
Amazon Kendra

Guía para desarrolladores



Amazon Kendra: Guía para desarrolladores

Copyright © 2023 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Las marcas comerciales y la imagen comercial de Amazon no se pueden utilizar en relación con ningún producto o servicio que no sea de Amazon de ninguna manera que pueda causar confusión entre los clientes y que menoscalice o desacredite a Amazon. Todas las demás marcas comerciales que no sean propiedad de Amazon son propiedad de sus respectivos propietarios, que pueden o no estar afiliados, conectados o patrocinados por Amazon.

Table of Contents

.....	ix
¿Qué es Amazon Kendra?	1
Consulta de Amazon Kendra	1
Ventajas de Amazon Kendra	1
Amazon KendraEdiciones	2
Precios de Amazon Kendra	3
¿Es la primera vez que usa Amazon Kendra?	3
Cómo funciona Amazon Kendra	4
Índice	4
UsoAmazon Kendracampos de documentos reservados o comunes	5
Búsqueda de índices	6
Documentos de	6
Tipos o formatos de documentos	6
Atributos o campos del documento	8
Orígenes de datos	10
Consultas	11
Etiquetas	11
Etiquetado de recursos de	12
Restricciones de las etiquetas	12
Configuración de Amazon Kendra	13
Registrarse en AWS	13
Regiones y puntos de enlace	13
Configuración de AWS CLI	13
Configuración de los AWS SDK	14
IAMfunciones de acceso para Amazon Kendra	15
IAMfunciones para índices	15
IAMfunciones para la BatchPutDocument API	17
IAMfunciones para fuentes de datos	19
IAMfunciones para preguntas frecuentes (FAQ)	64
IAMfunciones para sugerencias de consultas	65
IAMfunciones para el mapeo principal de usuarios y grupos	66
Roles de IAM para AWS IAM Identity Center (successor to AWS Single Sign-On)	68
IAMroles para Amazon Kendra experiencias	69
IAMfunciones para el enriquecimiento personalizado de documentos	71
Implementación de Amazon Kendra	74
Información general	74
Requisitos previos	75
Configuración del ejemplo	75
Página principal de búsqueda	76
Componente de búsqueda	76
Componente de resultados	76
Componente de facetas	76
Componente de paginación	76
Despliegue de una aplicación de búsqueda sin código	77
Cómo funciona la búsqueda Experience Builder	77
Diseña y ajusta tu experiencia de búsqueda	77
Proporcionar acceso a tu página de búsqueda	78
Configuración de una experiencia de búsqueda	79
Capacidad de ajuste	83
Capacidad de visualización	83
Añadir y eliminar capacidad	84
Amazon KendraCapacidad de clasificación inteligente	84
Capacidad de sugerencias de consultas	84
Amazon Kendracapacidad de experiencia	85

Capacidad de experiencia de búsqueda	85
Ráfaga de consultas adaptable	85
Introducción	86
Requisitos previos	86
Registro en una Cuenta de AWS	86
Crear un usuario administrativo	87
Amazon KendraRecursos:AWS CLI, SDK, consola	87
Empezar con elAmazon Kendraconsola	91
Introducción (AWS CLI)	92
Primeros pasos (SDK para Python (Boto3))	93
Primeros pasos (SDK para Java)	96
Introducción a S3 (consola)	98
Introducción a MySQL (consola)	99
Introducción a una fuente de identidad de IAM Identity Center (consola)	101
Cambiar la fuente de identidad de IAM Identity Center	103
Creación de un índice	104
Añadir documentos directamente a un índice con carga por lotes	107
Añadir documentos con la BatchPutDocument API	107
Añadir documentos desde un bucket de S3	109
Añadir preguntas frecuentes (FAQs) a un índice	111
Archivo CSV básico	112
Archivo CSV personalizado	112
Archivo JSON	113
Uso del archivo de preguntas frecuentes	115
Archivos de preguntas frecuentes en idiomas distintos del inglés	116
Creación de campos de documentos personalizados	116
Añadir atributos o campos personalizados con la BatchPutDocument API	117
Agregar atributos o campos personalizados a una fuente Amazon S3 de datos	117
Controlar el acceso de los usuarios a los documentos con identificadores	117
Uso de OpenID	118
Uso de un token web JSON (JWT) con un secreto compartido	120
Uso de un token web JSON (JWT) con una clave pública	122
Uso de JSON	124
Creación de un conector de fuente de datos	127
Establecer un cronograma de actualizaciones	127
Configuración de un idioma	127
Conectores de fuentes de datos	128
Esquemas de plantillas de fuentes de datos	128
Adobe Experience Manager	283
Alfresco	289
Amazon RDS/Aurora	294
Amazon FSx	299
Amazon S3	303
Amazon KendraRastreador web	312
Amazon WorkDocs	325
Box	328
Confluence	333
Conector de fuente de datos personalizado	344
Dropbox	350
GitHub	355
Gmail	360
Google Drive	365
Jira	376
Microsoft Exchange	380
MicrosoftOneDrive	384
MicrosoftSharePoint	393
Equipos de Microsoft	414

