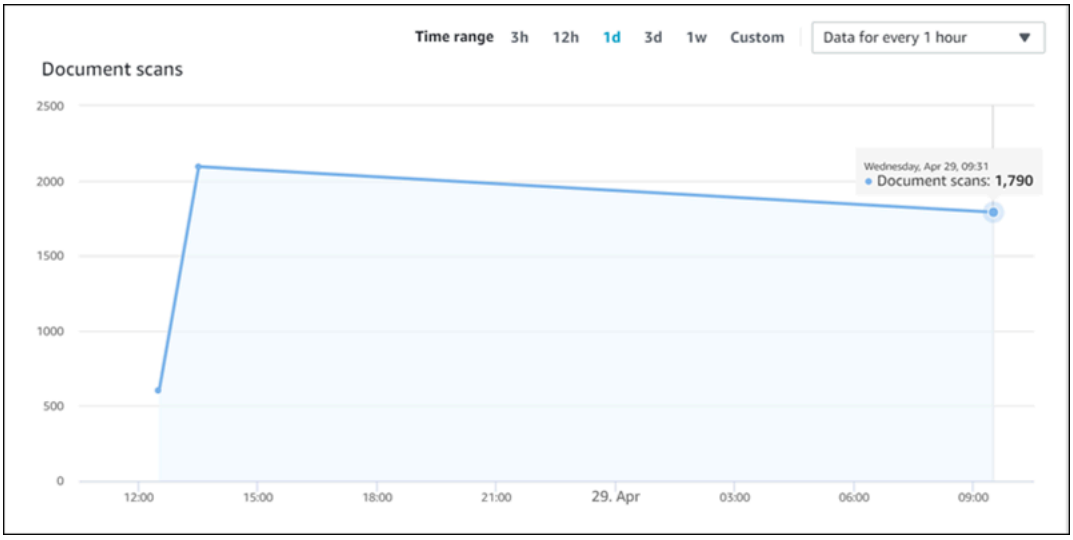
- 동기화 실행 기록 - 시작 및 종료 시간, 추가, 삭제 및 실패한 문서 수를 포함하여 동기화 실행에 대한 통계입니다. 동기화 실행이 실패할 경우 자세한 정보가 포함된 CloudWatch Logs 링크가 있습니다. 왼쪽 상단의 설정 아이콘을 선택하여 기록에 표시되는 열을 변경합니다. 이 정보를 사용하여 데이터 소스의 전반적인 상태를 확인할 수 있습니다.



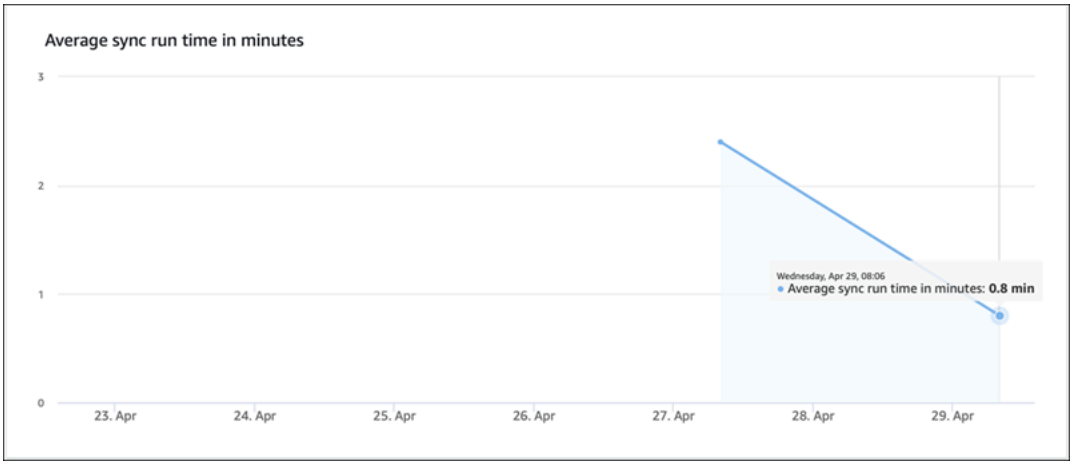
- 문서 수 - 이 데이터 소스에서 인덱싱된 총 문서 수입니다. 데이터 소스에 추가된 모든 문서의 합계에서 데이터 소스에서 삭제된 모든 문서의 총계를 뺀 값입니다. 이 정보를 사용하여 인덱스에 포함되는 이 데이터 소스의 문서 수를 확인할 수 있습니다.



- 문서 스캔 - 동기화 실행 중에 스캔한 총 문서 수입니다. 여기에는 추가, 업데이트, 삭제 또는 변경되지 않은 문서를 포함하여 데이터 소스의 모든 문서가 포함됩니다. 이 정보를 사용하여 Amazon Kendra가 데이터 소스의 모든 문서를 스캔하고 있는지 확인할 수 있습니다. 스캔한 문서 수는 서비스에 청구되는 금액에 영향을 줍니다.



- 평균 동기화 실행 시간 (분) - 동기화 실행이 완료되는 데 걸리는 평균 시간입니다. 데이터 소스를 동기화하는 데 걸리는 시간은 서비스에 청구되는 금액에 영향을 줍니다.



AWS CloudTrail 로그를 사용하여 Amazon Kendra API 호출 로깅

Amazon Kendra는 Amazon Kendra에서 사용자, 역할 또는 AWS 서비스가 수행한 작업에 대한 레코드를 제공하는 서비스인 AWS CloudTrail과 통합됩니다. CloudTrail은 Amazon Kendra 콘솔의 호출과 Amazon Kendra API에 대한 코드 호출을 포함하여, Amazon Kendra의 모든 API 호출을 이벤트로 캡처합니다. 추적을 생성하면 Amazon Kendra 이벤트를 포함한 CloudTrail 이벤트를 지속적으로 Amazon S3 버킷에 배포할 수 있습니다. 추적을 구성하지 않은 경우에도 CloudTrail 콘솔의 이벤트 기록에서 최신 이벤트를 볼 수 있습니다. CloudTrail에서 수집한 정보를 사용하여 Amazon Kendra에 수행된 요청, 요청이 수행된 IP 주소, 요청을 수행한 사람, 요청이 수행된 시간 및 추가 세부 정보를 확인할 수 있습니다.

구성 및 활성화 방법을 비롯한 CloudTrail에 대한 자세한 내용은 AWS CloudTrail 사용 설명서를 참조하세요.

CloudTrail의 Amazon Kendra 정보

CloudTrail은 계정 생성 시 AWS 계정에서 활성화됩니다. Amazon Kendra에서 활동이 발생하면 해당 활동이 CloudTrail 이벤트 기록의 다른 AWS 서비스 이벤트와 함께 CloudTrail 이벤트에 기록됩니다. AWS 계정에서 최신 이벤트를 확인, 검색 및 다운로드할 수 있습니다. 자세한 내용은 CloudTrail 이벤트 기록을 사용하여 이벤트 보기를 참조하세요.

Amazon Kendra에 대한 이벤트를 포함하여 AWS 계정의 이벤트의 지속적인 레코드의 경우, 추적을 생성합니다. 추적이란 CloudTrail에서 지정된 S3 버킷에 이벤트를 로그 파일로 제공할 수 있도록 허용하는 구성입니다. 콘솔에서 추적을 생성하면 기본적으로 모든 AWS 리전에 추적이 적용됩니다. 추적은 AWS 파티션에 있는 모든 리전의 이벤트를 로깅하고 지정된 S3 버킷으로 로그 파일을 전송합니다. 또는 CloudTrail 로그에서 수집된 이벤트 데이터를 추가 분석 및 처리하도록 다른 AWS 서비스를 구성할 수 있습니다. 자세한 정보는 다음을 참조하세요.

- 추적 생성 개요
- CloudTrail 지원 서비스 및 통합
- CloudTrail에서 Amazon SNS 알림 구성
- 여러 리전에서 CloudTrail 로그 파일 받기 및 여러 계정에서 CloudTrail 로그 파일 받기

CloudTrail은 모든 Amazon Kendra 작업을 기록하며, 이는 API 참조에 문서화되어 있습니다. 예를 들어, `CreateIndex`, `CreateDataSource` 및 `Query` 작업에 대한 호출은 CloudTrail 로그 파일의 항목을 생성합니다.

모든 이벤트 및 로그 항목에는 요청을 생성한 사용자에 대한 정보가 들어 있습니다. 자세한 내용은 CloudTrail userIdentity 요소를 참조하세요.

예: Amazon Kendra 로그 파일 항목

추적이란 지정한 S3 버킷에 이벤트를 로그 파일로 전달하도록 허용하는 구성입니다. CloudTrail 로그 파일에는 하나 이상의 로그 항목이 포함될 수 있습니다. 이벤트는 모든 소스로부터의 단일 요청을 나타내며 요청 작업, 작업 날짜와 시간, 요청 파라미터 등에 대한 정보가 들어 있습니다. CloudTrail 로그 파일은 퍼블릭 API 호출의 주문 스택 트레이스가 아니므로 특정 순서로 표시되지 않습니다.

`Query` 작업을 호출하면 다음 항목이 생성됩니다.

```
       {
           "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"
    },
```

Amazon Kendra Intelligent Ranking API 호출을 AWS CloudTrail 로그와 함께 로깅

Amazon Kendra Intelligent Ranking은 Amazon Kendra Intelligent Ranking에서 사용자, 역할 또는 AWS 서비스가 수행한 작업에 대한 레코드를 제공하는 서비스인 AWS CloudTrail과 통합됩니다. CloudTrail은 Amazon Kendra Intelligent Ranking API에 대한 코드 호출을 포함하여 Amazon Kendra Intelligent Ranking에서 발생하는 모든 API 호출을 이벤트로 캡처합니다. 추적을 생성하면 Amazon Kendra Intelligent Ranking 이벤트를 포함한 CloudTrail 이벤트를 지속적으로 Amazon S3 버킷에 배포할 수 있습니다. 추적을 구성하지 않은 경우에도 CloudTrail 콘솔의 이벤트 기록에서 최신 이벤트를 볼 수 있습니다. CloudTrail에서 수집한 정보를 사용하여 Amazon Kendra Intelligent Ranking에 수행된 요청, 요청이 수행된 IP 주소, 요청을 수행한 사람, 요청이 수행된 시간 및 추가 세부 정보를 확인할 수 있습니다.

구성 및 활성화 방법을 비롯한 CloudTrail에 대한 자세한 내용은 AWS CloudTrail 사용 설명서를 참조하세요.

CloudTrail의 Amazon Kendra Intelligent Ranking 정보

CloudTrail은 계정 생성 시 AWS 계정에서 활성화됩니다. Amazon Kendra Intelligent Ranking에서 활동이 발생하면 해당 활동이 CloudTrail 이벤트 기록의 다른 AWS 서비스 이벤트와 함께 CloudTrail 이벤트에 기록됩니다. AWS 계정에서 최신 이벤트를 확인, 검색 및 다운로드할 수 있습니다. 자세한 내용은 CloudTrail 이벤트 기록을 사용하여 이벤트 보기를 참조하세요.

Amazon Kendra Intelligent Ranking에 대한 이벤트를 포함하여 AWS 계정의 이벤트의 지속적인 레코드의 경우, 추적을 생성합니다. 추적이란 CloudTrail에서 지정된 S3 버킷에 이벤트를 로그 파일로 제공할 수 있도록 허용하는 구성입니다. 콘솔에서 추적을 생성하면 기본적으로 모든 AWS 리전에 추적이 적용됩니다. 추적은 AWS 파티션에 있는 모든 리전의 이벤트를 로깅하고 지정된 S3 버킷으로 로그 파일을 전송합니다. 또는 CloudTrail 로그에서 수집된 이벤트 데이터를 추가 분석 및 처리하도록 다른 AWS 서비스를 구성할 수 있습니다. 자세한 정보는 다음을 참조하세요.

- 추적 생성 개요
- CloudTrail 지원 서비스 및 통합
- CloudTrail에서 Amazon SNS 알림 구성
- 여러 리전에서 CloudTrail 로그 파일 받기 및 여러 계정에서 CloudTrail 로그 파일 받기

CloudTrail은 모든 Amazon Kendra Intelligent Ranking 작업을 기록하며, 이는 API 참조에 문서화되어 있습니다. 예를 들어 `CreateRescoreExecutionPlan`를 호출하면 CloudTrail 로그 파일에 항목이 생성됩니다.

모든 이벤트 및 로그 항목에는 요청을 생성한 사용자에 대한 정보가 들어 있습니다. 자세한 내용은 CloudTrail userIdentity 요소를 참조하세요.

예: Amazon Kendra Intelligent Ranking 로그 파일 항목

추적이란 지정한 S3 버킷에 이벤트를 로그 파일로 전달하도록 허용하는 구성입니다. CloudTrail 로그 파일에는 하나 이상의 로그 항목이 포함될 수 있습니다. 이벤트는 모든 소스로부터의 단일 요청을 나타내며 요청 작업, 작업 날짜와 시간, 요청 파라미터 등에 대한 정보가 들어 있습니다. CloudTrail 로그 파일은 퍼블릭 API 호출의 주문 스택 트레이스가 아니므로 특정 순서로 표시되지 않습니다.

`CreateRescoreExecutionPlan` 작업을 호출하면 다음 항목이 생성됩니다.

```
{
            "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"
        }
    }
```

Amazon CloudWatch를 사용한 Amazon Kendra 모니터링

인덱스 상태를 추적하려면 Amazon CloudWatch를 사용하세요. CloudWatch를 사용하면 인덱스의 문서 동기화에 대한 지표를 얻을 수 있습니다. 또한 하나 이상의 지표가 정의한 임계값을 초과하는 경우 알리도록 경보를 설정할 수도 있습니다. 예를 들어 인덱싱하도록 제출된 문서 수 또는 인덱싱에 실패한 문서 수를 모니터링할 수 있습니다.

CloudWatch를 사용하여 Amazon Kendra를 모니터링하려면 적절한 권한이 있어야 합니다. 자세한 내용은 Amazon CloudWatch 사용 설명서의 Amazon CloudWatch에 대한 인증 및 액세스 제어를 참조하세요.

Amazon Kendra 지표 보기

CloudWatch 콘솔을 사용한 Amazon Kendra 지표 확인

지표를 보려면(CloudWatch 콘솔)

1. AWS Management Console에 로그인하고 https://console.aws.amazon.com/cloudwatch/에서 CloudWatch 콘솔을 엽니다.
2. 지표, 모든 지표, Kendra를 순서대로 선택합니다.
3. 차원과 측정치 이름을 선택한 다음 그래프에 추가를 선택합니다.
4. 날짜 범위 값을 선택합니다. 선택한 날짜 범위에 대한 측정치 개수가 그래프에 표시됩니다.

경보 생성

CloudWatch 경보는 지정한 기간 동안 단일 지표를 감시하고, Amazon Simple Notification Service(SNS) top 또는 Auto Scaling 정책에 알림 보내기와 같은 하나 이상의 작업을 수행합니다. 이러한 작업은 지정한 여러 기간 동안 지정된 임계값에 따른 지표의 값을 기반으로 합니다. 또한 CloudWatch는 경보로 인해 상태가 변경되면 Amazon SNS 메시지를 전송할 수 있습니다.

경보는 상태가 변경되어 지정한 기간 동안 지속되는 경우에만 작업을 호출합니다.

경보를 설정하려면

1. AWS Management Console에 로그인하고 https://console.aws.amazon.com/cloudwatch/에서 CloudWatch 콘솔을 엽니다.
2. 경보를 선택한 다음 경보 생성 버튼을 선택합니다.
3. 지표를 선택합니다. 인덱스 및 데이터 소스에 대한 Kendra 지표를 선택하세요. 또한 시간을 설정된 시간, 일, 주 또는 사용자 지정 수로 설정하세요.
4. 통계를 선택하세요. 예: 평균. 또한 경보 트리거 기간을 설정된 분, 시간, 일별 또는 사용자 지정으로 선택하세요.
5. 경보를 트리거할 임계값, 정적 값 또는 대역 사용 여부, 임계값 충족 조건 등을 선택합니다.
6. 지표가 설정된 임계값을 벗어나야 하는지 아니면 다른 상태인지 등 트리거의 경보 상태를 선택합니다. 경보를 받을 사람/이메일을 선택합니다.
7. 경보가 만족스러우면 경보 생성을 선택합니다.

> **Note**
>
> CloudWatch 경보의 이름을 제공해야 합니다.

인덱스 동기화 작업을 위한 CloudWatch 지표

다음 표에서는 데이터 소스 동기화 작업에 대한 Amazon Kendra 지표를 설명합니다.

API 또는 CLI를 사용하는 경우 GetMetricStatistics API를 사용할 때 선택한 `MetricName` 외에도 `Namespace`를 'AWS/Kendra'로 지정해야 합니다.

지표	설명
`DocumentsCrawled`	동기화 작업이 실행 중에 스캔하거나 발견한 문서 수입니다. 차원: • IndexId • DataSourceId

지표	설명
	단위: 수
DocumentsSubmittedForIndexing	동기화 작업이 인덱스에 제출한 문서 수. 차원: • IndexId • DataSourceId 단위: 수
DocumentsSubmittedForIndexingFailed	인덱싱에 실패한 문서 수. 자세한 내용은 동기화 작업에 대한 CloudWatch 로그의 내용을 확인하세요. 차원: • IndexId • DataSourceId 단위: 수
DocumentsSubmittedForDeletion	동기화 작업에서 인덱스에서 제거하도록 요청한 문서 수. 차원: • IndexId • DataSourceId 단위: 수

지표	설명
DocumentsSubmittedForDeletionFailed	삭제에 실패한 문서 수. 자세한 내용은 동기화 작업에 대한 CloudWatch 로그의 내용을 확인하세요. 차원: • IndexId • DataSourceId 단위: 수

Amazon Kendra 데이터 소스에 대한 지표

다음 표에서는 데이터 소스 동기화 작업에 대한 Amazon Kendra 지표를 설명합니다. 별표(*)로 표시된 지표는 Amazon S3 데이터 소스에 대해서만 사용됩니다.

API 또는 CLI를 사용하는 경우 GetMetricStatistics API를 사용할 때 선택한 MetricName 외에도 Namespace를 'AWS/Kendra'로 지정해야 합니다.

지표	설명
DocumentsSkippedNoChange *	검토한 결과 변경되지 않은 것으로 확인되어 인덱싱을 위해 제출되지 않은 문서 수입니다. 차원: • IndexId • DataSourceId 단위: 수
DocumentsSkippedInvalidMetadata *	관련 메타데이터 파일에 문제가 발생하여 건너뛰었던 문서 수입니다. 자세한 내용은 동기화 실행에 대한 CloudWatch 로그의 내용을 확인하세요.

지표	설명
	차원: • IndexId • DataSourceId 단위: 수
DocumentsCrawled	검사한 문서 파일 수. 차원: • IndexId • DataSourceId 단위: 수
DocumentsSubmittedForDeletion	검사한 문서 중 데이터 소스에서 삭제되고 삭제를 위해 제출된 문서 수입니다. 차원: • IndexId • DataSourceId 단위: 수
DocumentsSubmittedForDeletionFailed	데이터 소스에서 삭제에 실패한 문서 수. 차원: • IndexId • DataSourceId 단위: 수

지표	설명
DocumentsSubmittedForIndexing	인덱싱을 위해 검사 및 제출된 문서 수. 차원: • IndexId • DataSourceId 단위: 수
DocumentsSubmittedForIndexingFailed	인덱싱을 위해 제출된 문서 중 인덱싱되지 못한 문서 수. 차원: • IndexId • DataSourceId 단위: 수

인덱싱된 문서에 대한 지표

다음 표에서는 인덱싱된 문서에 대한 Amazon Kendra 지표를 설명합니다. BatchPutDocument 작업을 사용하여 인덱싱된 문서의 경우 IndexId 차원만 지원됩니다.