Microsoft Yammer	420
Quip	424
Salesforce	427
ServiceNow	438
Slack	450
Zendesk	455
Asignación de campos de origen de datos	459
UsoAmazon Kendracampos de documentos reservados o comunes	5
Añadir documentos en idiomas distintos del inglés	463
ConfigurandoAmazon Kendrautilizar unAmazon VPC	465
Conexión a una base de datos en una VPC	466
Eliminar un índice, una fuente de datos o documentos cargados por lotes	468
Eliminar un índice	468
Eliminar una fuente de datos	469
Eliminar documentos subidos por lotes	470
Enriquecer sus documentos durante la ingestión	472
Cómo funciona Custom Document Enrichment	472
Operaciones básicas para cambiar los metadatos	473
Funciones lambda: extraer y cambiar metadatos o contenido	478
Contratos de datos para funciones Lambda	484
Formato del documento estructurado	485
Ejemplo de una función Lambda que se adhiere a los contratos de datos	486
Búsqueda en un índice	488
Consultar un índice	488
Requisitos previos	489
Búsqueda en un índice (consola)	489
Búsqueda en un índice (SDK)	489
Búsqueda en un índice (Postman)	491
Búsqueda con sintaxis de consulta avanzada	492
Búsqueda en idiomas	495
Recuperación de pasajes	498
Navegar por un índice	500
Incluye resultados de búsqueda	502
Búsqueda tabular de HTML	504
Sugerencias de consulta	507
Sugerencias de consulta mediante el historial de consultas	508
Sugerencias de consulta mediante campos de documentos	512
Bloquear determinadas consultas o el contenido de los campos del documento para que no aparezcan en las sugerencias	515
Corrector ortográfico de consultas	519
Uso del corrector ortográfico de consultas con límites predeterminados	520
Filtrado y búsqueda de facetas	520
Facetas	520
Uso de los atributos del documento para filtrar los resultados de la búsqueda	523
Filtrar los atributos de cada documento en los resultados de búsqueda	524
Filtrar según el contexto del usuario	525
Filtrar por token de usuario	525
Filtrar por ID de usuario y grupo	526
Filtrar por atributo de usuario	527
Filtrado de contexto de usuario para documentos añadidos directamente a un índice	528
Filtrado de contexto de usuario para preguntas frecuentes	528
Filtrado de contexto de usuario para fuentes de datos	528
Respuestas a consultas y tipos de respuestas	538
Respuestas a consultas	538
Tipos de respuesta	541
Ajustar y ordenar las respuestas	544
Ajustar las respuestas	544

Clasificación de respuestas	545
Ajustar la relevancia de la búsqueda	547
Ajuste de relevancia a nivel de índice	548
Ajuste de relevancia a nivel de consulta	548
Obtener información con el análisis de búsqueda	550
Métricas de búsqueda	550
Porcentaje de clics	551
Tasa de clics cero	551
Tasa cero de resultados de búsqueda	551
Tasa de respuesta instantánea	551
Consultas principales	551
Consultas principales con cero clics	552
Consultas principales con cero resultados de búsqueda	552
Documentos en los que más se ha hecho clic	552
Total de consultas	553
Documentos totales	553
Ejemplo de recuperación de datos métricos	553
Desde métricas hasta información procesable	554
Visualización y generación de informes de análisis de búsqueda	555
Gráfico de consultas totales	555
Gráfico de tasa de clics	555
Gráfico de tasa de clics cero	555
Gráfico de tasas de cero resultados de búsqueda	555
Gráfico de tasa de respuesta instantánea	556
Enviar comentarios para un aprendizaje incremental	557
Uso de la Amazon Kendra JavaScript biblioteca para enviar comentarios	558
Paso 1: Inserta una etiqueta de script en tu aplicación Amazon Kendra de búsqueda	558
Paso 2: Añade el token de valoración a los resultados de la búsqueda	560
Paso 3: Probar el script de comentarios	560
Uso de la Amazon Kendra API para enviar comentarios	561
Añadir sinónimos personalizados a un índice	563
Creación de un archivo de sinónimos	564
Añadir un diccionario de sinónimos a un índice	566
Actualización de un diccionario de sinónimos	569
Eliminar un diccionario de sinónimos	572
Aspectos destacados en los resultados de búsqueda	573
Tutorial: Creación de una solución de búsqueda inteligente	574
Requisitos previos	575
Paso 1: Añadir documentos	576
Descargar el conjunto de datos de muestra	576
Creación de un bucket de Amazon S3	577
Crear carpetas de datos y metadatos en su bucket de S3	579
Carga de los datos de entrada	581
Paso 2: Detectar entidades	583
Ejecución de un trabajo de análisis de entidades de Amazon Comprehend	583
Paso 3: Formatear los metadatos	589
Descargar y extraer la salida de Amazon Comprehend	589
Cargar la salida al bucket de S3	592
Conversion de la salida al formato de metadatos de Amazon Kendra	593
Limpiar su depósito de Amazon S3	596
Paso 4: Crear un índice e ingerir los metadatos	598
Creación de un índice de Amazon Kendra	598
Actualización del rol de IAM para el acceso a Amazon S3	604
Creación de campos de índice de búsqueda personalizados de Amazon Kendra	606
Añadir el bucket de Amazon S3 como fuente de datos para el índice	610
Sincronización del índice de Amazon Kendra	613
Paso 5: Consultar el índice	615

Consultar su índice de Amazon Kendra	615
Filtrar los resultados de búsqueda	619
Paso 6: Limpieza	622
Limpiar tus archivos	622
.....	623
Monitoreo y registro	624
Índices de monitorización	624
Supervisión de las llamadas a la API de Amazon Kendra con CloudTrail	627
Información sobre Amazon Kendra en CloudTrail	628
Ejemplo: entradas del archivo de registro de Amazon Kendra	628
Supervisión de las llamadas a la API de clasificación inteligente de Amazon Kendra con CloudTrail	629
Información sobre la clasificación inteligente de Amazon Kendra en CloudTrail	629
Ejemplo: entradas del archivo de registro de Amazon Kendra Intelligent Ranking	630
Supervisión de Amazon Kendra con CloudWatch	631
Visualización de las métricas de Amazon Kendra	631
Creación de una alarma	632
CloudWatchMétricas para trabajos de sincronización de índices	632
Métricas para las fuentes de datos de Amazon Kendra	633
Métricas para documentos indexados	635
Supervisión de Amazon Kendra con registros CloudWatch	636
Flujos de registro de fuentes de datos	636
Flujos de registro de documentos	637
Seguridad	639
Protección de los datos	639
Cifrado en reposo	640
Cifrado en tránsito	640
Administración de claves	640
Puntos de conexión de VPC (AWS PrivateLink)	641
Consideraciones sobre los endpoints de Amazon Kendra VPC	641
Creación de un endpoint de interfaz VPC para Amazon Kendra	641
Creación de una política de puntos de conexión de VPC para Amazon Kendra	642
Administración de identidades y accesos	642
Público	643
Autenticación con identidades	643
Administración de acceso mediante políticas	645
Cómo funciona Amazon Kendra con IAM	647
Ejemplos de políticas basadas en identidad	651
Políticas administradas por AWS	655
Solución de problemas	658
Prácticas recomendadas de seguridad	660
Aplicar el principio del mínimo privilegio	660
Permisos de control de acceso basados en roles (RBAC)	660
Registro y monitoreo en Amazon Kendra	660
Validación de conformidad	660
Resiliencia	661
Seguridad de infraestructuras	662
Configuración y análisis de vulnerabilidades	662
Cuotas	663
Regiones admitidas	663
Cuotas	663
Solución de problemas	668
Solución de problemas de fuentes de datos	668
Mis documentos no estaban indexados	668
Mi trabajo de sincronización ha fallado	668
Mi trabajo de sincronización está incompleto	669
Mi trabajo de sincronización se ha realizado correctamente, pero no hay documentos indexados ..	669
Tengo problemas de formato de archivo al sincronizar mi fuente de datos	669