API 또는 CLI를 사용하는 경우 GetMetricStatistics API를 사용할 때 선택한 MetricName 외에도 Namespace를 'AWS/Kendra'로 지정해야 합니다.

지표	설명
DocumentsIndexed	인덱싱된 문서 수. 차원: • IndexId • DataSourceId

지표	설명
	단위: 수
DocumentsFailedToIndex	인덱싱할 수 없는 문서 수. 자세한 내용은 CloudWatch 로그의 내용을 확인하세요. 차원: • IndexId • DataSourceId 단위: 수
IndexQueryCount	분당 인덱스 쿼리 수. 차원: • IndexId 단위: 수

Amazon CloudWatch Logs를 사용한 Amazon Kendra 모니터링

Amazon Kendra는 Amazon CloudWatch Logs를 사용하여 데이터 소스 운영에 대한 통찰을 제공합니다. Amazon Kendra는 문서가 인덱싱되는 동안 프로세스 세부 정보를 기록합니다. 문서가 인덱싱되는 동안 발생하는 데이터 소스의 오류를 기록합니다. Amazon CloudWatch Logs를 사용하여 로그 파일을 모니터링, 저장 및 액세스합니다.

CloudWatch Logs는 로그 그룹의 일부인 로그 스트림에 로그 이벤트를 저장합니다. Amazon Kendra는 다음과 같이 이러한 기능을 사용합니다.

- 로그 그룹 - Amazon Kendra는 모든 로그 스트림을 각 인덱스의 단일 로그 그룹에 저장합니다. Amazon Kendra는 인덱스가 생성될 때 로그 그룹을 생성합니다. 로그 그룹 식별자는 항상 “aws/kendra/”로 시작합니다.
- 로그 스트림 - Amazon Kendra는 실행하는 각 인덱스 동기화 작업에 대해 로그 그룹에 새 데이터 소스 로그 스트림을 생성합니다. 또한 스트림이 약 500개 항목에 도달하면 새 문서 로그 스트림을 생성합니다.

- 로그 항목 - Amazon Kendra는 문서를 인덱싱할 때 로그 스트림에 로그 항목을 생성합니다. 각 항목은 문서 처리 또는 발생한 오류에 대한 정보를 제공합니다.

CloudWatch Logs에 대한 자세한 내용은 Amazon CloudWatch Logs 사용 설명서의 Amazon CloudWatch Logs란 무엇입니까?를 참조하세요.

Amazon Kendra는 두 가지 유형의 로그 스트림을 생성합니다.

- 데이터 소스 로그 스트림
- 문서 로그 스트림

데이터 소스 로그 스트림

데이터 소스 로그 스트림은 인덱스 동기화 작업에 대한 항목을 게시합니다. 각 동기화 작업은 항목을 게시하는 데 사용하는 새 로그 스트림을 만듭니다. 오류 로그 스트림 이름은 다음과 같습니다.

```
data source id/YYYY-MM-DD-HH/data source sync job ID
```

각 동기화 작업 실행에 대해 새 로그 스트림이 생성됩니다.

데이터 소스 로그 스트림에 게시되는 로그 메시지에는 세 가지 유형이 있습니다.

- 인덱싱을 위해 전송하지 못한 문서에 대한 로그 메시지. 다음은 S3 데이터 소스의 문서에 대한 이 메시지의 예입니다.

```
{
    "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."
}
```

- 삭제를 위해 전송하지 못한 문서에 대한 로그 메시지. 다음은 이 메시지의 예입니다.

```
{
    "DocumentId": "document ID",
    "Message": "Failed to delete document via BatchDeleteDocument.",
    "ErrorCode": "InvalidRequest",
```

```
    "ErrorMessage": "Document can't be deleted because it doesn't exist."
}
```

- Amazon S3 버킷에서 문서에 대한 잘못된 메타데이터 파일이 발견된 경우의 로그 메시지. 다음은 이 메시지의 예입니다.

```
{
    "Message": "Found invalid metadata
 file bucket/prefix/filename.extension.metadata.json."
}
```

- SharePoint 및 데이터베이스 커넥터의 경우 Amazon Kendra는 문서를 인덱싱할 수 없는 경우에만 로그 스트림에 메시지를 기록합니다. 다음은 Amazon Kendra가 기록하는 오류 메시지의 예입니다.

```
{
    "DocumentID": "document ID",
    "IndexID": "index ID",
    "SourceURI": "",
    "CrawlStatus": "FAILED",
    "ErrorCode": "403",
    "ErrorMessage": "Access Denied",
    "DataSourceErrorCode": "403"
}
```

문서 로그 스트림

Amazon Kendra는 문서가 인덱싱되는 동안 문서 처리에 대한 정보를 기록합니다. Amazon S3 데이터 소스에 저장된 문서에 대한 메시지 세트를 기록합니다. Microsoft SharePoint 또는 데이터베이스 데이터 소스에 저장된 문서에 대해서만 오류를 기록합니다.

BatchPutDocument 작업을 사용하여 문서를 인덱스에 추가한 경우 로그 스트림의 이름은 다음과 같습니다.

```
YYYY-MM-DD-HH/UUID
```

데이터 소스를 사용하여 문서를 인덱스에 추가한 경우 로그 스트림의 이름은 다음과 같습니다.

```
dataSourceId/YYYY-MM-DD-HH/UUID
```

각 로그 스트림에는 최대 500개의 메시지가 포함됩니다.

문서 인덱싱에 실패하면 다음 메시지가 로그 스트림에 출력됩니다.

```
{
    "DocumentId": "document ID",
    "IndexName": "index name",
    "IndexId": "index ID"
    "SourceURI": "source URI"
    "IndexingStatus": "DocumentFailedToIndex",
    "ErrorCode": "400 | 500",
    "ErrorMessage": "message"
}
```

Amazon Kendra의 보안

클라우드 AWS 보안이 최우선 과제입니다. AWS 고객은 가장 보안에 민감한 조직의 요구 사항을 충족하도록 구축된 데이터 센터 및 네트워크 아키텍처의 혜택을 누릴 수 있습니다.

보안은 기업과 기업 간의 AWS 공동 책임입니다. 공동 책임 모델은 이 사항을 클라우드의 보안 및 클라우드 내 보안으로 설명합니다.

- 클라우드 보안 -AWS 클라우드에서 AWS 서비스를 실행하는 인프라를 보호하는 역할을 합니다. AWS AWS 또한 안전하게 사용할 수 있는 서비스를 제공합니다. Amazon Kendra에 적용되는 규정 준수 프로그램에 대해 자세히 알아보려면 규정 준수 프로그램별 범위 내 서비스 규정 준수 참조하십시오.
- 클라우드에서의 보안 — 사용하는 AWS 서비스에 따라 책임이 결정됩니다. 또한 귀하는 귀사의 데이터의 민감도, 귀사의 요구 사항, 관련 법률 및 규정을 비롯한 기타 요소에 대해서도 책임이 있습니다.

이 설명서는 Amazon Kendra를 사용할 때 공동 책임 모델을 적용하는 방법을 이해하는 데 도움이 됩니다. 다음 주제에서는 보안 및 규정 준수 목적에 맞게 Amazon Kendra를 구성하는 방법을 보여줍니다. 또한 Amazon Kendra 리소스를 모니터링하고 보호하는 데 도움이 되는 다른 AWS 서비스를 사용하는 방법도 알아봅니다.

주제

Amazon Kendra에서의 데이터 보호

AWS 공동 책임 모델 Amazon Kendra의 데이터 보호에 적용됩니다. 이 모델에 설명된 대로 은 (AWS 는) 모두를 실행하는 글로벌 인프라를 보호할 책임이 있습니다. AWS 클라우드사용자는 인프라에서 호스팅되는 콘텐츠를 관리해야 합니다. 사용하는 AWS 서비스 의 보안 구성과 관리 작업에 대한 책임도 사용자에게 있습니다. 데이터 프라이버시에 대한 자세한 내용은 데이터 프라이버시 FAQ를 참조하세요. 유럽의 데이터 보호에 대한 자세한 내용은 AWS 보안 블로그의 AWS 공동 책임 모델 및 GDPR 블로그 게시물을 참조하십시오.

데이터 보호를 위해 AWS 계정 자격 증명을 보호하고 AWS IAM Identity Center OR AWS Identity and Access Management (IAM) 을 사용하여 개별 사용자를 설정하는 것이 좋습니다. 이렇게 하면 개별 사용자에게 자신의 직무를 충실히 이행하는 데 필요한 권한만 부여됩니다. 또한 다음과 같은 방법으로 데이터를 보호하는 것이 좋습니다.

- 각 계정에 멀티 팩터 인증 설정(MFA)을 사용하세요.
- SSL/TLS를 사용하여 리소스와 통신할 수 있습니다. AWS TLS 1.2는 필수이며 TLS 1.3를 권장합니다.
- 를 사용하여 API 및 사용자 활동 로깅을 설정합니다. AWS CloudTrail
- 포함된 모든 기본 보안 제어와 함께 AWS 암호화 솔루션을 사용하십시오 AWS 서비스.
- Amazon S3에 저장된 민감한 데이터를 검색하고 보호하는 데 도움이 되는 Amazon Macie와 같은 고급 관리형 보안 서비스를 사용하세요.
- 명령줄 인터페이스 또는 API를 AWS 통해 액세스할 때 FIPS 140-2로 검증된 암호화 모듈이 필요한 경우 FIPS 엔드포인트를 사용하십시오. 사용 가능한 FIPS 엔드포인트에 대한 자세한 내용은 Federal Information Processing Standard(FIPS) 140-2를 참조하십시오.

고객의 이메일 주소와 같은 기밀 정보나 중요한 정보는 태그나 이름 필드와 같은 자유 양식 필드에 입력하지 않는 것이 좋습니다. 여기에는 콘솔 AWS CLI, API 또는 SDK를 사용하여 Amazon Kendra 또는 AWS 서비스 다른 사용자와 작업하는 경우가 포함됩니다. AWS 이름에 사용되는 태그 또는 자유 형식 텍스트 필드에 입력하는 모든 데이터는 청구 또는 진단 로그에 사용될 수 있습니다. 외부 서버에 URL을 제공할 때 해당 서버에 대한 요청을 검증하기 위해 보안 인증 정보를 URL에 포함시켜서는 안 됩니다.

저장된 데이터 암호화

Amazon Kendra는 사용자가 선택한 암호화 키를 사용하여 저장 데이터를 암호화합니다. 다음 중 하나를 선택할 수 있습니다.

- 소유한 KMS AWS키 AWS . 암호화 키를 지정하지 않으면 기본적으로 이 키로 데이터가 암호화됩니다.
- 계정의 AWS관리형 KMS 키. Amazon Kendra는 이 키를 고객 대신 생성, 관리 및 사용합니다. 키 이름은 `aws/kendra`입니다.
- 고객 관리형 키. 계정에서 생성한 암호화 키의 ARN을 제공할 수 있습니다. 고객 관리형 KMS 키를 사용하는 경우 Amazon Kendra가 키를 사용할 수 있도록 허용하는 키 정책을 키에 제공해야 합니다. 대칭 암호화 고객 관리형 KMS 키를 선택하세요. Amazon Kendra는 비대칭 KMS 키를 지원하지 않습니다. 자세한 정보는 키 관리을 참조하세요.

전송 중 데이터 암호화

Amazon Kendra는 HTTPS 프로토콜을 사용하여 클라이언트 애플리케이션과 통신합니다. 애플리케이션을 대신하여 HTTPS와 AWS 서명을 사용하여 다른 서비스와 통신합니다. VPC를 사용하는 경우 VPC와 Amazon Kendra 간에 프라이빗 연결을 설정하는 AWS PrivateLink 데 사용할 수 있습니다.

키 관리

Amazon Kendra는 세 가지 유형의 키 중 하나를 사용하여 인덱스의 콘텐츠를 암호화합니다. 다음 중 하나를 선택할 수 있습니다.

- 소유한 KMS. AWS AWS 이 값이 기본값입니다.
- AWS관리형 KMS 키. 이 키는 고객 계정에 생성되고, 고객 대신 Amazon Kendra이 관리 및 사용합니다.
- 고객 관리형 KMS 키. Amazon Kendra 인덱스 또는 데이터 소스를 생성할 때 키를 생성하거나 AWS KMS 콘솔을 사용하여 키를 생성할 수 있습니다. 대칭 암호화 고객 관리형 KMS 키를 선택합니다. Amazon Kendra는 비대칭 KMS 키를 지원하지 않습니다. 자세한 내용을 알아보려면 AWS Key Management Service 개발자 안내서의 대칭 및 비대칭 키 사용을 참조하세요.

아마존 켄드라 아마존 켄드라 인텔리전트 랭킹 및 인터페이스 VPC 엔드포인트 ()AWS PrivateLink

인터페이스 VPC 엔드포인트를 생성하여 VPC와 Amazon Kendra 간에 프라이빗 연결을 설정할 수 있습니다. 인터페이스 엔드포인트는 인터넷 게이트웨이 AWS PrivateLink, NAT 디바이스, VPN 연결 또는 Direct Connect 연결 없이 Amazon Kendra API에 비공개로 액세스할 수 있는 기술인 에 의해 구동됩니다. AWS VPC의 인스턴스는 Amazon Kendra API와 통신하는 데 퍼블릭 IP 주소를 필요로 하지 않습니다. VPC와 아마존 Kendra 간의 트래픽은 아마존 네트워크를 벗어나지 않습니다.

각 인터페이스 엔드포인트는 서브넷에서 하나 이상의 탄력적 네트워크 인터페이스로 표현됩니다.

아마존 켄드라 및 아마존 켄드라 인텔리전트 랭킹 VPC 엔드포인트에 대한 고려 사항

Amazon Kendra 또는 Amazon Kendra 인텔리전트 랭킹을 위한 인터페이스 VPC 엔드포인트를 설정하기 전에 Amazon VPC 사용 설명서의 사전 요구 사항을 검토하십시오.

Amazon Kendra 및 Amazon Kendra 인텔리전트 랭킹은 VPC에서 모든 API 작업에 대한 호출을 지원합니다.

아마존 켄드라 및 아마존 켄드라 인텔리전스 랭킹을 위한 인터페이스 VPC 엔드포인트 생성

Amazon VPC 콘솔 또는 () 를 사용하여 Amazon Kendra 또는 Amazon Kendra 인텔리전트 랭킹 서비스를 위한 VPC 엔드포인트를 생성할 수 있습니다. AWS Command Line Interface AWS CLI

다음 서비스 이름을 사용하여 Amazon Kendra에 대한 VPC 엔드포인트를 생성합니다.

- com.amazonaws.*region*.kendra

다음 서비스 이름을 사용하여 Amazon Kendra 인텔리전트 랭킹용 VPC 엔드포인트를 생성합니다.

- aws.api.*region*.kendra-ranking

VPC 엔드포인트를 생성한 후 `endpoint-url` 파라미터를 사용하여 Amazon Kendra API에 대한 인터페이스 엔드포인트를 지정하는 다음 예제 AWS CLI 명령을 사용할 수 있습니다.

```
aws kendra list-indices --endpoint-url https://VPC endpoint
```

VPC ###### 인터페이스 엔드포인트가 생성될 때 생성되는 DNS 이름입니다. 이 이름에는 VPC 엔드포인트 ID와 지역을 포함하는 Amazon Kendra 서비스 이름이 포함됩니다. 예를 들어 `vpce-1234-abcdef.kendra.us-west-2.vpce.amazonaws.com`입니다.

엔드포인트에 대한 프라이빗 DNS를 활성화하면 해당 지역의 기본 DNS 이름을 사용하여 Amazon Kendra에 API 요청을 할 수 있습니다. 예를 들어 `kendra.us-east-1.amazonaws.com`입니다.

자세한 내용은 Amazon VPC 사용 설명서의 인터페이스 엔드포인트 생성을 참조하세요.

아마존 켄드라 및 아마존 켄드라 인텔리전스 랭킹에 대한 VPC 엔드포인트 정책 생성

Amazon Kendra 또는 Amazon Kendra 인텔리전트 랭킹에 대한 액세스를 제어하는 엔드포인트 정책을 VPC 엔드포인트에 연결할 수 있습니다.

아마존 켄드라 또는 아마존 켄드라 인텔리전트 랭킹 정책에는 다음과 같은 정보가 명시되어 있습니다.

- 작업을 수행할 수 있는 주체/권한 있는 사용자.
- 수행할 수 있는 작업.
- 작업을 수행할 수 있는 리소스.

예제: Amazon Kendra 작업에 대한 VPC 엔드포인트 정책

다음은 Amazon Kendra에 대한 엔드포인트 정책의 예입니다. 엔드포인트에 연결되면 이 정책은 모든 리소스의 모든 주체/권한 있는 사용자에게 사용 가능한 모든 Amazon Kendra 작업에 대한 액세스 권한을 부여합니다.

```
{
   "Statement":[
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "kendra:*"
         ],
         "Resource":"*"
      }
   ]
}
```

예: Amazon Kendra 인텔리전트 랭킹 작업에 대한 VPC 엔드포인트 정책

다음은 Amazon Kendra 인텔리전트 랭킹에 대한 엔드포인트 정책의 예입니다. 엔드포인트에 연결되면 이 정책은 모든 리소스의 모든 주체/권한 있는 사용자에게 사용 가능한 모든 Amazon Kendra Intelligent Ranking 작업에 대한 액세스 권한을 부여합니다.

```
{
   "Statement":[
```

```
      {
         "Principal":"*",
         "Effect":"Allow",
         "Action":[
            "kendra-ranking:*"
         ],
         "Resource":"*"
      }
   ]
}
```

자세한 내용은 Amazon VPC 사용 설명서의 엔드포인트 정책을 사용한 VPC 엔드포인트 액세스 제어를 참조하십시오.

Amazon Kendra용 Identity and Access Management

AWS Identity and Access Management (IAM) 은 관리자가 리소스에 대한 액세스를 안전하게 제어하는 데 도움이 AWS 서비스 되는 도구입니다. AWS IAM 관리자는 Amazon Kendra 리소스를 사용할 수 있도록 인증(로그인)하고 권한을 부여(권한 보유)할 수 있는 사용자를 제어합니다. IAM은 추가 AWS 서비스 비용 없이 사용할 수 있습니다.

주제

- 고객
- ID를 통한 인증
- 정책을 사용한 액세스 관리
- Amazon Kendra에서 IAM을 사용하는 방법
- Amazon Kendra 자격 증명 기반 정책 예제
- AWS 아마존 켄드라의 관리형 정책
- Amazon Kendra 자격 증명 및 액세스 문제 해결

고객

사용 방법 AWS Identity and Access Management (IAM) 은 Amazon Kendra에서 수행하는 작업에 따라 다릅니다.

서비스 사용자 - Amazon Kendra 서비스를 사용하여 작업을 수행하는 경우 필요한 보안 인증 정보와 권한을 관리자가 제공합니다. 더 많은 Amazon Kendra 기능을 사용하여 작업을 수행한다면 추가 권한

이 필요할 수 있습니다. 액세스 권한 관리 방식을 이해하면 적절한 권한을 관리자에게 요청할 수 있습니다. Amazon Kendra의 기능에 액세스할 수 없다면 Amazon Kendra 자격 증명 및 액세스 문제 해결 섹션을 참조하세요.

서비스 관리자 – 회사에서 Amazon Kendra 리소스를 책임지고 있다면 Amazon Kendra에 대한 모든 액세스 권한이 있을 것입니다. 서비스 관리자는 서비스 사용자가 액세스해야 하는 Amazon Kendra 기능과 리소스를 결정합니다. 그런 다음, IAM 관리자에게 요청을 제출하여 서비스 사용자의 권한을 변경해야 합니다. 이 페이지의 정보를 검토하여 IAM의 기본 개념을 이해하십시오. 귀사에서 Amazon Kendra와 함께 IAM을 사용하는 방법에 대해 자세히 알아보려면 Amazon Kendra에서 IAM을 사용하는 방법 부분을 참조하세요.

IAM 관리자 - IAM 관리자라면 Amazon Kendra에 대한 액세스 권한 관리 정책 작성 방법을 자세히 알고 싶을 것입니다. IAM에서 사용할 수 있는 Amazon Kendra 자격 증명 기반 정책(identity-based policies)의 예제를 확인하시려면 자격 증명 기반 정책 예제 Amazon Kendra 자격 증명 기반 정책 예제 섹션을 참조하세요.

ID를 통한 인증

인증은 자격 증명을 AWS 사용하여 로그인하는 방법입니다. IAM 사용자로 인증 (로그인 AWS) 하거나 IAM 역할을 맡아 인증 (로그인) 해야 합니다. AWS 계정 루트 사용자

ID 소스를 통해 제공된 자격 증명을 사용하여 페더레이션 ID로 로그인할 수 있습니다. AWS AWS IAM Identity Center (IAM ID 센터) 사용자, 회사의 싱글 사인온 인증, Google 또는 Facebook 자격 증명이 페더레이션 ID의 예입니다. 페더레이션 ID로 로그인할 때 관리자가 이전에 IAM 역할을 사용하여 ID 페더레이션을 설정했습니다. 페더레이션을 사용하여 액세스하는 경우 AWS 간접적으로 역할을 맡게 됩니다.

사용자 유형에 따라 AWS Management Console 또는 AWS 액세스 포털에 로그인할 수 있습니다. 로그인에 대한 자세한 내용은 AWS 로그인 사용 설명서의 내 로그인 방법을 참조하십시오. AWS AWS 계정

AWS 프로그래밍 방식으로 액세스하는 경우 자격 증명을 사용하여 요청에 암호화 방식으로 서명할 수 있는 소프트웨어 개발 키트 (SDK) 와 명령줄 인터페이스 (CLI) 를 AWS 제공합니다. AWS 도구를 사용하지 않는 경우 요청에 직접 서명해야 합니다. 권장 방법을 사용하여 직접 요청에 서명하는 방법에 대한 자세한 내용은 IAM 사용 설명서의 AWS API 요청 서명을 참조하십시오.

사용하는 인증 방법에 상관없이 추가 보안 정보를 제공해야 할 수도 있습니다. 예를 들어, AWS 계정의 보안을 강화하기 위해 다단계 인증 (MFA) 을 사용할 것을 권장합니다. 자세한 내용은 AWS IAM Identity Center 사용 설명서의 다중 인증 및 IAM 사용 설명서의 AWS에서 다중 인증(MFA) 사용을 참조하십시오.

AWS 계정 루트 사용자

계정을 AWS 계정만들 때는 먼저 계정의 모든 AWS 서비스 리소스에 대한 완전한 액세스 권한을 가진 하나의 로그인 ID로 시작합니다. 이 ID를 AWS 계정 루트 사용자라고 하며, 계정을 만들 때 사용한 이메일 주소와 비밀번호로 로그인하여 액세스할 수 있습니다. 일상적인 태스크에 루트 사용자를 사용하지 않을 것을 강력히 권장합니다. 루트 사용자 보안 인증 정보를 보호하고 루트 사용자만 수행할 수 있는 태스크를 수행하는 데 사용하세요. 루트 사용자로 로그인해야 하는 전체 작업 목록은 IAM 사용 설명서의 Tasks that require root user credentials를 참조하십시오.

IAM 사용자 및 그룹

IAM 사용자는 단일 사용자 또는 애플리케이션에 대한 특정 권한을 가진 사용자 내의 자격 증명입니다. AWS 계정 가능하면 암호 및 액세스 키와 같은 장기 보안 인증이 있는 IAM 사용자를 생성하는 대신 임시 보안 인증을 사용하는 것이 좋습니다. 하지만 IAM 사용자의 장기 보안 인증이 필요한 특정 사용 사례가 있는 경우, 액세스 키를 교체하는 것이 좋습니다. 자세한 내용은 IAM 사용 설명서의 장기 보안 인증이 필요한 사용 사례의 경우 정기적으로 액세스 키 교체를 참조하십시오.

IAM 그룹은 IAM 사용자 컬렉션을 지정하는 자격 증명입니다. 사용자는 그룹으로 로그인할 수 없습니다. 그룹을 사용하여 여러 사용자의 권한을 한 번에 지정할 수 있습니다. 그룹을 사용하면 대규모 사용자 집합의 권한을 더 쉽게 관리할 수 있습니다. 예를 들어, IAMAdmins라는 그룹이 있고 이 그룹에 IAM 리소스를 관리할 권한을 부여할 수 있습니다.

사용자는 역할과 다릅니다. 사용자는 한 사람 또는 애플리케이션과 고유하게 연결되지만, 역할은 해당 역할이 필요한 사람이라면 누구나 수임할 수 있습니다. 사용자는 영구적인 장기 보안 인증 정보를 가지고 있지만, 역할은 임시 보안 인증만 제공합니다. 자세한 내용은 IAM 사용 설명서의 IAM 사용자를 만들어야 하는 경우(역할이 아님)를 참조하십시오.

IAM 역할

IAM 역할은 특정 권한을 가진 사용자 AWS 계정 내의 자격 증명입니다. IAM 사용자와 유사하지만, 특정 개인과 연결되지 않습니다. 역할을 AWS Management Console 전환하여 에서 일시적으로 IAM 역할을 맡을 수 있습니다. AWS CLI 또는 AWS API 작업을 호출하거나 사용자 지정 URL을 사용하여 역할을 수임할 수 있습니다. 역할 사용 방법에 대한 자세한 내용은 IAM 사용 설명서의 IAM 역할 사용을 참조하십시오.

임시 보안 인증이 있는 IAM 역할은 다음과 같은 상황에서 유용합니다.

- 페더레이션 사용자 액세스 - 페더레이션 ID에 권한을 부여하려면 역할을 생성하고 해당 역할의 권한을 정의합니다. 페더레이션 ID가 인증되면 역할이 연결되고 역할에 정의된 권한이 부여됩니다. 페더레이션 역할에 대한 자세한 내용은 IAM 사용 설명서의 서드 파티 ID 공급자의 역할 생성 단원을 참조

하십시오. IAM Identity Center를 사용하는 경우, 권한 집합을 구성합니다. 인증 후 ID가 액세스할 수 있는 항목을 제어하기 위해 IAM Identity Center는 권한 세트를 IAM의 역할과 연관짓습니다. 권한 세트에 대한 자세한 내용은 AWS IAM Identity Center 사용 설명서의 권한 세트를 참조하십시오.

- 임시 IAM 사용자 권한 - IAM 사용자 또는 역할은 IAM 역할을 수임하여 특정 태스크에 대한 다양한 권한을 임시로 받을 수 있습니다.
- 크로스 계정 액세스 - IAM 역할을 사용하여 다른 계정의 사용자(신뢰할 수 있는 보안 주체)가 내 계정의 리소스에 액세스하도록 허용할 수 있습니다. 역할은 계정 간 액세스를 부여하는 기본적인 방법입니다. 그러나 일부 AWS 서비스경우에는 역할을 프록시로 사용하는 대신 정책을 리소스에 직접 연결할 수 있습니다. 계정 간 액세스에 대한 역할과 리소스 기반 정책의 차이점을 알아보려면 IAM 사용 설명서의 IAM의 교차 계정 리소스 액세스를 참조하십시오.
- 서비스 간 액세스 — 일부는 다른 기능을 사용합니다. AWS 서비스 AWS 서비스예를 들어 서비스에서 직접적 호출을 수행하면 일반적으로 해당 서비스는 Amazon EC2에서 애플리케이션을 실행하거나 Amazon S3에 객체를 저장합니다. 서비스는 직접적으로 호출하는 보안 주체의 권한을 사용하거나, 서비스 역할을 사용하거나, 또는 서비스 연결 역할을 사용하여 이 태스크를 수행할 수 있습니다.
  - 순방향 액세스 세션 (FAS) — IAM 사용자 또는 역할을 사용하여 작업을 수행하는 경우 보안 AWS 주체로 간주됩니다. 일부 서비스를 사용하는 경우 다른 서비스에서 다른 작업을 시작하는 작업을 수행할 수 있습니다. FAS는 전화를 거는 주체의 권한을 다운스트림 AWS 서비스서비스에 AWS 서비스 요청하기 위한 요청과 결합하여 사용합니다. FAS 요청은 다른 서비스 AWS 서비스 또는 리소스와의 상호 작용이 필요한 요청을 서비스가 수신한 경우에만 이루어집니다. 이 경우 두 작업을 모두 수행할 수 있는 권한이 있어야 합니다. FAS 요청 시 정책 세부 정보는 전달 액세스 세션을 참조하세요.
  - 서비스 역할 - 서비스 역할은 서비스가 사용자를 대신하여 태스크를 수행하기 위해 맡는 IAM 역할입니다. IAM 관리자는 IAM 내에서 서비스 역할을 생성, 수정 및 삭제할 수 있습니다. 자세한 내용은 IAM 사용 설명서의 AWS 서비스에 대한 권한을 위임할 역할 생성을 참조하십시오.
  - 서비스 연결 역할 — 서비스 연결 역할은 에 연결된 서비스 역할의 한 유형입니다. AWS 서비스서비스는 사용자를 대신하여 작업을 수행하기 위해 역할을 수임할 수 있습니다. 서비스 연결 역할은 사용자에게 AWS 계정 표시되며 해당 서비스가 소유합니다. IAM 관리자는 서비스 링크 역할의 권한을 볼 수 있지만 편집은 할 수 없습니다.
- Amazon EC2에서 실행되는 애플리케이션 — IAM 역할을 사용하여 EC2 인스턴스에서 실행되고 API 요청을 AWS CLI 하는 애플리케이션의 임시 자격 증명을 관리할 수 있습니다. AWS 이는 EC2 인스턴스 내에 액세스 키를 저장할 때 권장되는 방법입니다. EC2 인스턴스에 AWS 역할을 할당하고 모든 애플리케이션에서 사용할 수 있게 하려면 인스턴스에 연결된 인스턴스 프로필을 생성합니다. 인스턴스 프로파일에는 역할이 포함되어 있으며 EC2 인스턴스에서 실행되는 프로그램이 임시 보안 인증을 얻을 수 있습니다. 자세한 내용은 IAM 사용 설명서의 IAM 역할을 사용하여 Amazon EC2 인스턴스에서 실행되는 애플리케이션에 권한 부여를 참조하십시오.

IAM 역할을 사용할지 또는 IAM 사용자를 사용할지를 알아보려면 IAM 사용 설명서의 IAM 역할(사용자 대신)을 생성하는 경우를 참조하십시오.

정책을 사용한 액세스 관리

정책을 생성하고 이를 AWS ID 또는 리소스에 AWS 연결하여 액세스를 제어할 수 있습니다. 정책은 ID 또는 리소스와 연결될 때 AWS 해당 권한을 정의하는 객체입니다. AWS 주도자 (사용자, 루트 사용자 또는 역할 세션) 가 요청할 때 이러한 정책을 평가합니다. 정책에서 권한은 요청이 허용되거나 거부되는 지를 결정합니다. 대부분의 정책은 JSON 문서로 AWS 저장됩니다. JSON 정책 문서의 구조와 콘텐츠에 대한 자세한 내용은 IAM 사용 설명서의 JSON 정책 개요를 참조하십시오.

관리자는 AWS JSON 정책을 사용하여 누가 무엇에 액세스할 수 있는지 지정할 수 있습니다. 즉, 어떤 보안 주체가 어떤 리소스와 어떤 조건에서 작업을 수행할 수 있는지를 지정할 수 있습니다.

기본적으로, 사용자와 역할에는 어떠한 권한도 없습니다. 사용자에게 사용자가 필요한 리소스에서 작업을 수행할 권한을 부여하려면 IAM 관리자가 IAM 정책을 생성하면 됩니다. 그런 다음 관리자가 IAM 정책을 역할에 추가하고, 사용자가 역할을 수임할 수 있습니다.

IAM 정책은 작업을 수행하기 위해 사용하는 방법과 상관없이 작업에 대한 권한을 정의합니다. 예를 들어, `iam:GetRole` 작업을 허용하는 정책이 있다고 가정합니다. 해당 정책을 사용하는 사용자는 AWS Management Console, AWS CLI, 또는 AWS API에서 역할 정보를 가져올 수 있습니다.

보안 인증 기반 정책

ID 기반 정책은 IAM 사용자, 사용자 그룹 또는 역할과 같은 ID에 연결할 수 있는 JSON 권한 정책 문서입니다. 이러한 정책은 사용자와 역할이 어떤 리소스와 어떤 조건에서 어떤 태스크를 수행할 수 있는지를 제어합니다. ID 기반 정책을 생성하는 방법을 알아보려면 IAM 사용 설명서의 IAM 정책 생성을 참조하십시오.

보안 인증 기반 정책은 인라인 정책 또는 관리형 정책으로 한층 더 분류할 수 있습니다. 인라인 정책은 단일 사용자, 그룹 또는 역할에 직접 포함됩니다. 관리형 정책은 내 여러 사용자, 그룹 및 역할에 연결할 수 있는 독립형 정책입니다. AWS 계정관리형 정책에는 AWS 관리형 정책과 고객 관리형 정책이 포함됩니다. 관리형 정책 또는 인라인 정책을 선택하는 방법을 알아보려면 IAM 사용 설명서의 관리형 정책과 인라인 정책의 선택을 참조하십시오.

리소스 기반 정책

리소스 기반 정책은 리소스에 연결하는 JSON 정책 설명서입니다. 리소스 기반 정책의 예는 IAM 역할 신뢰 정책과 Amazon S3 버킷 정책입니다. 리소스 기반 정책을 지원하는 서비스에서 서비스 관리자는

이러한 정책을 사용하여 특정 리소스에 대한 액세스를 통제할 수 있습니다. 정책이 연결된 리소스의 경우, 정책은 지정된 보안 주체가 해당 리소스와 어떤 조건에서 어떤 태스크를 수행할 수 있는지를 정의합니다. 리소스 기반 정책에서 보안 주체를 지정해야 합니다. 보안 주체에는 계정, 사용자, 역할, 연동 사용자 등이 포함될 수 있습니다. AWS 서비스

리소스 기반 정책은 해당 서비스에 있는 인라인 정책입니다. IAM의 AWS 관리형 정책은 리소스 기반 정책에 사용할 수 없습니다.

액세스 제어 목록(ACL)

액세스 제어 목록(ACL)은 어떤 보안 주체(계정 멤버, 사용자 또는 역할)가 리소스에 액세스할 수 있는 권한을 가지고 있는지를 제어합니다. ACLs는 JSON 정책 문서 형식을 사용하지 않지만 리소스 기반 정책과 유사합니다.

ACL을 지원하는 서비스의 예로는 아마존 S3와 아마존 VPC가 있습니다. AWS WAF ACL에 대해 자세히 알아보려면 Amazon Simple Storage Service 개발자 가이드의 ACL(액세스 제어 목록) 개요를 참조하십시오.

기타 정책 타입

AWS 일반적이지 않은 추가 정책 유형을 지원합니다. 이러한 정책 타입은 더 일반적인 정책 타입에 따라 사용자에게 부여되는 최대 권한을 설정할 수 있습니다.

- 권한 경계 – 권한 경계는 자격 증명 기반 정책에 따라 IAM 엔터티(IAM 사용자 또는 역할)에 부여할 수 있는 최대 권한을 설정하는 고급 기능입니다. 개체에 대한 권한 경계를 설정할 수 있습니다. 그 결과로 얻는 권한은 개체의 보안 인증 기반 정책과 그 권한 경계의 교집합입니다. `Principal` 필드에서 사용자나 역할을 지정하는 리소스 기반 정책은 권한 경계를 통해 제한되지 않습니다. 이러한 정책 중 하나에 포함된 명시적 거부는 허용을 재정의합니다. 권한 경계에 대한 자세한 내용은 IAM 사용 설명서의 IAM 엔터티에 대한 권한 경계를 참조하십시오.
- 서비스 제어 정책 (SCP) - SCP는 조직 또는 조직 단위 (OU) 에 대한 최대 권한을 지정하는 JSON 정책입니다. AWS Organizations AWS Organizations 사업체가 소유한 여러 AWS 계정 개를 그룹화하고 중앙에서 관리하는 서비스입니다. 조직에서 모든 기능을 활성화할 경우, 서비스 제어 정책(SCP)을 임의의 또는 모든 계정에 적용할 수 있습니다. SCP는 구성원 계정의 엔티티 (각 엔티티 포함) 에 대한 권한을 제한합니다. AWS 계정 루트 사용자조직 및 SCP에 대한 자세한 내용은 AWS Organizations 사용 설명서의 SCP 작동 방식을 참조하십시오.
- 세션 정책 - 세션 정책은 역할 또는 페더레이션 사용자에 대해 임시 세션을 프로그래밍 방식으로 생성할 때 파라미터로 전달하는 고급 정책입니다. 결과적으로 얻는 세션의 권한은 사용자 또는 역할의 보안 인증 기반 정책의 교차와 세션 정책입니다. 또한 권한을 리소스 기반 정책에서 가져올 수도 있

습니다. 이러한 정책 중 하나에 포함된 명시적 거부는 허용을 재정의합니다. 자세한 내용은 IAM 사용 설명서의 세션 정책을 참조하십시오.

여러 정책 타입

여러 정책 유형이 요청에 적용되는 경우, 결과 권한은 이해하기가 더 복잡합니다. 여러 정책 유형이 관련된 경우 요청을 허용할지 여부를 AWS 결정하는 방법을 알아보려면 IAM 사용 설명서의 정책 평가 로직을 참조하십시오.

Amazon Kendra에서 IAM을 사용하는 방법

IAM을 사용하여 Amazon Kendra에 대한 액세스를 관리하기 전에 Amazon Kendra에서 사용할 수 있는 IAM 기능을 이해해야 합니다. Amazon Kendra AWS 및 기타 서비스가 IAM과 함께 작동하는 방식을 개괄적으로 살펴보려면 IAM AWS 사용 설명서의 IAM과 함께 작동하는 서비스를 참조하십시오.

주제

- Amazon Kendra 자격 증명 기반 정책
- Amazon Kendra 리소스 기반 정책
- 액세스 제어 목록(ACL)
- Amazon Kendra 태그 기반 권한 부여
- Amazon Kendra IAM 역할

Amazon Kendra 자격 증명 기반 정책

IAM ID 기반 정책을 사용하면 허용되거나 거부되는 작업과 리소스뿐 아니라 작업이 허용되거나 거부되는 조건을 지정할 수 있습니다. Amazon Kendra는 특정 작업, 리소스, 조건 키를 지원합니다. JSON 정책에서 사용하는 모든 요소에 대해 알고 싶다면 IAM 사용 설명서의 IAM JSON 정책 요소 참조를 참조하세요.

작업

관리자는 AWS JSON 정책을 사용하여 누가 무엇에 액세스할 수 있는지 지정할 수 있습니다. 즉, 어떤 보안 주체가 어떤 리소스와 어떤 조건에서 작업을 수행할 수 있는지를 지정할 수 있습니다.

JSON 정책의 `Action`요소는 정책에서 액세스를 허용하거나 거부하는 데 사용할 수 있는 태스크를 설명합니다. 정책 작업은 일반적으로 관련 AWS API 작업과 이름이 같습니다. 일치하는 API 작업이 없는 권한 전용 작업 같은 몇 가지 예외도 있습니다. 정책에서 여러 작업이 필요한 몇 가지 작업도 있습니다. 이러한 추가 작업을 일컬어 종속 작업이라고 합니다.

연결된 작업을 수행할 수 있는 권한을 부여하기 위한 정책에 작업을 포함하십시오.