¿Cuánto tiempo lleva sincronizar una fuente de datos?	670
¿Cuánto cuesta sincronizar una fuente de datos?	670
Voy a recibir unAmazon EC2error de autorización	670
No puedo usar los enlaces del índice de búsqueda para abrir miAmazon S3objetos	670
Voy a recibir unAccessDeniedAl utilizar un archivo de certificado SSLmensaje de error	670
Recibo un error de autorización al utilizar unSharePointfuente de datos	671
Mi índice no rastrea los documentos de mi fuente de datos de Confluence	671
Solución de problemas de resultados de búsqueda de documentos	671
Los resultados de mi búsqueda no son relevantes para mi consulta de búsqueda	671
¿Por qué solo veo 100 resultados?	672
¿Por qué faltan los documentos que espero ver?	672
¿Por qué veo documentos que tienen una política de ACL?	672
Solución de problemas generales	672
Amazon KendraClasificación inteligente	674
Clasificación inteligente para autogestionados OpenSearch	674
Cómo funciona el complemento de búsqueda inteligente	674
Configuración del complemento de búsqueda inteligente	675
Interactuar con el complemento de búsqueda inteligente	678
Comparar OpenSearch los resultados con Amazon Kendra los resultados	682
Clasificación semántica de los resultados de un servicio de búsqueda	683
Historial de documentos	690
Referencia de la API	699
Glosario de AWS	700

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la versión original de inglés, prevalecerá la versión en inglés.

¿Qué es Amazon Kendra?

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

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

Note

También puedes usar las funciones Amazon Kendra de búsqueda semántica para volver a clasificar los resultados de otro servicio de búsqueda. Consulte [Clasificación Amazon Kendra inteligente](#) para obtener más información.

Con Amazon Kendra él, puede crear una experiencia de búsqueda unificada conectando varios repositorios de datos a un índice e ingiriendo y rastreando documentos. Puede utilizar los metadatos del documento para crear una experiencia de búsqueda personalizada y rica en funciones para sus usuarios, ayudándoles a encontrar de manera eficiente las respuestas correctas a sus consultas.

[¿Qué es Amazon Kendra?](#)

Consulta de Amazon Kendra

Puede realizar Amazon Kendra los siguientes tipos de consultas:

Preguntas factoides: preguntas sencillas sobre quién, qué, cuándo o dónde, como ¿dónde está el centro de servicio más cercano a Seattle? Las preguntas factoides tienen respuestas basadas en hechos que se pueden devolver como una sola palabra o frase. La respuesta se obtiene de una sección de preguntas frecuentes o de sus documentos indexados.

Preguntas descriptivas: preguntas en las que la respuesta puede ser una oración, un pasaje o un documento completo. Por ejemplo, ¿cómo conecto mi Echo Plus a mi red? O bien, ¿cómo puedo obtener beneficios tributarios para las familias de bajos ingresos?

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

Ventajas de Amazon Kendra

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

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

Conectividad: Amazon Kendra puede conectarse a repositorios de datos o fuentes de datos de terceros, como Microsoft SharePoint. Puede indexar y buscar documentos fácilmente utilizando su fuente de datos.

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

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

Amazon KendraEdiciones

Amazon Kendra tiene dos versiones: Developer Edition y Enterprise Edition. La siguiente tabla describe sus características y las diferencias entre las dos.

Amazon KendraEdición para desarrolladores	Amazon KendraEdición empresarial
Amazon KendraLa Edición para desarrolladores ofrece todas las funciones de Amazon Kendra a un costo menor.	Amazon KendraEnterprise Edition ofrece todas las funciones de los contextos de producción Amazon Kendra y está diseñada para ellos.
Caso de uso ideal	Caso de uso ideal
<ul style="list-style-type: none">Explorar cómo Amazon Kendra indexa tus documentosProbando funcionesDesarrollo de aplicaciones que utilicen Amazon Kendra	<ul style="list-style-type: none">Indexación de toda la biblioteca de documentos empresarialesImplementación de la aplicación en un entorno de producción
Características	Características
<ul style="list-style-type: none">Un nivel gratuito con 750 horas de uso incluidasHasta 5 índices con hasta 5 fuentes de datos cada uno10.000 documentos o 3 GB de texto extraídoAproximadamente 4.000 consultas por día o 0.05 consultas por segundoSe ejecuta en 1 zona de disponibilidad (AZ); consulte Zonas de disponibilidad (centros de datos en AWS regiones)	<ul style="list-style-type: none">Hasta 5 índices con hasta 50 fuentes de datos cada uno100 000 documentos o 30 GB de texto extraídoAproximadamente 8.000 consultas por día o 0,1 consultas por segundoSe ejecuta en 3 zonas de disponibilidad (AZ); consulte Zonas de disponibilidad (centros de datos en AWS regiones)
Limitaciones	Note
<ul style="list-style-type: none">No apto para aplicaciones de producciónNo hay garantías de latencia ni disponibilidad	Puede aumentar esta cuota mediante la consola de cuotas de servicio .
Limitaciones	Limitaciones
	<ul style="list-style-type: none">Ninguno

Note

Para obtener una lista de regiones, puntos de conexión y cuotas de servicio compatibles Amazon Kendra, consulte [Amazon Kendra puntos de conexión y cuotas](#).