Amazon Kendra의 정책 작업은 작업 앞에 `kendra:` 접두사를 사용합니다. 예를 들어, API 작업을 통해 Amazon Kendra 인덱스를 나열할 수 있는 권한을 누군가에게 부여하려면 ListIndices해당 사용자의 정책에 작업을 포함해야 `kendra:ListIndices` 합니다. 정책 문에는 `Action` 또는 `NotAction` 요소가 포함되어야 합니다. Amazon Kendra는 이 서비스로 수행할 수 있는 태스크를 설명하는 고유한 작업 세트를 정의합니다.

명령문 하나에 여러 태스크를 지정하려면 다음과 같이 쉼표로 구분합니다.

```
"Action": [
      "kendra:action1",
      "kendra:action2"
```

와일드카드(*)를 사용하여 여러 작업을 지정할 수 있습니다. 예를 들어, `Describe`라는 단어로 시작하는 모든 태스크를 지정하려면 다음 태스크를 포함합니다.

```
"Action": "kendra:Describe*"
```

Amazon Kendra 작업의 목록을 보려면 IAM 사용 설명서의 Amazon Kendra에서 정의한 작업을 참조하세요.

리소스

관리자는 AWS JSON 정책을 사용하여 누가 무엇에 액세스할 수 있는지 지정할 수 있습니다. 즉, 어떤 보안 주체가 어떤 리소스와 어떤 조건에서 작업을 수행할 수 있는지를 지정할 수 있습니다.

`Resource` JSON 정책 요소는 작업이 적용되는 하나 이상의 개체를 지정합니다. 문장에는 `Resource`또는 `NotResource`요소가 반드시 추가되어야 합니다. 모범 사례에 따라 Amazon 리소스 이름(ARN)을 사용하여 리소스를 지정합니다. 리소스 수준 권한이라고 하는 특정 리소스 유형을 지원하는 작업에 대해 이 태스크를 수행할 수 있습니다.

작업 나열과 같이 리소스 수준 권한을 지원하지 않는 작업의 경우, 와일드카드(*)를 사용하여 해당 문이 모든 리소스에 적용됨을 나타냅니다.

```
"Resource": "*"
```

Amazon Kendra 인덱스 리소스의 ARN은 다음과 같습니다.

```
arn:${Partition}:kendra:${Region}:${Account}:index/${IndexId}
```

ARN 형식에 대한 자세한 내용은 Amazon 리소스 이름 (ARN) 및 AWS 서비스 네임스페이스를 참조하십시오.

예를 들어 문에 인덱스를 지정하려면 다음 ARN에서 인덱스의 GUID를 사용합니다.

```
"Resource": "arn:aws:kendra:${Region}:${Account}:index/${GUID}"
```

특정 계정에 속하는 모든 인덱스를 지정하려면 와일드카드(*)를 사용합니다.

```
"Resource": "arn:aws:${Region}:${Account}:index/*"
```

리소스 생성 작업과 같은 일부 Amazon Kendra 작업은 특정 리소스에서 수행할 수 없습니다. 이러한 경우, 와일드카드(*)를 사용해야 합니다.

```
"Resource": "*"
```

Amazon Kendra 리소스 유형 및 해당 ARN의 목록을 보려면 IAM 사용 설명서의 Amazon Kendra에서 정의한 리소스를 참조하세요. 각 리소스의 ARN을 지정할 수 있는 작업을 알아보려면 Amazon Kendra에서 정의한 작업을 참조하세요.

조건 키

관리자는 AWS JSON 정책을 사용하여 누가 무엇에 액세스할 수 있는지 지정할 수 있습니다. 즉, 어떤 보안 주체가 어떤 리소스와 어떤 조건에서 작업을 수행할 수 있는지를 지정할 수 있습니다.

Condition 요소(또는 Condition 블록)를 사용하면 정책이 발효되는 조건을 지정할 수 있습니다. Condition 요소는 옵션입니다. 같거나 작음과 같은 조건 연산자를 사용하여 정책의 조건을 요청의 값과 일치시키는 조건식을 생성할 수 있습니다.

한 문에서 여러 Condition 요소를 지정하거나 단일 Condition 요소에서 여러 키를 지정하는 경우, AWS 는 논리적 AND 태스크를 사용하여 평가합니다. 단일 조건 키에 여러 값을 지정하는 경우 는 논리적 OR 연산을 사용하여 조건을 AWS 평가합니다. 명문의 권한을 부여하기 전에 모든 조건을 충족해야 합니다.

조건을 지정할 때 자리 표시자 변수를 사용할 수도 있습니다. 예컨대, IAM 사용자에게 IAM 사용자 이름으로 태그가 지정된 경우에만 리소스에 액세스할 수 있는 권한을 부여할 수 있습니다. 자세한 내용은 IAM 사용 설명서의 IAM 정책 요소: 변수 및 태그를 참조하십시오.

AWS 글로벌 조건 키 및 서비스별 조건 키를 지원합니다. 모든 AWS 글로벌 조건 키를 보려면 IAM 사용 AWS 설명서의 글로벌 조건 컨텍스트 키를 참조하십시오.

Amazon Kendra는 서비스별 조건 키를 제공하지 않지만, 일부 전역 조건 키 사용을 지원합니다. 모든 AWS 글로벌 조건 키를 보려면 IAM 사용 설명서의AWS 글로벌 조건 컨텍스트 키를 참조하십시오.

예

Amazon Kendra 자격 증명 기반 정책 예제를 보려면 Amazon Kendra 자격 증명 기반 정책 예제 단원을 참조하세요.

Amazon Kendra 리소스 기반 정책

Amazon Kendra는 리소스 기반 정책을 지원하지 않습니다.

액세스 제어 목록(ACL)

Amazon Kendra는 AWS 서비스 및 리소스에 대한 액세스를 위한 액세스 제어 목록(ACL)을 지원하지 않습니다.

Amazon Kendra 태그 기반 권한 부여

태그를 특정 유형의 Amazon Kendra 리소스와 연결하여 해당 리소스에 대한 액세스를 승인할 수 있습니다. 태그를 기반으로 액세스를 제어하려면 `aws:RequestTag/key-name` 또는 `aws:TagKeys` 조건 키를 사용하여 정책의 조건 요소에 태그 정보를 제공합니다.

다음 표에는 태그 기반 액세스 제어에 필요한 작업, 해당 리소스 유형과 조건 키가 나열되어 있습니다. 각 작업은 해당 리소스 유형과 연결된 태그를 기반으로 권한이 부여됩니다.

작업	리소스 유형	조건 키
CreateData소스		`aws:RequestTag` , `aws:TagKeys`
CreateFaq		`aws:RequestTag` , `aws:TagKeys`
CreateIndex		`aws:RequestTag` , `aws:TagKeys`

작업	리소스 유형	조건 키
API_ ListTags ForResource	데이터 소스, FAQ, 인덱스	
TagResource	데이터 소스, FAQ, 인덱스	`aws:RequestTag` , `aws:TagKeys`
UntagResource	데이터 소스, FAQ, 인덱스	`aws:TagKeys`

Amazon Kendra 리소스 태그 지정에 대한 자세한 내용은 Tags 섹션을 참조하세요. 리소스 태그를 기반으로 리소스에 대한 액세스를 제한하는 자격 증명 기반 정책의 예는 태그 기반 정책 예제 단원을 참조하세요. 태그를 사용하여 리소스에 대한 액세스를 제한하는 방법에 대한 자세한 내용은 IAM 사용 설명서의 태그를 사용한 액세스 제어를 참조하세요.

Amazon Kendra IAM 역할

IAM 역할은 AWS 계정 내에서 특정 권한을 가진 엔티티입니다.

Amazon Kendra에서 임시 자격 증명 사용

임시 보안 인증을 사용하여 페더레이션을 통해 로그인하거나, IAM 역할을 맡거나, 교차 계정 역할을 맡을 수 있습니다. AssumeRole또는 GetFederation토큰과 같은 AWS STS API 작업을 호출하여 임시 보안 자격 증명을 얻습니다.

Amazon Kendra는 임시 자격 증명 사용을 지원합니다.

서비스 역할

이 기능을 사용하면 서비스가 사용자를 대신하여 서비스 역할을 수임할 수 있습니다. 이 역할을 사용하면 서비스가 다른 서비스의 리소스에 액세스해 사용자를 대신해 작업을 완료할 수 있습니다. 서비스 역할은 IAM 계정에 나타나고, 해당 계정이 소유합니다. 즉, IAM 관리자가 이 역할에 대한 권한을 변경할 수 있습니다. 그러나 권한을 변경하면 서비스의 기능이 손상될 수 있습니다.

Amazon Kendra는 서비스 역할을 지원합니다.

Amazon Kendra에서 IAM 역할 선택

인덱스를 생성하거나, `BatchPutDocument` 작업을 호출하거나, 데이터 소스나 FAQ를 생성할 때는 Amazon Kendra가 사용자 대신 필요한 리소스에 액세스하는 데 사용하는 Amazon 리소스 이름(ARN) 액세스 역할을 제공해야 합니다. 이전에 역할을 생성한 경우 Amazon Kendra 콘솔이 선택할 수 있는

역할 목록을 제공합니다. 필요한 리소스에 대한 액세스를 허용하는 역할을 선택하는 것이 중요합니다. 자세한 정보는 IAM 액세스 역할 Amazon Kendra을 참조하세요.

Amazon Kendra 자격 증명 기반 정책 예제

기본적으로 사용자 및 역할은 Amazon Kendra 리소스를 생성하거나 수정할 수 있는 권한이 없습니다. 또한 AWS Management Console AWS CLI, 또는 AWS API를 사용하여 작업을 수행할 수 없습니다. IAM 관리자는 지정된 리소스에서 특정 API 작업을 수행할 수 있는 권한을 사용자와 역할에게 부여하는 IAM 정책을 생성해야 합니다. 그런 다음 관리자는 해당 권한이 필요한 사용자 또는 그룹에 이러한 정책을 연결해야 합니다.

이러한 예제 JSON 정책 문서를 사용하여 IAM 자격 증명 기반 정책을 생성하는 방법을 알아보려면 IAM 사용 설명서의 JSON 탭에서 정책 생성을 참조하십시오.

주제

- 정책 모범 사례
- Amazon Kendra용AWS 관리형 (미리 정의된) 정책
- 사용자가 자신의 고유한 권한을 볼 수 있도록 허용
- Amazon Kendra 인덱스에 액세스
- 태그 기반 정책 예제

정책 모범 사례

ID 기반 정책에 따라 계정에서 사용자가 Amazon Kendra 리소스를 생성, 액세스 또는 삭제할 수 있는지 여부가 결정됩니다. 이 작업으로 인해 AWS 계정에 비용이 발생할 수 있습니다. ID 기반 정책을 생성하거나 편집할 때는 다음 지침과 권장 사항을 따릅니다.

- AWS 관리형 정책으로 시작하여 최소 권한 권한으로 이동 — 사용자와 워크로드에 권한을 부여하려면 여러 일반 사용 사례에 권한을 부여하는 AWS 관리형 정책을 사용하세요. 에서 사용할 수 있습니다. AWS 계정사용 사례에 맞는 AWS 고객 관리형 정책을 정의하여 권한을 더 줄이는 것이 좋습니다. 자세한 내용은 IAM 사용 설명서의 AWS 관리형 정책 또는 직무에 대한AWS 관리형 정책을 참조하십시오.
- 최소 권한 적용 – IAM 정책을 사용하여 권한을 설정하는 경우, 태스크를 수행하는 데 필요한 권한만 부여합니다. 이렇게 하려면 최소 권한으로 알려진 특정 조건에서 특정 리소스에 대해 수행할 수 있는 작업을 정의합니다. IAM을 사용하여 권한을 적용하는 방법에 대한 자세한 정보는 IAM 사용 설명서의 IAM의 정책 및 권한을 참조하십시오.

- IAM 정책의 조건을 사용하여 액세스 추가 제한 – 정책에 조건을 추가하여 작업 및 리소스에 대한 액세스를 제한할 수 있습니다. 예를 들어 SSL을 사용하여 모든 요청을 전송해야 한다고 지정하는 정책 조건을 작성할 수 있습니다. 예를 AWS 서비스들어 특정 작업을 통해 서비스 작업을 사용하는 경우 조건을 사용하여 서비스 작업에 대한 액세스 권한을 부여할 수도 AWS CloudFormation있습니다. 자세한 내용은 IAM 사용 설명서의 IAM JSON 정책 요소: 조건을 참조하십시오.
- IAM Access Analyzer를 통해 IAM 정책을 확인하여 안전하고 기능적인 권한 보장 - IAM Access Analyzer에서는 IAM 정책 언어(JSON)와 모범 사례가 정책에서 준수되도록 신규 및 기존 정책을 확인합니다. IAM Access Analyzer는 100개 이상의 정책 확인 항목과 실행 가능한 추천을 제공하여 안전하고 기능적인 정책을 작성하도록 돕습니다. 자세한 내용은 IAM 사용 설명서의 IAM Access Analyzer 정책 검증을 참조하십시오.
- 멀티 팩터 인증 (MFA) 필요 - IAM 사용자 또는 루트 사용자가 필요한 시나리오가 있는 경우 추가 보안을 위해 AWS 계정 MFA를 활성화하십시오. API 작업을 직접적으로 호출할 때 MFA가 필요하면 정책에 MFA 조건을 추가합니다. 자세한 내용은 IAM 사용 설명서의 MFA 보호 API 액세스 구성을 참조하십시오.

IAM의 모범 사례에 대한 자세한 내용은 IAM 사용 설명서의 IAM의 보안 모범 사례를 참조하십시오.

Amazon Kendra용AWS 관리형 (미리 정의된) 정책

AWS 에서 생성하고 관리하는 독립형 IAM 정책을 제공하여 많은 일반적인 사용 사례를 해결합니다. AWS이러한 정책을 관리형 정책이라고 합니다 AWS . AWS 관리형 정책을 사용하면 정책을 직접 작성할 때보다 더 쉽게 사용자, 그룹 및 역할에 권한을 할당할 수 있습니다. 자세한 내용은 IAM 사용 설명서의 사용자에게 권한 추가를 참조하십시오.

계정의 그룹 및 역할에 연결할 수 있는 다음과 같은 AWS 관리형 정책은 Amazon Kendra에만 적용됩니다.

- AmazonKendraReadOnly— Amazon Kendra 리소스에 대한 읽기 전용 액세스 권한을 부여합니다.
- AmazonKendraFullAccess— 모든 Amazon Kendra 리소스를 생성, 읽기, 업데이트, 삭제, 태그 지정 및 실행할 수 있는 전체 액세스 권한을 부여합니다.

콘솔의 경우 역할에 `iam:CreateRole`, `iam:CreatePolicy`, `iam:AttachRolePolicy` 및 `s3:ListBucket` 권한도 있어야 합니다.

> **Note**
>
> IAM 콘솔에 로그인하고 이 콘솔에서 특정 정책을 검색하여 이러한 권한을 검토할 수 있습니다.

Amazon Kendra API 작업에 대한 권한을 허용하는 고유의 사용자 지정 정책을 생성할 수도 있습니다. 해당 권한이 필요한 IAM 역할 또는 그룹에 이러한 사용자 지정 정책을 연결할 수 있습니다. Amazon Kendra의 IAM 정책 예제는 Amazon Kendra 자격 증명 기반 정책 예제 섹션을 참조하세요.

사용자가 자신의 고유한 권한을 볼 수 있도록 허용

이 예제는 IAM 사용자가 자신의 사용자 ID에 연결된 인라인 및 관리형 정책을 볼 수 있도록 허용하는 정책을 생성하는 방법을 보여줍니다. 이 정책에는 콘솔에서 또는 API를 사용하여 프로그래밍 방식으로 이 작업을 완료할 수 있는 AWS CLI 권한이 포함됩니다. 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": "*"
        }
    ]
}
```

Amazon Kendra 인덱스에 액세스

이 예시에서는 AWS 계정의 사용자에게 색인을 쿼리할 수 있는 액세스 권한을 부여하려고 합니다.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "QueryIndex",
            "Effect": "Allow",
            "Action": [
                "kendra:Query"
            ],
            "Resource": "arn:aws:kendra:${Region}:${Account}:index/${Index ID}"
        }
    ]
}
```

태그 기반 정책 예제

태그 기반 정책은 보안 주체가 태그가 지정된 리소스에 대해 수행할 수 있는 작업을 지정하는 JSON 정책 문서입니다.

예제: 태그를 사용하여 리소스 액세스

이 예제 정책은 AWS 계정의 사용자 또는 역할에게 **department** 키와 값으로 **finance** 태그가 지정된 모든 리소스에서 Query 작업을 사용할 수 있는 권한을 부여합니다.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kendra:Query"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/department": "finance"
                }
            }
        }
```

```
    ]
}
```

예: 태그를 사용하여 Amazon Kendra 작업 활성화

이 예제 정책은 AWS 계정의 사용자 또는 역할에게 Amazon Kendra 작업을 사용할 권한을 부여합니다. 단, 키와 값으로 태그가 지정된 리소스를 사용하는 작업은 TagResource 예외입니다. **department finance**

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra:*",
            "Resource": "*"
        },
        {
            "Effect": "Deny",
            "Action": [
                "kendra:TagResource"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/department": "finance"
                }
            }
        }
    ]
}
```

예: 태그를 사용하여 작업에 대한 액세스 제한

이 예제 정책은 AWS 계정 내 사용자 또는 역할이 CreateIndex 작업을 사용할 수 있도록 액세스를 제한합니다. 단, 사용자가 **department** 태그를 제공하고 태그에 허용된 값 및 가 있는 경우는 예외입니다. **finance 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 아마존 켄드라의 관리형 정책

사용자, 그룹 및 역할에 권한을 추가하려면 정책을 직접 작성하는 것보다 AWS 관리형 정책을 사용하는 것이 더 쉽습니다. 팀에 필요한 권한만 제공하는 IAM 고객 관리형 정책을 생성하기 위해서는 시간과 전문 지식이 필요합니다. 빠르게 시작하려면 AWS 관리형 정책을 사용할 수 있습니다. 이러한 정책은 일반적인 사용 사례를 다루며 AWS 계정에서 사용할 수 있습니다. AWS 관리형 정책에 대한 자세한 내용은 IAM 사용 설명서의AWS 관리형 정책을 참조하십시오.

AWS 서비스는 AWS 관리형 정책을 유지 관리하고 업데이트합니다. AWS 관리형 정책에서는 권한을 변경할 수 없습니다. 서비스에서 때때로 추가 권한을 AWS 관리형 정책에 추가하여 새로운 기능을 지

원합니다. 이 유형의 업데이트는 정책이 연결된 모든 ID(사용자, 그룹 및 역할)에 적용됩니다. 서비스는 새로운 기능이 시작되거나 새 작업을 사용할 수 있을 때 AWS 관리형 정책에 업데이트됩니다. 서비스는 AWS 관리형 정책에서 권한을 제거하지 않으므로 정책 업데이트로 인해 기존 권한이 손상되지 않습니다.

또한 여러 서비스에 걸친 작업 기능에 대한 관리형 정책을 AWS 지원합니다. 예를 들어 ReadOnly액세스 AWS 관리형 정책은 모든 AWS 서비스와 리소스에 대한 읽기 전용 액세스를 제공합니다. 서비스가 새 기능을 시작하면 새 작업 및 리소스에 대한 읽기 전용 권한이 AWS 추가됩니다. 직무 정책의 목록과 설명은 IAM 사용 설명서의 직무에 관한AWS 관리형 정책을 참조하십시오.

AWS 관리형 정책: AmazonKendraReadOnly

Amazon Kendra 리소스에 대한 읽기 전용 액세스 권한을 부여합니다. 이 정책에는 다음 권한이 포함되어 있습니다.