Precios de Amazon Kendra

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

Cuando finalice la prueba, se le cobrará por todos los Amazon Kendra índices aprovisionados, incluso si están vacíos y no se ejecuta ninguna consulta. Cuando finalice la versión de prueba, se cobrarán cargos adicionales por escanear y sincronizar documentos mediante las fuentes de Amazon Kendra datos.

Para obtener una lista completa de cargos y precios, consulte [Amazon Kendra precios](#).

¿Es la primera vez que usa Amazon Kendra?

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

1	2	3	4	5	6
Cómo funciona Amazon Kendra (p. 4)	Introducción (p. 86)	Creación de un índice (p. 104)	Añadir documentos directamente a un índice con carga por lotes (p. 107)	Creación de un conector de fuente de datos (p. 127)	Búsqueda en un índice (p. 488)
Presenta Amazon Kendra los componentes y describe cómo utilizarlos para crear una solución de búsqueda.	Explica cómo configurar tu cuenta y probar la API Amazon Kendra de búsqueda.	Explica cómo se utiliza Amazon Kendra para crear un índice de búsqueda y agregar fuentes de datos para sincronizar los documentos.	Explica cómo añadir documentos directamente a un Amazon Kendra índice.	Explica cómo añadir documentos del repositorio de datos a un Amazon Kendra índice.	Explica cómo usar la API Amazon Kendra de búsqueda para buscar en un índice.

Cómo funciona Amazon Kendra

Amazon Kendraproporciona funciones de búsqueda a su aplicación. Indexa sus documentos directamente o desde su repositorio de documentos de terceros y proporciona información relevante de forma inteligente a sus usuarios. Puedes usarAmazon Kendrapara crear un índice actualizable de documentos de diversos tipos. Para obtener una lista de los tipos de documentos admitidos porAmazon Kendraver[Tipos de documentos](#).

Amazon Kendrase integra con otros servicios. Por ejemplo, puedes alimentar[Amazon Lexbots de chat](#)conAmazon Kendrabusque para proporcionar respuestas útiles a las preguntas de los usuarios. Puede utilizar un[Amazon Simple Storage Servicecangilón](#)como fuente de datos paraAmazon Kendrapara conectarse a sus documentos e indexarlos. Además, puede configurar políticas de acceso o permisos a los recursos mediante[AWS Identity and Access Management](#).

Amazon Kendra tiene los siguientes componentes:

- Un[índice](#)que contiene sus documentos y permite buscarlos.
- UN[fuente de datos](#)que almacena tus documentos yAmazon Kendrase conecta a. Puede sincronizar automáticamente una fuente de datos con unAmazon Kendraindex para que su índice se mantenga actualizado con su repositorio de origen.
- UN[API de adición de documentos](#)que añade documentos directamente a un índice.

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

Temas

- [Índice \(p. 4\)](#)
- [Documentos de \(p. 6\)](#)
- [Orígenes de datos \(p. 10\)](#)
- [Consultas \(p. 11\)](#)
- [Etiquetas \(p. 11\)](#)

Índice

Un índice contiene el contenido de los documentos y está estructurado de manera que se puedan buscar en los documentos. La forma de añadir documentos al índice depende de cómo almacene los documentos.

- Si almacena sus documentos en algún tipo de repositorio, como unAmazon S3bucket o un MicrosoftSharePointsitio, usas un[conector de fuente de datos](#)para indexar tus documentos de tu repositorio.
- Si no almacena sus documentos en un repositorio, utilice el[BatchPutDocument](#)API para indexar directamente tus documentos.
- Para las preguntas y respuestas de las preguntas frecuentes, que deben almacenarse en unAmazon Kendra(Amazon S3) bucket, los subes desde el bucket

Puede crear índices conAmazon Kendraconsola, laAWS CLI, o unAWSSDK. Para obtener información sobre los tipos de documentos que se pueden indexar, consulte[Tipos de documentos](#).

UsoAmazon Kendracampos de documentos reservados o comunes

Con el[UpdateIndexAPI](#), puede crear campos reservados o comunes mediante[DocumentMetadataConfigurationUpdates](#) y especificando elAmazon Kendranombre de campo de índice reservado para asignarlo a su atributo/nombre de campo de documento equivalente. También puede crear campos personalizados. Si utiliza un conector de fuente de datos, la mayoría incluye asignaciones de campos que asignan los campos del documento de fuente de datos aAmazon Kendracampos de índice. Si utiliza la consola, actualiza los campos seleccionando la fuente de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de mapeos de campos para configurar la fuente de datos.

Puede configurar el[Searchobject](#) para establecer un campo como visualizable, facetable, buscable y ordenable. Puede configurar el[Relevanceobject](#) para establecer el orden de clasificación de un campo, la duración del aumento o el período de tiempo que se aplicará a los valores de aumento, frescura, valor de importancia y valores de importancia mapeados a valores de campo específicos. Si utiliza la consola, puede establecer la configuración de búsqueda de un campo seleccionando la opción de faceta en el menú de navegación. Para configurar el ajuste de relevancia, seleccione la opción de buscar en el índice en el menú de navegación, introduzca una consulta y utilice las opciones del panel lateral para ajustar la relevancia de la búsqueda. No puede cambiar el tipo de campo una vez creado el campo.