- `kendra` – 사용자가 항목 목록 또는 항목에 대한 세부 정보를 반환하는 작업을 수행할 수 있습니다. 여기에는 `Describe`, `List`, `Query`, `BatchGetDocumentStatus`, `GetQuerySuggestions` 또는 `GetSnapshots`로 시작하는 API 작업이 포함됩니다.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "kendra:Describe*",
                "kendra:List*",
                "kendra:Query",
                "kendra:BatchGetDocumentStatus",
                "kendra:GetQuerySuggestions",
                "kendra:GetSnapshots"
            ],
            "Effect": "Allow",
            "Resource": "*"
        }
    ]
}
```

AWS 관리형 정책: AmazonKendraFullAccess

Amazon Kendra 리소스 모두를 생성하고, 읽고, 업데이트하고, 삭제하고, 태그를 지정하고 실행할 모든 액세스 권한을 부여합니다. 이 정책에는 다음 권한이 포함되어 있습니다.

- `kendra` - 보안 주체에게 Amazon Kendra의 모든 작업에 대한 읽기 및 쓰기 액세스 권한을 허용합니다.
- `s3` - 보안 주체가 Amazon S3 버킷 위치를 가져오고 버킷을 나열하도록 허용합니다.
- `iam` - 보안 주체가 역할을 전달하고 나열하도록 허용합니다.
- `kms`—주도자가 AWS KMS 키와 별칭을 설명하고 나열할 수 있습니다.
- `secretsmanager` - 보안 주체가 암호를 생성, 설명 및 나열하도록 허용합니다.
- `ec2` - 보안 주체가 보안 그룹, VCP(가상 사설 클라우드) 및 서브넷에 대해 설명하도록 허용합니다.
- `cloudwatch` - 보안 주체가 Cloud Watch 지표를 볼 수 있습니다.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "iam:PassRole",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "iam:PassedToService": "kendra.amazonaws.com"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "iam:ListRoles"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeVpcs",
```

```
                "ec2:DescribeSubnets"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kms:ListKeys",
                "kms:ListAliases",
                "kms:DescribeKey"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:ListAllMyBuckets",
                "s3:GetBucketLocation"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:ListSecrets"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "cloudwatch:GetMetricData"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "secretsmanager:CreateSecret",
                "secretsmanager:DescribeSecret"
            ],
            "Resource": "arn:aws:secretsmanager:*:*:secret:AmazonKendra-*"
        },
        {
```

```
            "Effect": "Allow",
            "Action": "kendra:*",
            "Resource": "*"
        }
    ]
}
```

Amazon Kendra, 관리형 정책 업데이트 AWS

이 서비스가 이러한 변경 사항을 추적하기 시작한 이후 Amazon Kendra의 AWS 관리형 정책 업데이트에 대한 세부 정보를 확인하십시오. 이 페이지의 변경 사항에 대한 자동 알림을 받아보려면 Amazon Kendra 문서 기록 페이지에서 RSS 피드를 구독하세요.

변경 사항	설명	날짜
AmazonKendraReadOnly—지원 권한 추가, API GetSnapshots BatchGetDocumentStatus	Amazon Kendra는 새로운 API `GetSnapshots` , `BatchGetDocumentStatus` 를 추가했습니다. `GetSnapshots` 은 사용자가 검색 애플리케이션과 상호 작용하는 방식을 보여주는 데이터를 제공합니다. `BatchGetDocumentStatus` 는 문서 인덱싱 진행 상황을 모니터링합니다.	2022년 1월 3일
AmazonKendraReadOnly—운영 지원을 위한 권한 추가 GetQuerySuggestions	Amazon Kendra는 인기 검색 쿼리에 대한 쿼리 제안을 받을 수 있는 액세스를 허용하는 새 API `GetQuerySuggestions` 를 추가하여 사용자의 검색을 안내하는 데 도움을 줍니다. 사용자가 검색 쿼리를 입력하면 추천 쿼리가 검색을 자동 완성하는 데 도움이 됩니다.	2021년 5월 27일

변경 사항	설명	날짜
Amazon Kendra에서 변경 사항 추적 시작	Amazon Kendra는 관리형 정책의 변경 사항을 추적하기 시작했습니다. AWS	2021년 5월 27일

Amazon Kendra 자격 증명 및 액세스 문제 해결

다음 정보를 사용하여 Amazon Kendra 및 IAM에서 작업할 때 발생할 수 있는 일반적인 문제를 진단하고 수정할 수 있습니다.

주제

- Amazon Kendra에서 태스크를 수행할 권한이 없음
- 저는 IAM을 수행할 권한이 없습니다. PassRole
- 관리자이며 다른 사용자가 Amazon Kendra에 액세스하도록 허용하려고 합니다.
- 내 AWS 계정 외부의 사용자가 내 Amazon Kendra 리소스에 액세스할 수 있도록 허용하고 싶습니다.

Amazon Kendra에서 태스크를 수행할 권한이 없음

작업을 수행할 권한이 없다는 AWS Management Console 메시지가 표시되면 관리자에게 도움을 요청해야 합니다. 관리자는 로그인 보안 인증 정보를 제공한 사람입니다.

다음 오류 예제는 `mateojackson` 사용자가 이 콘솔을 사용하여 인덱스에 대한 세부 정보를 보려고 하지만 `kendra:DescribeIndex` 권한이 없는 경우에 발생합니다.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 kendra:DescribeIndex on resource: index ARN
```

이 경우 Mateo는 `index` 작업을 사용하여 `kendra:DescribeIndex` 리소스에 액세스하도록 허용하는 정책을 업데이트하라고 관리자에게 요청합니다.

저는 IAM을 수행할 권한이 없습니다. PassRole

`iam:PassRole` 작업을 수행할 수 있는 권한이 없다는 오류가 수신되면 Amazon Kendra에 역할을 전달할 수 있도록 정책을 업데이트해야 합니다.

새 서비스 역할 또는 서비스 연결 역할을 만드는 대신 기존 역할을 해당 서비스에 전달할 AWS 서비스 수 있는 기능도 있습니다. 이렇게 하려면 사용자가 서비스에 역할을 전달할 수 있는 권한을 가지고 있어야 합니다.

다음 예제 오류는 marymajor라는 IAM 사용자가 콘솔을 사용하여 Amazon Kendra에서 태스크를 수행하려고 하는 경우에 발생합니다. 하지만 작업을 수행하려면 서비스 역할이 부여한 권한이 서비스에 있어야 합니다. Mary는 서비스에 역할을 전달할 수 있는 권한을 가지고 있지 않습니다.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

이 경우, Mary가 iam:PassRole 작업을 수행할 수 있도록 Mary의 정책을 업데이트해야 합니다.

도움이 필요하면 관리자에게 문의하세요. AWS 관리자는 로그인 자격 증명을 제공한 사람입니다.

관리자이며 다른 사용자가 Amazon Kendra에 액세스하도록 허용하려고 합니다.

다른 사용자가 Amazon Kendra에 액세스하도록 허용하려면 액세스 권한이 필요한 사용자 또는 애플리케이션에 대한 IAM 엔터티(사용자 또는 역할)를 생성해야 합니다. 다른 사용자들은 해당 엔터티에 대한 보안 인증을 사용해 AWS에 액세스합니다. 그런 다음 Amazon Kendra에서 올바른 권한을 부여하는 정책을 엔터티에 연결해야 합니다.

바로 시작하려면 IAM 사용 설명서의 첫 번째 IAM 위임 사용자 및 그룹 생성을 참조하십시오.

내 AWS 계정 외부의 사용자가 내 Amazon Kendra 리소스에 액세스할 수 있도록 허용하고 싶습니다.

다른 계정의 사용자 또는 조직 외부의 사람이 리소스에 액세스할 때 사용할 수 있는 역할을 생성할 수 있습니다. 역할을 수임할 신뢰할 수 있는 사람을 지정할 수 있습니다. 리소스 기반 정책 또는 액세스 제어 목록(ACL)을 지원하는 서비스의 경우 이러한 정책을 사용하여 다른 사람에게 리소스에 대한 액세스 권한을 부여할 수 있습니다.

자세히 알아보려면 다음을 참조하십시오.

- Amazon Kendra에서 이러한 기능을 지원하는지 여부를 알아보려면 Amazon Kendra에서 IAM을 사용하는 방법 섹션을 참조하세요.
- 소유하고 AWS 계정 있는 모든 리소스에 대한 액세스를 제공하는 방법을 알아보려면 IAM 사용 설명서의 다른 AWS 계정 IAM 사용자에게 액세스 권한 제공을 참조하십시오.
- 제3자에게 리소스에 대한 액세스 권한을 제공하는 방법을 알아보려면 IAM 사용 설명서의 타사 AWS 계정AWS 계정 소유에 대한 액세스 제공을 참조하십시오.

- ID 페더레이션을 통해 액세스 권한을 제공하는 방법을 알아보려면 IAM 사용 설명서의 외부에서 인증된 사용자에게 액세스 권한 제공(자격 증명 페더레이션)을 참조하십시오.
- 교차 계정 액세스에 대한 역할 사용과 리소스 기반 정책의 차이점을 알아보려면 IAM 사용 설명서의 IAM의 교차 계정 리소스 액세스를 참조하십시오.

보안 모범 사례

Amazon Kendra는 자체 보안 정책을 개발하고 구현할 때 고려해야 할 여러 보안 기능을 제공합니다. 다음 모범 사례는 일반적인 지침이며 완벽한 보안 솔루션을 나타내지는 않습니다. 이러한 모범 사례는 환경에 적절하지 않거나 충분하지 않을 수 있으므로 참고용으로만 사용하십시오.

최소 권한의 원칙 적용

Amazon Kendra는 역할을 사용하는 애플리케이션에 대한 세분화된 액세스 정책을 제공합니다. IAM 애플리케이션 조치와 로그 대상에 대한 액세스 등 작업에 필요한 최소한의 권한만 역할에 부여하는 것이 좋습니다. 또한 애플리케이션이 변경될 때마다 정기적으로 권한을 확인하기 위해 작업을 감사하는 것이 좋습니다.

역할 기반 액세스 제어(RBAC) 권한

관리자는 Amazon Kendra 애플리케이션에 대한 역할 기반 액세스 제어(RBAC) 권한을 엄격하게 제어해야 합니다.

Amazon Kendra의 로깅 및 모니터링

모니터링은 Amazon Kendra 애플리케이션의 안정성, 가용성 및 성능을 유지하는 중요한 역할을 합니다. Amazon Kendra API 호출을 모니터링하려면 를 사용할 수 있습니다. AWS CloudTrail작업 상태를 모니터링하려면 Amazon CloudWatch Logs를 사용하십시오.

- Amazon CloudWatch Alarms — CloudWatch 경보를 사용하면 지정한 기간 동안 단일 지표를 관찰할 수 있습니다. 지표가 정책을 초과하는 경우. CloudWatch 지표가 특정 상태일 때는 경보가 작업을 호출하지 않습니다. 대신, 상태가 변경되어 지정된 기간 동안 유지되어야 합니다. 자세한 정보는 Amazon CloudWatch를 사용한 Amazon Kendra 모니터링을 참조하세요.
- AWS CloudTrail 로그 — Amazon Kendra 또는 Amazon Kendra 인텔리전트 랭킹에서 사용자, 역할 또는 AWS 서비스가 수행한 작업에 대한 기록을 CloudTrail 제공합니다. 에서 수집한 정보를 사용하여 Amazon Kendra에 이루어진 요청 CloudTrail, 요청이 이루어진 IP 주소, 요청한 사람, 요청 시기

및 추가 세부 정보를 확인할 수 있습니다. 자세한 내용은 AWS CloudTrail 로그를 사용하여 Amazon Kendra API 호출 로깅 및 Amazon Kendra Intelligent Ranking API 호출을 AWS CloudTrail 로그와 함께 로깅 섹션을 참조하세요.

Amazon Kendra에 대한 규정 준수 확인

타사 감사자는 여러 Amazon Kendra 규정 준수 프로그램의 일환으로 Amazon Kendra의 보안 및 규정 준수를 평가합니다. Amazon Kendra는 다음을 준수합니다.

- HIPAA(미국 건강 보험 양도 및 책임에 관한 법)
- SOC(시스템 및 조직 제어) 2
- Information Security Registered Assessors Program(IRAP)
- 미국 동부/서부 리전의 연방 위험 및 권한 부여 관리 프로그램(FedRAMP) Moderate
- GovCloud AWS (미국 서부) 지역의 연방 위험 및 권한 관리 프로그램 (FedRAMP) 상위권

특정 규정 준수 프로그램 범위 내 AWS 서비스 목록은 규정 준수 프로그램별 범위 내 서비스 규정 준수 참조하십시오. 일반 정보는 AWS 규정 준수 프로그램AWS 보증 프로그램 규정AWS 참조하십시오.

를 사용하여 AWS Artifact타사 감사 보고서를 다운로드할 수 있습니다. 자세한 내용은 Artifact에서 보고서 다운로드 AWS Artifact에서 참조하십시오. AWS

Amazon Kendra 사용 시 규정 준수 책임은 데이터의 민감도, 회사의 규정 준수 목표 및 관련 법률과 규정에 따라 결정됩니다. AWS 는 규정 준수를 지원할 다음과 같은 리소스를 제공합니다.

- 보안 및 규정 준수 퀵 스타트 가이드 - 이 배포 가이드에서는 아키텍처 고려 사항을 설명하고 보안 및 규정 준수에 중점을 둔 기본 환경을 배포하기 위한 단계를 제공합니다. AWS
- HIPAA 보안 및 규정 준수를 위한 설계 백서 —이 백서는 기업이 HIPAA 준수 애플리케이션을 개발하는 데 사용할 수 있는 방법을 설명합니다. AWS
- AWS 규정 준수 리소스 규정 —이 통합 문서 및 가이드 모음은 해당 산업 및 지역에 적용될 수 있습니다.
- AWS Config 개발자 안내서의 규칙을 통한 리소스 평가 —이 AWS Config 서비스는 리소스 구성이 내부 관행, 업계 지침 및 규정을 얼마나 잘 준수하는지 평가합니다.
- AWS Security Hub—이 AWS 서비스는 보안 업계 표준 및 모범 사례를 준수하는지 확인하는 데 도움이 AWS 되는 내부 보안 상태를 종합적으로 보여줍니다.

Amazon Kendra의 복원성

AWS 글로벌 인프라는 지역 및 가용 영역을 중심으로 AWS 구축됩니다. AWS 지역은 물리적으로 분리되고 격리된 여러 가용 영역을 제공하며, 이러한 가용 영역은 지연 시간이 짧고 처리량이 높으며 중복성이 높은 네트워킹으로 연결됩니다. 가용 영역을 사용하면 중단 없이 영역 간에 자동으로 장애 극복 조치가 이루어지는 애플리케이션 및 데이터베이스를 설계하고 운영할 수 있습니다. 가용 영역은 기존의 단일 또는 다중 데이터 센터 인프라보다 가용성, 내결함성, 확장성이 뛰어납니다.

AWS 지역 및 가용 영역에 대한 자세한 내용은 글로벌 인프라를 참조하십시오AWS .

AWS 글로벌 인프라를 갖춘 Amazon Kendra 엔터프라이즈 에디션은 내결함성, 확장성, 고가용성을 제공합니다. 이전 버전의 인덱스로 롤백하는 것은 현재 지원되지 않지만 기존 데이터 소스를 삭제하고 인덱스에 다시 추가하여 인덱스의 일부를 새로 고치거나 다시 생성할 수 있습니다.

Amazon Kendra의 인프라 보안

관리형 서비스인 Amazon Kendra는 글로벌 네트워크 보안의 보호를 AWS 받습니다. AWS 보안 서비스 및 인프라 AWS 보호 방법에 대한 자세한 내용은 AWS 클라우드 보안을 참조하십시오. 인프라 보안 모범 사례를 사용하여 AWS 환경을 설계하려면 Security Pillar AWS Well-Architected Framework의 인프라 보호를 참조하십시오.

AWS 게시된 API 호출을 사용하여 네트워크를 통해 Amazon Kendra에 액세스할 수 있습니다. 고객은 다음을 지원해야 합니다.

- 전송 계층 보안(TLS) TLS 1.2는 필수이며 TLS 1.3을 권장합니다.
- DHE(Ephemeral Diffie-Hellman) 또는 ECDHE(Elliptic Curve Ephemeral Diffie-Hellman)와 같은 완전 전송 보안(PFS)이 포함된 암호 제품군 Java 7 이상의 최신 시스템은 대부분 이러한 모드를 지원합니다.

또한 요청은 액세스 키 ID 및 IAM 주체와 관련된 비밀 액세스 키를 사용하여 서명해야 합니다. 또는 AWS Security Token Service(AWS STS)를 사용하여 임시 보안 인증을 생성하여 요청에 서명할 수 있습니다.

의 구성 및 취약성 분석 AWS Identity and Access Management

AWS 게스트 운영 체제 (OS) 및 데이터베이스 패치, 방화벽 구성, 재해 복구와 같은 기본 보안 작업을 처리합니다. 적합한 제3자가 이 절차를 검토하고 인증하였습니다. 자세한 내용은 다음 리소스를 참조하세요.

- 공동 책임 모델
- AWS: 보안 프로세스 개요(백서)

다음 리소스는 AWS Identity and Access Management ()IAM의 구성 및 취약성 분석도 다룹니다.

- 규정 준수 검증: AWS Identity and Access Management
- 에서 보안 모범 사례 및 사용 사례를 참조하십시오 AWS Identity and Access Management.

에 대한 할당량 Amazon Kendra

지원되는 리전

사용 가능한 AWS Amazon Kendra 지역 목록은 Amazon Web Services 일반 참조의Amazon Kendra 지역 및 엔드포인트를 참조하십시오.

할당량

서비스 할당량 (한도라고도 함) 은 계정에 사용할 수 있는 최대 서비스 리소스 수입니다. AWS 자세한 내용을 알아보려면 AWS 일반 참조의 Amazon Kendra 서비스 할당량을 참조하세요.

인덱스 할당량

설명	기본값	Edition	조정 가능
계정당 최대 인덱스 수	10	개발자, 엔터프라이즈	예
단일 단위의 색인에 대해 추출된 텍스트의 양 (개발자). Developer Edition에서는 텍스트 추출을 위한 추가 단위를 추가할 수 없습니다.	3GB	개발자	아니요
단일 단위로 인덱스에 대해 추출된 텍스트의 양 (엔터프라이즈). Enterprise Edition용 텍스트 추출을 위해 최대 100개의 추가 단위를 추가하거나 지원팀에 문의하기만 하면 됩니다.	30GB	엔터프라이즈	예

데이터 소스 커넥터 할당량

설명	기본값	Edition	조정 가능
인덱스당 최대 데이터 소스 커넥터 수 (개발자)	5	개발자	아니요
인덱스당 최대 데이터 소스 커넥터 수 (엔터프라이즈)	50	엔터프라이즈	예
데이터 소스 커넥터 사용 시 단일 문서 또는 원시 파일의 최대 크기	50MB	개발자, 엔터프라이즈	예
Amazon S3 데이터 소스 커넥터에 포함된 액세스 제어 목록 구성 파일의 최대 S3 접두사 수	100	개발자, 엔터프라이즈	아니요
Amazon S3 데이터 소스 커넥터에 포함된 액세스 제어 목록 구성 파일의 최대 크기	50MB	개발자, 엔터프라이즈	예

FAQ 쿼터

설명	기본값	Edition	조정 가능
인덱스당 최대 FAQ 수	30	개발자, 엔터프라이즈	예
1 FAQ의 최대 크기	5MB	개발자, 엔터프라이즈	예
FAQ에 반환되는 최대 결과 수	4	개발자, 엔터프라이즈	예

설명	기본값	Edition	조정 가능
FAQ 질문에 허용되는 최대 문자 수	300	개발자, 엔터프라이즈	아니요
FAQ 답변의 최대 글자 수	2000	개발자, 엔터프라이즈	아니요

사전 할당량

설명	기본값	Edition	조정 가능
인덱스당 최대 동의어 사전 수	1	개발자, 엔터프라이즈	아니요
사전 파일의 최대 크기	5MB	개발자, 엔터프라이즈	예
사전당 최대 동의어 규칙 수	10,000개	개발자, 엔터프라이즈	예
인덱스에 포함된 모든 동의어 사전의 용어당 최대 동의어 수	10	개발자, 엔터프라이즈	아니요

Amazon Kendra 경험 할당량

설명	기본값	Edition	조정 가능
인덱스당 최대 Amazon Kendra 경험 수	50	개발자, 엔터프라이즈	예

쿼리 및 검색 결과 할당량

설명	기본값	Edition	조정 가능
단일 단위의 인덱스에 대한 초당 쿼리 수(개발자). Developer Edition에서는 쿼리를 위한 추가 단위를 추가할 수 없습니다.	0.05	개발자	아니요
단일 단위 인덱스의 초당 쿼리 양 (엔터프라이즈). Enterprise Edition용 쿼리를 위해 최대 100개의 추가 단위를 추가하거나 지원팀에 문의하기만 하면 됩니다.	0.1	엔터프라이즈	예
쿼리 텍스트당 최대 문자 수	1000	개발자, 엔터프라이즈	예
쿼리당 최대 검색 결과 수. 기본값은 100입니다. 100개 이상의 결과를 허용하려면 지원팀에 문의하면 됩니다.	100	개발자, 엔터프라이즈	예
페이지당 최대 검색 결과 수.	100	개발자, 엔터프라이즈	예
잘라내기 전의 쿼리 텍스트당 최대 토큰 단어 수. 기본값은 30입니다. 30개 넘는 단어를	30	개발자, 엔터프라이즈	예

설명	기본값	Edition	조정 가능
허용하려면 지원 팀에 문의하면 됩니다.			
쿼리 속성당 최대 사용자 그룹 목록 크기	1000	개발자, 엔터프라이즈	예
쿼리 속성당 최대 문자열 목록 크기	10	개발자, 엔터프라이즈	예

쿼리 제안 및 할당량

설명	기본값	Edition	조정 가능
제안 호출당 GetQuery 반환되는 최대 쿼리 제안 수	10	개발자, 엔터프라이즈	예
제안 호출당 쿼리 제안의 최대 필드/속성 수 GetQuery	10	개발자, 엔터프라이즈	예
제안 호출당 쿼리 제안을 위한 추가 필드/속성의 최대 수 GetQuery	5	개발자, 엔터프라이즈	예
인덱스당 최대 차단 목록 수	1	개발자, 엔터프라이즈	아니요
차단 목록 텍스트 파일의 최대 크기	2MB	개발자, 엔터프라이즈	예
차단 목록에 있는 최대 항목(단어 또는 문구) 수	20,000건	개발자, 엔터프라이즈	예

설명	기본값	Edition	조정 가능
Query API 직접 호출에서 반환할 맞춤법 수정 쿼리 제안의 최대 수.	1	개발자, 엔터프라이즈	예

문서 할당량

설명	기본값	Edition	조정 가능
단일 단위의 색인에 대해 추출된 텍스트의 양 (개발자). Developer Edition에서는 텍스트 추출을 위한 추가 단위를 추가할 수 없습니다.	3GB	개발자	아니요
단일 단위로 인덱스에 대해 추출된 텍스트의 양 (엔터프라이즈). Enterprise Edition용 텍스트 추출을 위해 최대 100개의 추가 단위를 추가하거나 지원팀에 문의하기만 하면 됩니다.	30GB	엔터프라이즈	예
데이터 소스 커넥터 사용 시 단일 문서 또는 원시 파일의 최대 크기	50MB	개발자, 엔터프라이즈	예
BatchPutDocument API 사용	5MB	개발자, 엔터프라이즈	예

설명	기본값	Edition	조정 가능
시 단일 문서 또는 원시 파일의 최대 크기			
단일 문서에서 추출할 수 있는 최대 텍스트 양	5MB	개발자, 엔터프라이즈	아니요
인덱스당 최대 사용자 지정 필드/속성 수	500	개발자, 엔터프라이즈	아니요

추천 검색 결과 할당량

설명	기본값	Edition	조정 가능
추천 결과 세트당 추천 문서의 최대 수	4	엔터프라이즈	예
추천 결과 세트당 최대 쿼리 텍스트 수	49	엔터프라이즈	아니요
추천 결과 세트의 쿼리 텍스트당 최대 문자 수	1000	엔터프라이즈	예
인덱스당 추천 결과 세트의 최대 수	50	엔터프라이즈	예

검색 결과 할당량 재점수/순위 조정

설명	기본값	Edition	조정 가능
재평가 실행 계획 또는 단일 용량 단위에 대한 초당 최대 `Rescore` 요청 수. 최대 1000 단	0.01	엔터프라이즈	아니요

설명	기본값	Edition	조정 가능
위를 추가할 수 있습니다.			
계정당 최대 재평가 실행 계획 수.	50	엔터프라이즈	예
Rescore 요청 시 문서 하나의 Title에 포함되는 최대 토큰 수.	100	엔터프라이즈	아니요
Rescore 요청 시 문서 하나의 Body에 포함되는 최대 토큰 수.	200	엔터프라이즈	아니요
Rescore 요청당 최대 문서 수.	25	엔터프라이즈	아니요
Rescore 요청 내 그룹당 최대 문서 수.	3	엔터프라이즈	아니요

서비스 할당량에 대한 자세한 내용과 할당량 증가를 요청하려면 Amazon Kendra Service Quotas를 참조하십시오.

문제 해결

이 섹션은 작업할 때 발생할 수 있는 일반적인 문제를 해결하는 데 도움이 될 수 Amazon Kendra있습니다.

주제

데이터 소스 문제 해결

이 섹션에서는 Amazon Kendra 데이터 소스 커넥터를 구성하고 사용할 때 발생하는 일반적인 문제를 해결하는 데 도움이 될 수 있습니다.

내 문서가 인덱싱되지 않았습니다.

Amazon Kendra 색인을 데이터 원본과 동기화할 때 문서가 색인되지 않는 문제가 발생할 수 있습니다. 인덱싱은 두 단계로 이루어져 있습니다. 먼저 데이터 소스에서 인덱스를 생성할 새 문서 및 업데이트된 문서가 있는지 확인하고 인덱스에서 제거할 문서를 찾습니다. 둘째, 문서 수준에서 각 문서에 액세스하고 인덱싱합니다.

두 단계 중 하나에서 오류가 발생할 수 있습니다. 데이터 소스 수준 오류는 콘솔의 데이터 소스 세부정보 페이지의 동기화 실행 기록 섹션에 보고됩니다. 동기화 작업의 상태는 성공, 미완료 또는 실패일 수 있습니다. 또한 작업 중에 인덱싱되고 삭제된 문서 수를 볼 수 있습니다. 상태가 실패인 경우 세부 정보 열에 메시지가 표시됩니다.

문서 수준 오류는 에서 보고됩니다. Amazon CloudWatch Logs CloudWatch 콘솔을 사용하여 오류를 확인할 수 있습니다.

문서 동기화 상태 보고서를 생성하려면 내 문서에 대한 동기화 상태 보고서를 생성하고자 함을 참조하세요.

동기화 작업이 실패함

동기화 작업은 일반적으로 인덱스 또는 데이터 소스에 구성 오류가 있는 경우 실패합니다. 콘솔의 세부정보 열 아래에 있는 데이터 소스 세부 정보 페이지의 동기화 실행 기록 섹션에서 오류 메시지를 찾을

수 있습니다. 문서 수준 오류는 Amazon CloudWatch Logs에 보고됩니다. 오류 메시지는 무엇이 잘못되었는지에 대한 정보를 제공합니다. 문제는 일반적으로 인덱스 또는 데이터 원본에 적절한 IAM 권한이 없다는 것입니다. 오류 메시지에 누락된 권한이 설명되어 있습니다. 수신할 수 있는 몇 가지 오류 메시지는 다음과 같습니다.