Amazon Kendratiene los siguientes campos de documento reservados o comunes que puede utilizar:

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

En el caso de los campos personalizados, estos campos se crean mediante[DocumentMetadataConfigurationUpdates](#) con el[UpdateIndexAPI](#), igual que cuando

se crea un campo reservado o común. Debe establecer el tipo de datos adecuado para el campo personalizado. Si utiliza la consola, actualiza los campos seleccionando la fuente de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de mapeos de campos para configurar la fuente de datos. Algunas fuentes de datos no admiten la adición de campos nuevos o personalizados. No puede cambiar el tipo de campo una vez creado el campo.

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

- Fecha
- Número
- Cadena
- Lista de cadenas

Si ha añadido documentos al índice utilizando [BatchPutDocument API](#), `Attributes` enumera los campos/atributos de sus documentos y puede crear campos mediante el `DocumentAttribute` objeto.

Para documentos indexados desde un Amazon S3 fuente de datos, los campos se crean mediante un [Archivo de metadatos JSON](#) que incluye la información de los campos.

Si utiliza una base de datos compatible como fuente de datos, puede configurar los campos mediante la [opción de mapeos de campos](#).

Búsqueda de índices

Después de crear un índice, puede empezar a buscar en los documentos. Para obtener más información, consulte [Búsqueda de índices](#).

Documentos de

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

Temas

- [Tipos o formatos de documentos \(p. 6\)](#)
- [Atributos o campos del documento \(p. 8\)](#)

Tipos o formatos de documentos

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

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

Los documentos se pueden organizar en filas y columnas. Por ejemplo, cada documento es una fila y cada campo o atributo del documento, como el título y el contenido del cuerpo, es una columna. Por ejemplo, si utiliza una base de datos como fuente de datos, los datos deben estar estructurados u organizados en filas y columnas.

Puede añadir documentos a su índice de las siguientes maneras:

- API de [BatchPutDocument](#)
- [Conector de fuente de datos](#)

Si desea añadir un archivo de preguntas frecuentes, utilice el[CreateFaq](#)API para añadir el archivo almacenado en unAmazon S3balde. Puede elegir entre un formato CSV básico, un formato CSV que incluye campos/atributos personalizados en un encabezado y un formato JSON que incluye campos personalizados. El formato predeterminado es CSV básico.

A continuación se proporciona información sobre cada formato de documento admitido y sobre cómoAmazon Kendra trata cada formato al indexar documentos.

Formato de documento	Tratados como	Cómo se trata el documento	Estructura original
Formato de documento portátil (PDF)	HTML	Se convierte a HTML y, a continuación, se extrae el contenido.	No estructurado
HyperTextLenguaje de marcado (HTML)	HTML	Las etiquetas HTML se filtran para extraer el contenido. El contenido debe estar entre los principalesHTMLetiquetas de inicio y cierre (<HTML>content</HTML>).	Semiestructurado
Lenguaje de marcado extensible (XML)	XML	Las etiquetas XML se filtran para extraer el contenido.	Semiestructurado
Transformación extensible del lenguaje de hojas de estilo (XSLT)	XSLT	Las etiquetas se filtran para extraer el contenido.	Semiestructurado
MarkDown(MARYLAND)	Texto no cifrado	El contenido se extrae conMarkDownsintaxis incluida.	Semiestructurado
Valores separados por comas (CSV)	CSV	Contenido extraído de cada celda, con un único archivo tratado como un único resultado de documento.	Estructurado para archivos de preguntas frecuentes, de lo contrario semiestructurado
Microsoft Excel (XLSX)	XLSX	Contenido extraído de cada celda, con un único archivo tratado como un único resultado de documento.	Semiestructurado
JavaScriptNotación de objetos (JSON)	Texto no cifrado	El contenido se extrae con la sintaxis JSON incluida.	Semiestructurado

Formato de documento	Tratados como	Cómo se trata el documento	Estructura original
Formato de texto enriquecido (RTF)	RTF	La sintaxis RTF se filtra para extraer el contenido.	Semiestructurado
Microsoft PowerPoint(PPT)	PPT	Solo se extrae el contenido del textoPowerPointdiapositivas para buscar. Las imágenes y otros contenidos no se extraen.	No estructurado
Microsoft Word (DOCX)	DOCX	Solo se extrae el contenido de texto de las páginas de Word para su búsqueda. Las imágenes y otros contenidos no se extraen.	No estructurado
Texto plano (TXT)	TXT	Se extrae todo el texto del documento de texto.	No estructurado

Atributos o campos del documento

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

Para poder utilizar un campo de documento en una consulta, debe estar asignado a un campo de índice. Por ejemplo, el campo de título se puede asignar al campo_document_title. Para obtener más información, consulte [Campos de mapeo](#). Para añadir un campo nuevo, debe crear un campo de índice al que asignar el campo. Los campos de índice se crean mediante la consola o mediante la [UpdateIndex API](#).

Puede utilizar los campos del documento para filtrar las respuestas y crear resultados de búsqueda por facetas. Por ejemplo, puede filtrar una respuesta para que solo devuelva una versión específica de un documento, o puede filtrar las búsquedas para que solo devuelvan el tipo 1099 de documentos fiscales que coincidan con el término de búsqueda. Para obtener más información, consulte [Filtrado y búsqueda de facetas](#).

También puede utilizar los campos del documento para ajustar manualmente la respuesta a la consulta. Por ejemplo, puede optar por aumentar la importancia del campo de título para aumentar el peso que Amazon Kendra asigna al campo al determinar qué documentos se devolverán en la respuesta. Para obtener más información, consulte [Ajustar la relevancia de la búsqueda](#).

Si va a añadir un documento directamente a un índice, especifique los campos en [Documento](#) parámetro de entrada al [BatchPutDocument API](#). Los valores de los campos personalizados se especifican en un [DocumentAttribute](#) matriz de objetos. Si utiliza una fuente de datos, el método que utilice para agregar los campos del documento depende de la fuente de datos. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).

Uso de Amazon Kendra campos de documentos reservados o comunes

Con el [UpdateIndexAPI](#), puede crear campos reservados o comunes mediante `DocumentMetadataConfigurationUpdates` especificando el Amazon Kendra nombre de campo de índice reservado para asignarlo a su atributo/nombre de campo de documento equivalente. También puede crear campos personalizados. Si utiliza un conector de fuente de datos, la mayoría incluye asignaciones de campos que asignan los campos del documento de fuente de datos a Amazon Kendra campos de índice. Si utiliza la consola, actualiza los campos seleccionando la fuente de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de mapeos de campos para configurar la fuente de datos.

Puede configurar el `SearchObject` para establecer un campo como visualizable, facetable, buscable y ordenable. Puede configurar el `RelevanceObject` para establecer el orden de clasificación de un campo, la duración del aumento o el período de tiempo que se aplicará a los valores de aumento, frescura, valor de importancia y valores de importancia mapeados a valores de campo específicos. Si utiliza la consola, puede establecer la configuración de búsqueda de un campo seleccionando la opción de faceta en el menú de navegación. Para configurar el ajuste de relevancia, seleccione la opción de buscar en el índice en el menú de navegación, introduzca una consulta y utilice las opciones del panel lateral para ajustar la relevancia de la búsqueda. No puede cambiar el tipo de campo una vez creado el campo.

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

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

En el caso de los campos personalizados, estos campos se crean mediante `DocumentMetadataConfigurationUpdates` con el `UpdateIndexAPI`, igual que cuando se crea un campo reservado o común. Debe establecer el tipo de datos adecuado para el campo

personalizado. Si utiliza la consola, actualiza los campos seleccionando la fuente de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de mapeos de campos para configurar la fuente de datos. Algunas fuentes de datos no admiten la adición de campos nuevos o personalizados. No puede cambiar el tipo de campo una vez creado el campo.

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

- Fecha
- Número
- Cadena
- Lista de cadenas

Si ha añadido documentos al índice utilizando [BatchPutDocument API](#), `Attributes` enumera los campos/atributos de sus documentos y puede crear campos mediante el `DocumentAttribute` objeto.

Para documentos indexados desde un Amazon S3 fuente de datos, los campos se crean mediante un [Archivo de metadatos JSON](#) que incluye la información de los campos.

Si utiliza una base de datos compatible como fuente de datos, puede configurar los campos mediante la [opción de mapeos de campos](#).

Orígenes de datos

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

Las fuentes de datos compatibles son:

- [Adobe Experience Manager](#)
- [Al aire libre](#)
- [Amazon FSx](#)
- Amazon RDSPara MySQL, Amazon RDSPara PostgreSQL, Amazon Aurora MySQL, Amazon Aurora PostgreSQL [bases de datos](#)
- [Amazon S3cubos](#)
- [Amazon KendraRastreador web](#)
- [Amazon WorkDocs](#)
- [Box](#)
- [Confluencia](#)
- [Fuentes de datos personalizadas](#)
- [Dropbox](#)
- [GitHub](#)
- [Gmail](#)
- [Google Workspace Drives](#)
- [Jira](#)
- [Microsoft Exchange](#)
- [MicrosoftOneDrive](#)

- [Microsoft SharePoint](#)
- [Equipo de Microsoft](#)
- [Microsoft Yammer](#)
- [Sofismo](#)
- [Fuerza de ventas](#)
- [ServiceNow](#)
- [Slack](#)
- [Zendesk](#)

Para obtener una lista de los tipos o formatos de documentos admitidos por Amazon Kendra ver [Tipos de documentos](#). Primero debe crear un índice antes de crear un conector de fuente de datos para indexar los documentos desde la fuente de datos.

Note

Para crear un índice de documentos, no es necesario utilizar una fuente de datos. Puede añadir documentos directamente a un índice mediante la carga por lotes. Para obtener más información, consulte [Añadir documentos directamente a un índice](#).

Para ver un tutorial sobre el uso de Amazon Kendra consola, la AWS CLI o SDK, consulte [Primeros pasos](#).

Consultas

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

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

Una consulta también puede contener criterios sobre cómo filtrar la respuesta para que Amazon Kendra devuelva únicamente los documentos que cumplen los criterios de filtrado. Por ejemplo, si ha creado un campo de índice denominado departamento, puede filtrar la respuesta para que solo los documentos con el campo del departamento establecido en jurídico se devuelvan. Para obtener más información, consulte [Filtrar búsqueda](#).

Puede influir en los resultados de una consulta ajustando la relevancia de los campos individuales del índice. El ajuste cambia la importancia de un campo en los resultados. Por ejemplo, si planteas la importancia de los documentos con la categoría nuevo, es más probable que los documentos de esta categoría se incluyan en la respuesta. Para obtener más información, consulte [Ajustar la relevancia de la búsqueda](#).

Para obtener más información sobre el uso de consultas, consulte [Búsqueda en un índice](#).

Etiquetas

Administre sus índices, fuentes de datos y preguntas frecuentes mediante la asignación de etiquetas o rótulos. Puede utilizar etiquetas para clasificar sus Amazon Kendra recursos de diversas formas. Por ejemplo, por propósito, propietario, aplicación o cualquier combinación. Cada etiqueta consta de una clave y un valor, ambos definidos por el usuario.

Las etiquetas le ayudan a:

- Identificar y organizar sus recursos de AWS. Muchos AWS los servicios admiten el etiquetado, por lo que puede asignar la misma etiqueta a los recursos de diferentes servicios para indicar que los recursos están relacionados. Por ejemplo, puede etiquetar un índice y el Amazon Lexbot que usa el índice con la misma etiqueta.
- Asigne costos. Las etiquetas se activan en AWS Billing and Cost Management al picadero. AWS usa etiquetas para clasificar sus costos y entregarle un informe mensual de asignación de costos. Para obtener más información, consulte [Asignación de costos y etiquetado](#) en Acerca de AWS Administración de facturación y costos.
- Controle el acceso a los recursos de . Puede utilizar etiquetas en AWS Identity and Access Management (IAM) políticas que controlan el acceso a Amazon Kendra recursos. Puede adjuntar estas políticas a un IAM rol o usuario para activar el control de acceso basado en etiquetas. Para obtener más información, consulte [Autorización basada en etiquetas](#).