```
Failed to create log group for job. Please make sure that the IAM role
provided has sufficient permissions.
```

인덱스 역할에 사용 CloudWatch권한이 없는 경우 데이터 원본은 CloudWatch 로그를 만들 수 없습니다. 이 오류가 발생하면 인덱스 역할에 CloudWatch 권한을 추가해야 합니다.

```
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.
```

Amazon S3 데이터 원본을 사용하는 경우 문서가 포함된 버킷에 액세스할 수 있는 권한이 Amazon Kendra 있어야 합니다. 버킷을 읽을 수 Amazon Kendra 있는 권한을 데이터 원본 IAM 역할에 추가해야 합니다.

```
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 인덱스 및 데이터 소스 IAM 역할을 맡을 권한이 필요합니다. `sts:AssumeRole` 작업에 대한 권한이 있는 역할에 신뢰 정책을 추가해야 합니다.

데이터 원본을 Amazon Kendra 인덱싱해야 하는 IAM 정책에 대해서는 IAM 역할을 참조하십시오.

문서 동기화 상태 보고서를 생성하려면 내 문서에 대한 동기화 상태 보고서를 생성하고자 함을 참조하세요.

동기화 작업이 완료되지 않음

데이터 소스 수준 프로세스를 완료했지만 문서 수준 프로세스에서 약간의 오류가 발생한 경우, 작업은 일반적으로 완료되지 않습니다. 작업이 완료되지 않은 경우 일부 문서의 인덱스가 제대로 생성되지 않았을 수 있습니다. Amazon S3 데이터 소스의 경우 미완료 작업은 일반적으로 다음과 같은 원인으로 발생합니다.

- 하나 이상의 문서에 대한 메타데이터가 잘못되었습니다.
- 인덱싱을 위해 문서를 제출했지만 하나 이상의 문서가 제출되지 않은 경우.

- 인덱스에서 삭제하기 위해 문서를 제출했지만 하나 이상의 문서가 제출되지 않은 경우.

불완전한 동기화 작업의 문제를 해결하려면 먼저 로그를 살펴보세요. CloudWatch

1. 세부 정보 열에서 세부 정보 보기를 선택합니다. CloudWatch
2. 오류 메시지를 검토하여 문서 오류의 원인을 확인하세요.

문서 동기화 상태 보고서를 생성하려면 내 문서에 대한 동기화 상태 보고서를 생성하고자 함을 참조하세요.

동기화 작업은 성공했지만 인덱싱된 문서가 없음

인덱스 동기화 작업 실행이 성공한 것으로 표시되지만 예상했던 인덱싱된 새 문서나 업데이트된 문서가 없는 경우가 있습니다. 가능한 이유는 다음과 같습니다.

- CloudWatch `DocumentsSubmittedForIndexingFailed`지표를 확인하여 동기화에 실패한 문서가 있는지 확인하십시오. CloudWatch 로그에서 자세한 내용을 확인하세요.
- Amazon S3 데이터 원본의 경우 잘못된 버킷 이름이나 접두사를 Amazon Kendra 지정했을 수 있습니다. 사용 중인 Amazon Kendra 버킷이 인덱싱할 문서를 포함하는 버킷인지 확인하십시오.
- 이전 작업에서 인덱싱하지 못한 문서를 다시 인덱싱할 때는 문서 또는 관련 메타데이터 파일을 변경하지 않는 한, Amazon Kendra 가 인덱싱하지 않습니다.

문서 동기화 상태 보고서를 생성하려면 내 문서에 대한 동기화 상태 보고서를 생성하고자 함을 참조하세요.

데이터 소스를 동기화하는 동안 파일 형식 문제가 발생함

데이터 소스에 파일을 추가하거나 데이터 소스를 동기화하는 동안 파일 형식 문제가 발생하는 경우 문서 유형이 Amazon Kendra 를 지원하는지 확인하세요. 에서 지원하는 문서 유형 목록은 문서 유형 또는 형식을 Amazon Kendra 참조하십시오.

일반 텍스트 파일과 함께 `BatchPutDocument` API를 사용하는 경우 콘텐츠 유형으로 `PLAIN_TEXT`를 지정하세요.

내 문서에 대한 동기화 기록 보고서를 생성하고자 함

Amazon Kendra 데이터 소스 커넥터를 동기화할 때 데이터 원본의 각 문서에 대한 동기화 상태 보고서를 생성하여 Amazon S3 버킷에 복사할 Amazon Kendra 수 있습니다. 이 과정에서 데이터는 AWS

KMS 키를 사용하여 암호화되며 사용자만 볼 수 있습니다. 보고된 문서 상태는 실패, 완료 또는 성공(오류 있음) 중 하나일 수 있습니다.

동기화 상태 보고서를 생성하는 경우 먼저 다음을 수행해야 합니다.

- Amazon S3 액세스 정책에 다음 Amazon Kendra 서비스 보안 주체를 추가합니다.

```
{
    “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/*”
        }
    ]
}
```

- 다음과 같은 액세스 권한이 있는 Amazon S3 버킷을 생성합니다. Amazon Kendra

콘솔을 사용하여 동기화 상태 보고서를 생성하는 경우 데이터 소스 세부 정보 페이지에서 동기화 기록 생성 옵션을 활성화하도록 선택합니다. 그런 다음 Amazon S3 버킷 위치를 입력하고 사용 가능한 구성 옵션 중에서 선택합니다. 보고서 생성을 활성화하면 다음번 동기화에서 보고서가 생성됩니다.

Amazon S3 버킷을 삭제하면 로그 데이터가 손실되므로 새 동기화 보고서를 저장할 새 버킷을 설정해야 합니다.

동기화 보고서 생성 상태는 현재 오직 Amazon S3 커넥터에서만 지원됩니다.

데이터 소스를 동기화하는 데 시간이 얼마나 걸리나요?

문서가 업데이트되지 않는 경우 Amazon Kendra 인덱스의 동기화 시간은 문서 수에 비례하여 선형적으로 증가합니다. 예를 들어 업데이트가 없는 문서 1,000개는 동기화하는 데 약 5분이 걸리고, 업데이트가 없는 문서 2,000개는 동기화하는 데 약 10분이 걸립니다. 문서가 업데이트된 경우 업데이트된 문서 수에 따라 동기화 시간이 늘어납니다.

데이터 소스를 동기화하는 데 드는 요금은 얼마인가요?

색인을 동기화할 때 필요한 연결을 Amazon EC2 설정하기 위해 워밍업하고 활성화하는 데 2분이 걸립니다. 이 프로세스에는 요금이 청구되지 않습니다. 사용량 측정기는 동기화 작업이 시작된 후에만 시작됩니다. Amazon Kendra 요금에 대한 자세한 내용은 가격 Amazon Kendra 책정을 참조하십시오.

Amazon EC2 인증 오류가 발생합니다.

VPC (가상 사설 클라우드) 데이터 원본에 대한 동기화 중에 Amazon EC2 무단 작업 오류가 발생하는 경우 VPC IAM 역할에 필요한 권한이 없을 수 있습니다. 데이터 소스에 사용하는 IAM 역할에 연결된 권한이 있는지 확인하세요. 자세한 내용은 가상 사설 클라우드 IAM 역할을 참조하십시오.

검색 색인 링크를 사용하여 Amazon S3 객체를 열 수 없습니다.

Amazon Kendra 색인은 Amazon S3 데이터 원본에서 액세스 권한을 부여한 파일에만 액세스할 수 있습니다. 예를 들어, 객체를 공개할지 아니면 암호화할지를 결정하는 Amazon S3 권한을 수정할 수 Amazon Kendra 없습니다. Amazon Kendra 또한 Amazon S3 개체에 대해 서명된 링크를 만들거나 반환할 수 있는 기본 권한도 없습니다. Amazon Kendra 색인의 Amazon S3 객체에 대해 서명된 링크를 활성화하려는 경우 다음 두 가지 옵션이 있습니다.

- 결과를 검색 페이지에 반환하기 전에 소스 uri 객체를 사용하여 인덱스 쿼리 결과에 서명할 수 있습니다. 이 프로세스를 step-by-step 단계별로 살펴보려면 미리 서명된 URL을 사용한 개체 공유를 참조하십시오.
- Amazon S3 객체 메타데이터 소스 uri를 재정의하고 버킷에 연결된 CDN (CloudFront 콘텐츠 전송 네트워크) 을 통해 서비스를 사용할 수 있도록 할 수 있습니다. Amazon S3 또는 미리 서명된 URL을 반환하고 해당 URL로 리디렉션하는 API Gateway 프록시 엔드포인트를 사용할 수 있습니다.

SSL 인증서 파일 사용 AccessDenied 시 오류 메시지가 나타납니다.

데이터 원본과 함께 SSL 인증서를 사용할 때 액세스 거부 오류가 발생하는 경우, 해당 IAM 역할에 지정된 위치에 있는 SSL 인증서 파일에 액세스할 수 있는 권한이 있는지 확인하세요. 인증서가 키로 암호화된 경우 IAM 역할에는 AWS KMS 키를 사용하여 암호를 해독할 수 있는 권한도 있어야 합니다 AWS KMS . 자세한 내용은 AWS KMS에 대한 인증 및 액세스 제어를 참조하세요.

데이터 원본을 사용할 때 인증 오류가 발생합니다. SharePoint

색인을 SharePoint 데이터 원본과 동기화하는 동안 권한 부여 오류가 발생하는 경우 사이트 관리자 역할이 할당되었는지 확인하세요. SharePoint

내 인덱스가 Confluence 데이터 소스의 문서를 크롤링하지 않음

동기화 프로세스 중에 Amazon Kendra 색인이 Confluence 데이터 원본의 문서를 크롤링하지 않는 경우 Confluence의 관리자 그룹에 속해 있는지 확인하십시오.

문서 검색 결과의 문제 해결

이 섹션은 검색 결과의 문제를 해결하는 데 도움이 될 수 있습니다. Amazon Kendra

내 검색 결과가 내 검색 쿼리와 관련이 없음

검색 결과가 관련이 없어 보인다면 다음과 같은 이유 때문일 수 있습니다.

- 신뢰도가 LOW인 결과가 검색 결과에 포함됩니다. QueryResultItem's ScoreAttributes 필드를 사용하여 값이 인 모든 결과를 제외하면 결과를 LOW 확실하게 필터링할 수 LOW 있습니다. Amazon Kendra 각 결과에 VERY_HIGHHIGH, MEDIUM 및 LOW 중 하나의 신뢰 버킷 값을 할당합니다. 이 값은 결과가 쿼리와 얼마나 관련 있는지를 뜻하는 신뢰도 수준을 나타냅니다. 또한 신뢰 버켓에 관계없이 ANSWER (제안된 답변 발췌), (FAQ), QUESTION_ANSWER (문서 발췌) 의 순서로 세 가지 유형의 결과를 Amazon Kendra 반환합니다. DOCUMENT 따라서 LOW 신뢰도의 QUESTION_ANSWER 결과가 VERY_HIGH 신뢰도의 DOCUMENT 결과 위에 위치할 수 있습니다. 하지만 LOW 신뢰도의 QUESTION_ANSWER가 항상 VERY_HIGH 신뢰도 DOCUMENT보다 더 나은 결과는 아닙니다.
- 특정 메타데이터 필드 또는 속성은 매우 높은 값으로 부스트되어 결과 순위에 영향을 줍니다. Amazon Kendra 문서 제목, 텍스트, 날짜, 사용자 지정 텍스트 필드 또는 속성 등의 여러 매개 변수를 사용하여 색인을 검색합니다. 다양한 부스팅 값을 실험하여 모든 쿼리에서 최상의 결과를 얻을 수 있습니다. 또한 쿼리 수준에서 동적 관련성 조정을 사용하여 각 쿼리마다 다른 부스팅 값을 사용할 수도 있습니다.
- 사용자들은 정보를 쿼리할 때 특수 용어를 사용하고 있으나 이러한 특수 용어를 처리하기 위해 인덱스에 설정된 사용자 지정 동의어가 없습니다. 동의어 사용 방법 및 시기에 대한 자세한 내용은 인덱스에 사용자 지정 동의어 추가를 참조하세요.