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

Etiquetado de recursos de

Si está utilizando el Amazon Kendra consola, puede etiquetar los recursos al crearlos o agregarlos más adelante. También puede utilizar la consola para actualizar o eliminar etiquetas.

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

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

Restricciones de las etiquetas

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

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

Configuración de Amazon Kendra

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

Esta guía incluye ejemplos AWS CLI de Java y Python.

Temas

- [Registrarse en AWS \(p. 13\)](#)
- [Regiones y puntos de enlace \(p. 13\)](#)
- [Configuración de AWS CLI \(p. 13\)](#)
- [Configuración de los AWS SDK \(p. 14\)](#)

Registrarse en AWS

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

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

Para inscribirse en AWS

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

Regiones y puntos de enlace

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

Configuración de AWS CLI

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

1. Para instalarloAWS CLI, siga las instrucciones de [Instalación de la interfaz de línea de AWS comandos](#) de la Guía del usuario de la interfaz de línea de AWS comandos.

2. Para configurar AWS CLI y configurar un perfil al que llamar AWS CLI, siga las instrucciones que aparecen en la Guía [del AWS CLI](#) usuario de la interfaz de línea de AWS comandos.
3. Para confirmar que el perfil de la AWS CLI está configurado correctamente, ejecute el comando siguiente:

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

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

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

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

```
aws kendra help
```

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

Configuración de los AWS SDK

Descargue e instale los SDK de AWS que desee utilizar. Esta guía proporciona ejemplos de Python. Para obtener más información sobre otros SDK de AWS, consulte [Herramientas para Amazon Web Services](#).

El paquete del SDK de Python se denomina Boto3.

Antes de ejecutar los siguientes comandos de Python, primero debe descargar e instalar [Python 3.6 o una versión posterior](#) para su sistema operativo. La compatibilidad con Python 3.5 y versiones anteriores está en desuso. Si no incluye pip en su directorio de scripts de Python, puede descargar [get-pip.py](#) y almacenarlo en su directorio de scripts. También puede configurar su directorio de Python como una [ruta o variable de entorno](#) mediante un programa de terminal.

```
# Install the latest Boto3 release via pip  
pip install boto3  
  
# You can install a specific version of Boto3 for compatibility reasons  
# Install Boto3 version 1.0 specifically  
pip install boto3==1.0.0  
  
# Make sure Boto3 is no older than version 1.15.0  
pip install boto3>=1.15.0  
  
# Avoid versions of Boto3 newer than version 1.15.3  
pip install boto3<=1.15.3
```

Para usar Boto3, debe configurar las credenciales de autenticación para su AWS cuenta mediante la consola de [IAM](#).

IAMfunciones de acceso para Amazon Kendra

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

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

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

Temas

- [IAMfunciones para índices \(p. 15\)](#)
- [IAMfunciones para la BatchPutDocument API \(p. 17\)](#)
- [IAMfunciones para fuentes de datos \(p. 19\)](#)
- [IAMfunciones para preguntas frecuentes \(FAQ\) \(p. 64\)](#)
- [IAMfunciones para sugerencias de consultas \(p. 65\)](#)
- [IAMfunciones para el mapeo principal de usuarios y grupos \(p. 66\)](#)
- [Roles de IAM para AWS IAM Identity Center \(successor to AWS Single Sign-On\) \(p. 68\)](#)
- [IAMroles para Amazon Kendra experiencias \(p. 69\)](#)
- [IAMfunciones para el enriquecimiento personalizado de documentos \(p. 71\)](#)

IAMfunciones para índices

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

IAMfunciones para índices

Una política de roles que permite acceder Amazon Kendra a un CloudWatch registro.

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Effect": "Allow",  
            "Action": "cloudwatch:PutMetricData",  
            "Resource": "*",  
            "Condition": {  
                "StringEquals": {  
                    "cloudwatch:namespace": "AWS/Kendra"  
                }  
            }  
        }  
    ]  
}
```

```
},
{
    "Effect": "Allow",
    "Action": "logs:DescribeLogGroups",
    "Resource": "*"
},
{
    "Effect": "Allow",
    "Action": "logs>CreateLogGroup",
    "Resource": "arn:aws:logs:your-region:your-account-id:log-group:/aws/kendra/*"
},
{
    "Effect": "Allow",
    "Action": [
        "logs:DescribeLogStreams",
        "logs>CreateLogStream",
        "logs:PutLogEvents"
    ],
    "Resource": "arn:aws:logs:your-region:your-account-id:log-group:/aws/kendra/*
*:log-stream:)"
}
]
```

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

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

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

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

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

IAMfunciones para la BatchPutDocument API

Warning

Amazon Kendra utiliza una política de bucket que otorgue permisos a un Amazon Kendra director para interactuar con un bucket de S3. En su lugar, usa IAM roles. Asegúrese de que Amazon Kendra no esté incluido como miembro de confianza en su política de bucket para evitar cualquier problema de seguridad de los datos al conceder permisos accidentalmente a directores arbitrarios. Sin embargo, puedes añadir una política de bucket para usar un Amazon S3 bucket en diferentes cuentas. Para obtener más información, consulte [Políticas para usar Amazon S3 en todas las cuentas](#). Para obtener información sobre las IAM funciones de las fuentes de datos de S3, consulte las [IAMfunciones](#).

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

IAMfunciones para la BatchPutDocument API

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

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

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

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

Se recomienda incluir `aws:sourceAccount` y `aws:sourceArn` en la política de confianza. Esto limita los permisos y comprueba de forma segura si `aws:sourceAccount` `aws:sourceArn` son los mismos que los que se proporcionan en la política de IAM roles para la `sts:AssumeRole` acción. Esto evita que entidades no autorizadas accedan a sus IAM funciones y a sus permisos. Para obtener más información, consulte la AWS Identity and Access Management guía sobre el [problema del diputado confuso](#).

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