결과가 100개만 표시되는 이유는 무엇인가요?

Amazon Kendra 관련 문서의 총 개수를 반환합니다. 기본적으로 쿼리당 상위 100개가 반환됩니다. 쿼리 결과가 페이지 매김됩니다. PageNumber를 사용하여 다양한 페이지에 액세스할 수 있습니다.

쿼리당 최대 1,000개의 문서 또는 검색 결과를 반환하고 페이지당 최대 100개의 결과를 Amazon Kendra 반환하도록 구성할 수 있습니다. 100개가 넘는 결과를 반환하려면 할당량 지원 팀에 문의하여 결과를 요청할 수 있습니다. 검색 결과 수를 늘리면 지연 시간에 영향을 미칠 수 있습니다.

표시될 것으로 예상되는 문서가 왜 누락되었나요?

Amazon Kendra 사용자 및 그룹을 기반으로 하는 ACL (액세스 제어 목록) 을 지원합니다. Amazon Kendra 커넥터를 통해 ACL 정책을 수집합니다. 인덱스가 ACL을 구성하지 않는 경우 사용자 및 그룹의 속성 필터와 일치하는 문서만 표시됩니다. 사용자 또는 그룹 속성 필터를 제공하는 경우 ACL이 없는 문서는 표시되지 않습니다.

토큰 기반 액세스 제어를 사용하는 경우 ACL 정책이 없는 문서, 사용자 및 그룹과 일치하는 문서가 표시됩니다.

ACL 정책이 적용된 문서가 표시되는 이유는 무엇입니까?

인덱스가 액세스 제어 정책을 구성하지 않는 경우 필터를 통해 사용자 및 그룹을 제공할 수 있습니다. 사용자 및 그룹 필터를 적용하지 않으면 모든 관련 문서가 반환됩니다. 모든 ACL 정책은 무시됩니다.

일반적인 문제 해결

Amazon Kendra CloudWatch 지표와 로그를 사용하여 데이터 소스 동기화에 대한 통찰력을 제공합니다. 지표와 로그를 사용하여 동기화 실행에서 발생한 문제와 해결 방법을 확인할 수 있습니다.

일반적인 문제 해결의 경우 CloudWatch 지표부터 시작하세요.

- `DocumentsCrawled` 지표를 확인하여 데이터 소스에서 확인한 문서 수를 확인하세요. Amazon S3 버킷의 경우 숫자가 예상보다 적으면 데이터 소스가 올바른 버킷을 가리키고 있는지 확인하세요.
- `DocumentsSkippedNoChange` 지표를 확인하여 마지막 동기화 이후 변경되지 않아 건너뛰었던 문서 수를 확인하세요. 숫자가 예상과 일치하지 않는 경우 리포지토리가 올바르게 업데이트되었는지 확인하세요.
- `DocumentsSkippedInvalidMetadata` 지표를 확인하여 잘못된 메타데이터가 있는 문서 수를 확인하세요. CloudWatch 로그를 확인하여 발생한 특정 오류를 확인하세요.
- `DocumentsSubmittedForIndexingFailed` 지표를 확인하여 데이터 소스에서 인덱스로 전송되었지만 인덱싱에 실패한 문서 수를 확인하세요. 예를 들어 사용자 지정 인덱스 필드로 정의되지 않은 Amazon S3 데이터 소스의 메타데이터 속성을 사용하면 문서가 인덱싱되지 않습니다. CloudWatch 로그를 확인하여 발생한 특정 오류를 확인하세요.
- `DocumentsSubmittedForDeletionFailed` 지표를 확인하여 데이터 소스가 인덱스에서 제거하려고 시도했으나 인덱스에서 삭제되지 못한 문서 수를 확인하세요. CloudWatch 로그를 확인하여 발생한 특정 오류를 확인하세요.

특정 동기화 실행의 CloudWatch 로그를 보면 실행 중에 발생한 오류의 세부 정보를 얻을 수 있습니다. 를 사용한 CloudWatch 로그에 대한 자세한 내용은 Amazon Kendra을 참조하십시오 CloudWatch Logs.

Amazon Kendra Intelligent Ranking

Amazon Kendra Intelligent Ranking은 Amazon Kendra 시맨틱 검색 기능을 사용하여 검색 서비스 결과의 순위를 지능적으로 재지정합니다.

주제

Amazon Kendra 자체 관리를 위한 지능형 랭킹 OpenSearch

Amazon Kendra의 시맨틱 검색 기능을 활용하여 Apache 2.0 라이선스를 기반으로 하는 자체 관리형 오픈 소스 검색 서비스의 검색 결과를 개선할 수 있습니다. OpenSearch Amazon Kendra 인텔리전트 랭킹 플러그인은 를 사용하여 결과의 순위를 의미적으로 다시 매깁니다. OpenSearch Amazon Kendra 기본 검색 결과에서 문서 본문이나 제목과 같은 특정 필드를 사용하여 검색 쿼리의 의미와 컨텍스트를 이해함으로써 이를 수행합니다. OpenSearch

'main keynote address'라는 쿼리를 예로 들어 보겠습니다. '주소'에는 여러 가지 의미가 있으므로 쿼리의 의미를 유추하여 의도한 의미와 일치하는 관련 정보를 반환할 Amazon Kendra 수 있습니다. 이 맥락에서는, 컨퍼런스 기조 연설입니다. 예를 들어 단순한 검색 서비스는 의도를 고려하지 않아 메인 스트리트에 있는 도로명 주소에 대한 결과를 반환할 수 있습니다.

용 OpenSearch 인텔리전트 랭킹 플러그인은 OpenSearch (자체 관리형) 버전 2.4.0 이상에서 사용할 수 있습니다. 퀵 스타트 Bash 스크립트를 사용하여 플러그인을 설치하여 인텔리전트 랭킹 플러그인이 OpenSearch 포함된 새 Docker 이미지를 빌드할 수 있습니다. 지능형 검색 플러그인 설정 참조 - 빠르게 시작하고 실행할 수 있는 설정의 예입니다.

지능형 검색 플러그인 작동 방식

OpenSearch (자체 관리형) 용 인텔리전트 랭킹 플러그인의 전체 프로세스는 다음과 같습니다.

1. OpenSearch 사용자가 쿼리를 실행하고 쿼리 응답 또는 쿼리와 관련된 문서 목록을 OpenSearch 제공합니다.
2. Intelligent Ranking 플러그인은 쿼리 응답을 받아 문서에서 정보를 추출합니다.
3. 인텔리전트 랭킹 플러그인은 Amazon Kendra 인텔리전트 랭킹의 Rescore API를 호출합니다.

4. Rescore API는 문서에서 추출된 정보를 가져와 검색 결과의 순위를 의미적으로 다시 매깁니다.
5. Rescore API는 순위가 다시 매겨진 검색 결과를 플러그인에 다시 보냅니다. 플러그인은 새로운 시맨틱 순위를 반영하도록 검색 응답의 OpenSearch 검색 결과를 재정렬합니다.

Intelligent Ranking 플러그인은 “본문” 및 “제목” 필드를 사용하여 결과의 순위를 다시 매깁니다. 이러한 플러그인 필드를 문서 본문 및 제목의 정의에 가장 적합한 OpenSearch 색인의 필드에 매핑할 수 있습니다. 예를 들어 인덱스에 “chapter_heading” 및 “chapter_contents”와 같은 필드가 있는 책의 장이 포함되어 있는 경우 전자를 “제목”에 매핑하고 후자를 “본문”에 매핑하여 최상의 결과를 얻을 수 있습니다.

지능형 검색 플러그인 설정

다음은 인텔리전트 랭킹 플러그인을 사용하여 신속하게 설정 OpenSearch (자체 관리) 하는 방법을 간략하게 설명합니다.

인텔리전트 랭킹 플러그인으로 설정 OpenSearch (자체 관리) (빠른 설정)

이미 Docker 이미지를 opensearch:2.4.0 사용하고 있다면 이 Dockerfile을 사용하여 인텔리전트 랭킹 플러그인으로 OpenSearch 2.4.0의 새 이미지를 빌드할 수 있습니다. docker-compose.yml 파일 또는 opensearch.yml 파일에 새 이미지를 위한 컨테이너를 포함합니다. 재평가 실행 계획을 만들 때 생성된 재평가 실행 계획 ID를 리전 및 엔드포인트 정보와 함께 포함시킬 수도 있습니다. 재평가 실행 계획을 만들려면 2단계를 참조하세요.

2.4.0 이전의 opensearch 도커 이미지 버전을 이전에 다운로드한 경우 도커 이미지 opensearch:2.4.0 이상을 사용하고 Intelligent Ranking 플러그인이 포함된 새 이미지를 빌드해야 합니다.

1. 사용 중인 운영 체제용 Docker Desktop을 다운로드하여 설치합니다. Docker Desktop에는 Docker Compose 및 Docker Engine이 포함되어 있습니다. 컴퓨터가 Docker 설치 세부 정보에 언급된 시스템 요구 사항을 충족하는지 확인하는 것이 좋습니다.

   Docker 데스크톱의 설정에서 메모리 사용 요구 사항을 늘릴 수도 있습니다. 무료로 제공되는 Docker 서비스 사용 한도 이외의 Docker 사용 요구 사항은 사용자의 책임입니다. Docker 구독을 참조하세요.

   Docker 데스크톱 상태가 “실행 중”인지 확인하세요.
2. Amazon Kendra 인텔리전트 랭킹 및 용량 요구 사항을 제공하세요. Amazon Kendra Intelligent Ranking을 프로비저닝하면 설정된 용량 단위를 기준으로 시간당 요금이 부과됩니다. 프리 티어 및 요금 정보를 참조하세요.

CreateRescoreExecutionPlanAPI를 사용하여 프로비저닝합니다Rescore API. 단일 단위 기본값보다 더 많은 용량 단위가 필요하지 않은 경우 단위를 더 추가하지 말고 재평가 실행 계획의 이름만 제공하세요. UpdateRescoreExecutionPlanAPI를 사용하여 용량 요구 사항을 업데이트할 수도 있습니다. 자세한 내용은 검색 서비스 결과의 의미론적 순위 지정을 참조하세요.

필요한 경우 빠른 시작 Bash 스크립트를 실행할 때 3단계로 이동하여 기본 재평가 실행 계획을 만들 수 있습니다.

4단계의 경우 응답에 포함된 재평가 실행 계획 ID를 참고하세요.

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. 기본 브랜치 드롭다운에서 GitHub 버전 브랜치를 OpenSearch 선택하여 사용 중인 버전의 퀵 스타트 Bash 스크립트를 다운로드하십시오.

   이 스크립트는 스크립트용 GitHub 리포지토리에서 선택한 버전을 사용하는 Docker OpenSearch 이미지와 OpenSearch 대시보드를 사용합니다. 인텔리전트 랭킹 플러그인용 zip 파일을 다운로드하고 플러그인이 OpenSearch 포함된 새 Docker 이미지를 `Dockerfile` 빌드하기 위한 zip 파일을 생성합니다. 또한 인텔리전트 랭킹 플러그인 및 대시보드를 위한 OpenSearch 컨테이너가 포함된 docker-compose.yml 파일을 생성합니다. OpenSearch 스크립트는 재평가 실행 계획 ID, 리전 정보, 엔드포인트(리전 사용)를 docker-compose.yml 파일에 추가합니다. 그런 다음 스크립트를

`docker-compose up` 실행하여 인텔리전트 랭킹 및 대시보드가 포함된 컨테이너를 시작합니다. OpenSearch OpenSearch 컨테이너를 제거하지 않고 중지하려면 `docker-compose stop`를 실행하세요. 컨테이너를 제거하려면 `docker-compose down`를 실행합니다.

4. 터미널을 열고 Bash 스크립트의 디렉터리에서 다음 명령을 실행합니다.

```
bash search_processing_kendra_quickstart.sh -p <execution-plan-id> -r <region>
```

이 명령을 실행할 때는 Amazon Kendra 인텔리전트 랭킹을 프로비저닝할 때 2단계에서 기록해 둔 재점수 실행 계획 ID를 지역 정보와 함께 제공합니다. `--create-execution-plan` 옵션을 사용하여 Amazon Kendra Intelligent Ranking을 대신 프로비저닝할 수도 있습니다. 이렇게 하면 기본 이름과 기본 용량을 사용하여 재평가 실행 계획이 생성됩니다.

기본 임시 컨테이너를 제거해도 인덱스가 손실되지 않도록 하려면 `--volume-name` 옵션을 사용하여 데이터 볼륨 이름을 제공하여 실행 시에도 인덱스가 지속되도록 할 수 있습니다. 이전에 인덱스를 만든 경우 docker-compose.yml 또는 opensearch.yml 파일에서 볼륨을 지정할 수 있습니다. 볼륨을 그대로 두려면 `docker-compose down -v`를 실행하지 마세요.

퀵 스타트 배쉬 스크립트는 인텔리전트 랭킹에 연결되도록 OpenSearch 키스토어의 AWS 자격 증명을 구성합니다. Amazon Kendra 스크립트에 AWS 자격 증명을 제공하려면 `--profile` 옵션을 사용하여 프로필을 지정합니다. AWS `--profile`옵션이 지정되지 않은 경우 빠른 시작 Bash 스크립트는 환경 변수와 기본 프로필에서 AWS 자격 증명 (액세스/비밀 키, 선택적 세션 토큰) 을 읽으려고 시도합니다. AWS `--profile`옵션을 지정하지 않고 자격 증명을 찾을 수 없는 경우 스크립트는 자격 증명을 키 저장소에 전달하지 않습니다. OpenSearch OpenSearch 키스토어에 자격 증명이 지정되지 않은 경우에도 플러그인은 메타데이터 서비스를 통해 전달된 Amazon ECS 컨테이너 자격 증명 또는 인스턴스 프로필 자격 증명을 포함하여 기본 자격 증명 공급자 체인의 자격 증명을 계속 확인합니다. Amazon EC2

인텔리전트 랭킹을 Amazon Kendra 호출하는 데 필요한 권한이 있는 IAM 역할을 생성했는지 확인하십시오. 다음은 특정 재점수 실행 계획에 `Rescore` API 사용 권한을 부여하는 IAM 정책의 예입니다.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra-ranking:Rescore",
            "Resource": "arn:aws:kendra-ranking:${Region}:${Account}:rescore-
execution-plan/${RescoreExecutionPlanId}"
```

```
        }
    ]
}
```

docker-compose.yml의 예

인텔리전트 랭킹 플러그인 및 대시보드 2.4.0 이상에서 OpenSearch 2.4.0 이상을 사용하는 docker-compose.yml 파일의 예입니다. 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
```

Dockerfile 및 이미지 빌드의 예

인텔리전트 랭킹 플러그인과 함께 2.4.0 이상을 사용하는 예제입니다. `Dockerfile` OpenSearch

```
FROM opensearchproject/opensearch:<your-version>
RUN /usr/share/opensearch/bin/opensearch-plugin install --batch  https://github.com/
opensearch-project/search-processor/releases/download/<your-version>/search-
processor.zip
```

인텔리전트 랭킹 플러그인을 OpenSearch 사용하여 Docker 이미지를 빌드합니다.

```
docker build --tag=<Docker image tag name of OpenSearch with Intelligent Ranking
 plugin>
```

지능형 검색 플러그인과의 상호 작용

인텔리전트 랭킹 플러그인을 설치 OpenSearch (자체 관리) 한 후에는 curl 명령어 또는 OpenSearch 클라이언트 라이브러리를 사용하여 플러그인과 상호 작용할 수 있습니다. 인텔리전트 랭킹 플러그인으로 액세스하기 OpenSearch 위한 기본 자격 증명은 사용자 이름 'admin'과 암호 'admin'입니다.

인텔리전트 랭킹 플러그인 설정을 색인에 적용하려면: OpenSearch

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)
```

순위를 다시 매기는 데 사용할 기본 텍스트 필드의 이름(예: 문서 본문 또는 문서 컨텐츠 필드)을 포함해야 합니다. 문서 제목이나 문서 요약과 같은 다른 텍스트 필드도 포함할 수 있습니다.

이제 어떤 쿼리든 실행할 수 있으며 Intelligent Ranking 플러그인을 사용하여 결과의 순위가 매겨집니다.

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)
```

OpenSearch 인덱스의 인텔리전트 랭킹 플러그인 설정을 제거하려면:

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)
```

특정 쿼리에서 Intelligent Ranking 플러그인을 테스트하거나 특정 본문 및 제목 필드에서 테스트하려면:

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)
```

OpenSearch 결과와 Amazon Kendra 결과 비교

순위가 매겨진 결과를 side-by-side OpenSearch (자체 관리) Amazon Kendra의 순위가 다시 매겨진 결과와 비교할 수 있습니다. OpenSearch 대시보드 버전 2.4.0 이상에서는 side-by-side 검색 결과를 제공

하므로 문서 순위를 매기는 방식과 플러그인이 검색 OpenSearch 쿼리의 문서 순위를 매기는 방식을 Amazon Kendra 비교할 수 있습니다.

OpenSearch 순위가 매겨진 결과와 Amazon Kendra 순위가 다시 매겨진 결과를 비교하기 전에 Intelligent Ranking 플러그인이 설치된 OpenSearch 서버에서 OpenSearch 대시보드를 지원하는지 확인하십시오. Docker와 빠른 시작 Bash 스크립트를 사용하여 이를 설정할 수 있습니다. 지능형 검색 플러그인 설정 섹션을 참조하십시오.

다음은 대시보드에서 결과를 OpenSearch 비교하고 Amazon Kendra 검색하는 방법을 요약한 것입니다. OpenSearch 자세한 내용은 설명서를 참조하십시오. OpenSearch

OpenSearch 대시보드의 검색 결과 비교

1. http://localhost:5601 을 열고 OpenSearch 대시보드에 로그인합니다. 기본 보안 인증은 사용자 이름 'admin'과 비밀번호 'admin'입니다.
2. 탐색 메뉴의 OpenSearch 플러그인에서 검색 관련성을 선택합니다.
3. 검색 창에 검색 텍스트를 입력합니다.
4. 쿼리 1의 색인을 선택하고 OpenSearch 쿼리 DSL에 쿼리를 입력합니다. `%SearchText%` 변수를 사용하여 검색 창에 입력한 검색 텍스트를 참조할 수 있습니다. 이 쿼리의 예는 OpenSearch 설명서를 참조하십시오. 이 쿼리에 대해 반환되는 OpenSearch 결과는 인텔리전트 랭킹 플러그인을 사용하지 않은 결과입니다.
5. 쿼리 2에 대해 동일한 인덱스를 선택하고 OpenSearch 쿼리 DSL에 동일한 쿼리를 입력합니다. 또한 `kendra_intelligent_ranking`에 확장을 포함하고 순위를 매길 필수 `body_field`를 지정하세요. 제목 필드를 지정할 수도 있지만 본문 필드는 필수입니다. 이 쿼리의 예는 OpenSearch 설명서를 참조하십시오. 이 쿼리에 대해 반환되는 결과는 인텔리전트 Amazon Kendra 랭킹 플러그인을 사용하여 순위를 다시 매긴 결과입니다. 플러그인은 최대 25개의 결과 순위를 매깁니다.
6. 결과를 반환하고 비교하려면 검색을 선택합니다.

검색 서비스 결과의 의미론적 순위 지정

Amazon Kendra 인텔리전트 랭킹은 Amazon Kendra의 시맨틱 검색 기능을 사용하여 검색 서비스 결과의 순위를 재지정합니다. 검색 쿼리의 컨텍스트와 검색 서비스 문서에서 사용 가능한 모든 정보를 고려하여 이를 수행합니다. Amazon Kendra 지능형 랭킹은 단순한 키워드 매칭을 개선할 수 있습니다.

CreateRescoreExecutionPlanAPI는 Rescore API를 프로비저닝하는 데 사용되는 Amazon Kendra 인텔리전트 랭킹 리소스를 생성합니다. `Rescore`API는 OpenSearch (자체 관리형) 와 같은 검색 서비스의 검색 결과 순위를 다시 매깁니다.

CreateRescoreExecutionPlan를 호출하면 검색 서비스 결과의 순위를 다시 매기는 데 필요한 용량 단위를 설정합니다. 단일 단위 기본값 외에 더 많은 용량 단위가 필요하지 않은 경우 기본값을 변경하지 마세요. 재평가 실행 계획에는 이름만 입력하세요. 최대 1000 추가 단위를 설정할 수 있습니다. 단일 용량 단위에 포함되는 항목에 대한 자세한 내용은 용량 조정을 참조하세요. Amazon Kendra 인텔리전트 랭킹을 프로비저닝하면 설정된 용량 단위를 기준으로 시간당 요금이 부과됩니다. 프리 티어 및 요금 정보를 참조하세요.

재평가 실행 계획 ID가 생성되어 CreateRescoreExecutionPlan 호출 시 응답에 반환됩니다. Rescore API는 재평가 실행 계획 ID를 사용하여 사용자가 설정한 용량으로 검색 서비스 결과의 순위를 다시 매깁니다. 검색 서비스의 구성 파일에 재평가 실행 계획 ID를 포함합니다. 예를 들어 OpenSearch (자체 관리형) 을 사용하는 경우 docker-compose.yml 또는 opensearch.yml 파일에 재평가 실행 계획 ID를 포함합니다. 지능적으로 순위 지정 (셀프 서비스) 결과를 참조하십시오. OpenSearch

Amazon 리소스 이름(ARN)도 CreateRescoreExecutionPlan 호출 시 응답에 생성됩니다. 이 ARN을 사용하여 AWS Identity and Access Management (IAM) 에서 권한 정책을 생성하여 특정 재점수 실행 계획에 따라 특정 ARN에 대한 사용자 액세스를 제한할 수 있습니다. 특정 재점수 실행 계획에 Rescore API 사용 권한을 부여하는 IAM 정책의 예는 자체 관리를 위한 Amazon Kendra 지능형 순위 지정을 참조하십시오. OpenSearch

다음은 용량 단위를 1로 설정하여 재평가 실행 계획을 생성하는 예제입니다.

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.");
  }
}
```

다음은 재평가 실행 계획을 업데이트하여 용량 단위를 2로 설정하는 예제입니다.

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.");
  }
}
```

다음은 Rescore API 사용의 예입니다.

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());
  }
}
```

에 대한 문서 기록 Amazon Kendra

- 최신 설명서 업데이트: 2024년 2월 27일

다음 표에는 각 릴리스의 Amazon Kendra 중요한 변경 사항이 설명되어 있습니다. 이 설명서의 업데이트에 대한 알림을 받으려면 RSS 피드를 구독할 수 있습니다.

변경 사항	설명	날짜
새 기능	Amazon Kendra 이제 GitHub 데이터 소스 커넥터의 업데이트된 버전을 지원합니다. 자세한 내용은 을 참조하십시오 GitHub.	2024년 2월 27일
새 기능	Amazon Kendra 이제 Amazon FSx 데이터 소스 커넥터의 업데이트된 버전을 지원합니다. 자세한 내용은 Amazon FSx (Windows) 및 Amazon FSx (NetAppONTAP) 를 참조하십시오.	2024년 2월 8일
새 기능	Amazon Kendra 이제 Slack 데이터 소스 커넥터의 업데이트된 버전을 지원합니다. 자세한 정보는 Slack을 참조하세요.	2024년 1월 11일
새 기능	Amazon Kendra 이제 검색 결과 축소 및 확장이 지원됩니다. 자세한 내용은 검색 결과 축소/확장을 참조하세요.	2023년 10월 19일
새 기능	Amazon Kendra 이제 Aurora (MySQL) 데이터 소스 커넥터를 지원합니다. 자세한 내용은	2023년 9월 28일

	Aurora (MySQL)을 참조하세요.	
새 기능	Amazon Kendra 이제 Aurora (PostgreSQL) 데이터 소스 커넥터를 지원합니다. 자세한 내용은 Aurora (PostgreSQL)을 참조하세요.	2023년 9월 28일
새 기능	Amazon Kendra 이제 Amazon RDS (MySQL) 데이터 소스 커넥터를 지원합니다. 자세한 내용은 Amazon RDS (MySQL)을 참조하세요.	2023년 9월 28일
새 기능	Amazon Kendra 이제 Amazon RDS (Microsoft SQL Server) 데이터 소스 커넥터를 지원합니다. 자세한 내용은 Amazon RDS (Microsoft SQL Server)를 참조하세요.	2023년 9월 28일
새 기능	Amazon Kendra 이제 Amazon RDS (Oracle) 데이터 소스 커넥터를 지원합니다. 자세한 내용은 Amazon RDS (Oracle)을 참조하세요.	2023년 9월 28일
새 기능	Amazon Kendra 이제 Amazon RDS (PostgreSQL) 데이터 소스 커넥터를 지원합니다. 자세한 내용은 Amazon RDS (PostgreSQL)을 참조하세요.	2023년 9월 28일
새 기능	Amazon Kendra 이제 IBM DB2 데이터 소스 커넥터를 지원합니다. 자세한 내용은 IBM DB2를 참조하세요.	2023년 9월 28일

새 기능	Amazon Kendra 이제 Microsoft SQL Server 데이터 소스 커넥터를 지원합니다. 자세한 내용은 Microsoft SQL Server를 참조하세요.	2023년 9월 28일
새 기능	Amazon Kendra 이제 MySQL 데이터 소스 커넥터를 지원합니다. 자세한 내용은 MySQL을 참조하세요.	2023년 9월 28일
새 기능	Amazon Kendra 이제 Oracle 데이터베이스 데이터 소스 커넥터를 지원합니다. 자세한 내용은 Oracle Database를 참조하세요.	2023년 9월 28일
새 기능	Amazon Kendra 이제 PostgreSQL 데이터 소스 커넥터를 지원합니다. 자세한 내용은 PostgreSQL을 참조하세요.	2023년 9월 28일
새 기능	Amazon Kendra 이제 Drupal용 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Drupal을 참조하세요.	2023년 9월 6일
새 기능	검색 증강 생성(RAG) 시스템용 Amazon Kendra 검색 API를 사용하여 의미상 관련 있는 구절을 검색합니다.	2023년 6월 22일
새 기능	Amazon Kendra 이제 Amazon Kendra Web Crawler 데이터 소스 커넥터의 업데이트된 버전을 지원합니다. 자세한 내용은 Amazon Kendra 웹 크롤러 v2.0을 참조하세요.	2023년 6월 21일

리전 확장	Amazon Kendra 이제 유럽 (런던) (eu-west-2) 에서 사용할 수 있습니다.	2023년 6월 5일
새 기능	Amazon Kendra 이제 Alfresco 데이터 소스 커넥터의 업데이트된 버전을 지원합니다. 자세한 내용은 Alfresco를 참조하세요.	2023년 5월 16일
새 기능	Amazon Kendra 이제 Adobe Experience Manager를 위한 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Adobe Experience Manager를 참조하세요.	2023년 5월 11일
새 기능	Amazon Kendra 이제 호출 시 문서 필드/속성을 구성할 수 있습니다. GetQuerySuggestions 이제 문서 필드의 내용을 기반으로 쿼리 제안을 할 수 있습니다. 자세한 내용은 쿼리 제안을 참조하세요.	2023년 5월 2일
새 기능	Amazon Kendra 이제 Gmail용 데이터 소스 커넥터를 제공합니다. 자세한 정보는 Gmail을 참조하세요.	2023년 4월 13일
새 기능	Amazon Kendra 이제 Microsoft OneDrive 데이터 소스 커넥터의 업데이트된 버전을 지원합니다. 자세한 내용은 Microsoft OneDrive v2.0을 참조하십시오.	2023년 4월 3일

새 기능	주요 결과를 사용하여 사용자가 특정 쿼리를 입력할 때 새 문서의 가시성을 향상하거나 특정 문서를 홍보하세요.	2023년 3월 30일
새 기능	Amazon Kendra 이제 Microsoft용 업데이트된 데이터 소스 커넥터를 지원합니다 SharePoint. 자세한 내용은 Microsoft를 참조하십시오 SharePoint.	2023년 3월 2일
새 기능	Amazon Kendra 이제 Confluence 데이터 원본 커넥터의 업데이트된 버전을 지원합니다. 자세한 내용은 Confluence를 참조하세요.	2023년 3월 1일
리전 확장	Amazon Kendra 이제 아시아 태평양 (도쿄) (ap-northeast-1)에서 사용할 수 있습니다.	2023년 2월 7일
새 기능	Amazon Kendra 이제 Microsoft Exchange를 위한 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Microsoft Exchange를 참조하세요.	2023년 1월 12일
새 기능	Amazon Kendra 이제 Microsoft Yammer를 위한 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Microsoft Yammer를 참조하세요.	2023년 1월 12일

새 기능	Amazon Kendra 이제 RTF, XML, XSLT, MS_EXCEL, CSV, JSON 및 MD 문서 유형의 인덱싱을 지원합니다. 자세한 내용은 문서 유형을 참조하세요.	2023년 1월 11일
새 기능	Amazon Kendra 이제 데이터 소스 커넥터의 업데이트된 버전을 지원합니다. Amazon S3 자세한 설명은 Amazon S3 섹션을 참조하세요.	2023년 1월 10일
새 기능	OpenSearch(자체 관리형) 검색 결과는 Amazon Kendra 인텔리전트 랭킹을 사용하여 의미론적으로 순위를 매길 수 있습니다.	2023년 1월 9일
새 기능	Amazon Kendra 이제 Microsoft Teams를 위한 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Microsoft Teams를 참조하세요.	2023년 1월 5일
새 기능	Amazon Kendra Google 드라이브용 데이터 소스 커넥터가 업데이트되었습니다. 자세한 내용은 Google Drive를 참조하세요.	2023년 1월 5일
새 기능	Amazon Kendra 의 데이터 소스 커넥터가 업데이트되었습니다 ServiceNow. 자세한 내용은을 참조하십시오 ServiceNow.	2022년 12월 21일

새 기능	Amazon Kendra Salesforce용 데이터 소스 커넥터가 업데이트되었습니다. 자세한 내용은 Salesforce를 참조하세요.	2022년 12월 21일
리전 확장	Amazon Kendra 이제 아시아 태평양 (뭄바이) (ap-south-1)에서 사용할 수 있습니다.	2022년 12월 14일
새 기능	Amazon Kendra의 표 검색 기능은 HTML 문서에 포함된 표에서 답변을 검색하고 추출할 수 있습니다.	2022년 11월 27일
새 기능	Amazon Kendra 선택한 언어 집합에 대한 시맨틱 검색을 지원합니다.	2022년 11월 27일
새 기능	Amazon Kendra 이제 Dropbox용 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Dropbox를 참조하세요.	2022년 9월 27일
새 기능	Amazon Kendra 이제 Zendesk용 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Zendesk를 참조하세요.	2022년 8월 17일
새로운 특성	이제 문서를 인덱싱한 후 문서 수준 액세스 제어를 다시 구성할 수 있습니다. 자세한 내용은 액세스 제어 구성을 참조하세요.	2022년 7월 14일
새 기능	Amazon Kendra 이제 Alfresco용 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Alfresco를 참조하세요.	2022년 6월 30일

새 기능	Amazon Kendra 이제 에 대한 데이터 소스 커넥터를 제공합니다. GitHub 자세한 내용은 을 참조하십시오 GitHub.	2022년 6월 2일
새 기능	Amazon Kendra 이제 Jira용 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Jira를 참조하세요.	2022년 5월 12일
새 기능	패싯 내의 중첩된 패싯을 검색 결과에 표시할 수 있습니다. 자세한 내용은 패싯을 참조하세요.	2022년 5월 5일
새 기능	Amazon Kendra 이제 Quip용 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Quip을 참조하세요.	2022년 4월 19일
새 기능	Amazon Kendra 이제 Box용 데이터 소스 커넥터를 제공합니다. 자세한 내용은 Box를 참조하세요.	2022년 4월 6일
새 기능	Amazon Kendra 이제 Slack용 데이터 소스 커넥터를 제공합니다. 자세한 정보는 Slack을 참조하세요.	2022년 3월 14일
새 기능	Amazon Kendra 이제 데이터 소스 커넥터를 제공합니다. Amazon FSx 자세한 설명은 Amazon FSx 섹션을 참조하세요.	2022년 2월 8일

AWS 관리형 정책 업데이트 - 새 정책	Amazon Kendra 새 AWS 관리형 정책이 추가되었습니다. 자세한 내용은 Amazon Kendra에 대한 AWS 관리형 정책을 참조하세요.	2022년 1월 3일
새 기능	Amazon Kendra 프런트 엔드 코드 없이 클릭 몇 번으로 검색 애플리케이션을 배포할 수 있습니다. 자세한 내용은 코드 없이 검색 애플리케이션 배포를 참조하세요.	2021년 12월 1일
새 기능	문서 수집 프로세스 중 문서 메타데이터 및 콘텐츠를 보강할 수 있습니다. 자세한 내용을 알아보려면 수집 프로세스 중 문서 메타데이터 사용자 지정을 참조하세요.	2021년 12월 1일
새 기능	Amazon Kendra 검색 애플리케이션에 대한 유용한 통찰력을 얻을 수 있는 검색 분석을 제공합니다. 자세한 내용은 검색 분석을 통한 인사이트 확보를 참조하세요.	2021년 12월 1일
리전 확장	Amazon Kendra 이제 AWS GovCloud (미국 서부) (us-gov-west-1) 에서 사용할 수 있습니다.	2021년 10월 13일

새 기능	Amazon Kendra 이제 여러 언어로 문서를 인덱싱하고 검색 결과를 언어별로 필터링할 수 있습니다. 영어 이외의 언어로 문서 추가 및 언어 선택을 참조하세요.	2021년 10월 7일
새 기능	Amazon Kendra 이제 Identity Center 디렉터리와 통합되어 사용자 컨텍스트 필터링을 위한 그룹 및 사용자의 액세스 수준을 가져올 수 있습니다. IAM Identity Center의 사용자 그룹 구성을 참조하세요.	2021년 10월 6일
새 자습서	Amazon Kendra 이제 메타데이터가 풍부한 검색 솔루션을 구축하는 방법을 안내하는 자습서를 제공합니다. 지능형 검색 솔루션 구축을 참조하세요.	2021년 8월 13일
새 기능	Amazon Kendra 이제 에 대한 데이터 소스 커넥터를 제공합니다. Amazon WorkDocs 자세한 설명은 Amazon WorkDocs 섹션을 참조하세요.	2021년 7월 20일
새 기능	Amazon Kendra 이제 웹 페이지를 크롤링하고 인덱싱할 수 있는 웹 크롤러를 제공합니다. 자세한 내용은 웹 크롤러를 참조하세요.	2021년 6월 16일
리전 확장	Amazon Kendra 이제 캐나다(중부) (ca-central-1) 에서 사용할 수 있습니다.	2021년 6월 16일

리전 확장	Amazon Kendra 이제 미국 동부 (오하이오) (us-east-2) 에서 사용할 수 있습니다.	2021년 6월 7일
새 기능	Amazon Kendra 이제 검색어와 관련된 인기 검색어를 사용자에게 제안하는 검색어 제안을 지원합니다. 자세한 내용은 인기 검색 쿼리 제안을 참조하세요.	2021년 5월 27일
AWS 관리형 정책 업데이트 - 새 정책	Amazon Kendra 새 AWS 관리형 정책이 추가되었습니다. 자세한 내용은 Amazon Kendra에 대한 AWS 관리형 정책을 참조하세요.	2021년 5월 27일
리전 확장	Amazon Kendra 이제 아시아 태평양 (싱가포르) (ap-southeast-1) 에서 사용할 수 있습니다.	2021년 5월 5일
새 기능	Amazon Kendra 이제 인덱스 수준에서 설정된 조정 구성을 재정의하여 쿼리의 검색 관련성을 조정할 수 있습니다. 자세한 내용은 검색 관련성 조정 및 응답 조정을 참조하세요.	2021년 4월 20일
새 기능	Amazon Kendra 이제 OAuth 2.0 인증을 지원하고 ServiceNow 쿼리를 사용하여 인덱싱할 문서를 선택할 수 있습니다. 자세한 내용은 을 참조하십시오. ServiceNow	2021년 4월 1일

새 기능	Amazon Kendra 이제 FAQ 문서에 대한 증분 학습을 지원합니다. 자세한 내용은 점진적 학습을 위한 피드백 제출을 참조하세요.	2021년 2월 17일
새 기능	Amazon Kendra 이제 인덱스 동의어를 지원합니다. 자세한 내용은 인덱스에 동의어 추가를 참조하세요.	2020년 12월 10일
새 기능	Amazon Kendra 이제 Google 워크스페이스 드라이브용 데이터베이스 커넥터를 제공합니다. 자세한 내용은 Google Workspace Drive 데이터 소스 사용을 참조하세요.	2020년 12월 8일
새 기능	Amazon Kendra 이제 쿼리 피드백을 더 쉽게 제공할 수 있는 JavaScript 라이브러리가 제공됩니다 Amazon Kendra. 자세한 내용은 피드백 제출을 참조하세요.	2020년 12월 8일
새 기능	Amazon Kendra 이제 토큰 기반 사용자 액세스 제어를 지원합니다. 자세한 내용은 인덱스의 문서에 대한 액세스 제어를 참조하세요.	2020년 11월 5일
새 기능	이제 Amazon Kendra Confluence 데이터 소스 커넥터가 Confluence 클라우드와 함께 작동합니다. 자세한 내용을 알아보려면 Confluence 데이터 소스 사용을 참조하세요.	2020년 11월 5일

리전 확장	Amazon Kendra 이제 아시아 태평양 (시드니) 에서 사용할 수 있습니다 (ap-southeast-2).	2020년 11월 2일
새 기능	Amazon Kendra 이제 Confluence 서버용 데이터 소스 커넥터를 제공합니다. 자세한 내용을 알아보려면 Confluence 데이터 소스 사용을 참조하세요.	2020년 10월 26일
새 기능	Amazon Kendra 이제 사용자 지정 커넥터에 대한 통계를 생성하는 데 사용할 수 있는 데이터 원본을 제공합니다. 자세한 내용은 사용자 지정 데이터 소스 사용을 참조하세요.	2020년 10월 21일
새 기능	Amazon Kendra 이제 자주 묻는 질문에 대한 사용자 지정 속성을 지원합니다. 자세한 내용은 질문 및 답변 추가를 참조하세요.	2020년 9월 17일
새 기능	Amazon Kendra 이제 쿼리 결과에 대한 신뢰도 점수를 반환합니다. 자세한 내용은 을 참조하십시오 QueryResultItem.	2020년 9월 15일
새 기능	AWS CloudFormation 이제 지원됩니다 Amazon Kendra. 자세한 내용은 Amazon Kendra 리소스 유형 참조 -를 참조하십시오 AWS CloudFormation.	2020년 9월 10일

새 기능	Amazon Kendra 에 대한 지원을 추가합니다 AWS PrivateLink. 자세한 내용은 Amazon Kendra 및 인터페이스 VPC 엔드포인트(AWS PrivateLink)를 참조하세요.	2020년 7월 7일
새 안내서	이 문서는 첫 번째 Amazon Kendra 개발자 안내서 릴리스입니다.	2020년 5월 11일

API 참조

이제 API 참조 설명서가 별도의 안내서로 제공됩니다.

AWS 용어집

최신 AWS 용어는 AWS 용어집 참조서의 AWS 용어집을 참조하세요.

기계 번역으로 제공되는 번역입니다. 제공된 번역과 원본 영어의 내용이 상충하는 경우에는 영어 버전이 우선합니다.