```
    ]  
}
```

Una política de roles opcional que permite Amazon Kendra usar una clave maestra de AWS KMS cliente (CMK) para descifrar los documentos de un Amazon S3 bucket.

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

IAMfunciones para fuentes de datos

Cuando utilice la [CreateDataSource](#) API, debe asignar Amazon Kendra un IAM rol que tenga permiso para acceder a los recursos de la base de datos. Los permisos específicos necesarios dependen de la fuente de datos.

IAMfunciones para las fuentes de datos de Adobe Experience Manager

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

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

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

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

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

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

IAMfunciones para las fuentes de datos de Alfresco

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

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

```
{
```

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

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

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

IAMfunciones para fuentes Amazon S3 de datos

Warning

Amazon Kendra utiliza una política de bucket que otorgue permisos a un Amazon Kendra director para interactuar con un bucket de S3. En su lugar, usa IAM roles. Asegúrese de que Amazon Kendra no esté incluido como miembro de confianza en su política de bucket para evitar cualquier problema de seguridad de los datos al conceder permisos accidentalmente a directores arbitrarios. Sin embargo, puedes añadir una política de bucket para usar un Amazon S3 bucket en diferentes cuentas. Para obtener más información, consulte [Políticas para usar Amazon S3 en todas las cuentas \(p. 27\)](#) (desplácese hacia abajo).

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

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

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

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

Una política de roles opcional que permite Amazon Kendra usar una clave maestra de AWS KMS cliente (CMK) para descifrar los documentos de un Amazon S3 bucket.

```
{
```

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

Una política de roles opcional que permite acceder Amazon Kendra a un Amazon S3 bucket, mientras se usa un Amazon VPC AWS KMS permiso y sin activarlo AWS KMS ni compartirlo.

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

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

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::{bucket-name}/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3>ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::{bucket-name}"
            ],
            "Effect": "Allow"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:{your-region}:{your-account-id}:key/{key-id}"
            ],
            "Condition": {
                "StringLike": {
                    "kms:ViaService": [
                        "s3.{your-region}.amazonaws.com"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2>CreateNetworkInterface"
            ],
            "Resource": [
                "arn:aws:ec2:{your-region}:{your-account-id}:subnet/[[subnet-ids]]",
                "arn:aws:ec2:{your-region}:{your-account-id}:security-group/[[security-group]]"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2>CreateNetworkInterface"
            ],
            "Resource": "arn:aws:ec2:{your-region}:{your-account-id}:network-interface/*",
            "Condition": {
                "StringLike": {
                    "aws:RequestTag/AWS_KENDRA": "kendra_{your-account-id}_{index-id}_{data-source-id}_*"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "kendra:ListIndex"
            ],
            "Resource": [
                "arn:aws:kendra:{your-region}:{your-account-id}:index/*"
            ],
            "Condition": {
                "StringLike": {
                    "aws:RequestTag/AWS_KENDRA": "kendra_{your-account-id}_{index-id}_{data-source-id}_*"
                }
            }
        }
    ]
}
```

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

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

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

Políticas para usar Amazon S3 en todas las cuentas

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

Una política de roles para usar tu Amazon S3 bucket como fuente de datos cuando el bucket se encuentra en una cuenta diferente a la de tu Amazon Kendra índice.

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Action": [  
                "s3:GetObject"  
            ],  
            "Resource": [  
                "arn:aws:s3:::$bucket-in-other-account/*"  
            ],  
            "Effect": "Allow"  
        },  
        {  
            "Action": [  
                "s3>ListBucket"  
            ],  
            "Resource": [  
                "arn:aws:s3:::$bucket-in-other-account/*"  
            ],  
            "Effect": "Allow"  
        },  
        {  
            "Effect": "Allow",  
            "Action": [  
                "kendra:BatchPutDocument",  
                "kendra:BatchDeleteDocument"  
            ],  
            "Resource": [  
                "arn:aws:kendra:$your-region:$your-account-id:index/$index-id"  
            ]  
        },  
        {  
            "Effect": "Allow",  
            "Action": [  
                "s3:GetObject",  
                "s3:PutObject",  
                "s3:PutObjectAcl"  
            ],  
            "Resource": "arn:aws:s3:::$bucket-in-other-account/*"  
        }  
    ]  
}
```

```
        ]
    }
```

Una política de bucket para permitir que la función Amazon S3 de fuente de datos acceda al Amazon S3 bucket en todas las cuentas.

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

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

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

IAMfunciones para fuentes de datos de bases de datos

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

- Permisos para acceder al AWS Secrets Manager secreto que contiene el nombre de usuario y la contraseña del sitio de base de datos. Para obtener más información sobre el contenido del secreto, consulte [Fuentes de datos de bases de datos](#).
- Permisos para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el secreto de nombre de usuario y contraseña almacenado por Secrets Manager

- Permisos para utilizar las BatchDeleteDocument operaciones BatchPutDocument y para actualizar el índice.
- Permisos para acceder al Amazon S3 bucket que contiene el certificado SSL utilizado para comunicarse con el sitio de la base de datos.

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

Hay dos políticas opcionales que puede utilizar con una fuente de datos de base de datos.

Si ha cifrado el Amazon S3 bucket que contiene el certificado SSL utilizado para comunicarse con la base de datos, proporcione una política para dar Amazon Kendra acceso a la clave.

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

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

Si utilizas una VPC, proporciona una política que dé Amazon Kendra acceso a los recursos necesarios. Consulte las [IAMfunciones para las fuentes de datos y VPC](#) para ver la política requerida.

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

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

IAMfunciones para fuentes Amazon FSx de datos

Cuando lo usaAmazon FSx, proporciona un rol con las siguientes políticas.

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

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

```
        "arn:aws:kms:{your-region}:{{your-account-id}}:key/{{key-id}}"
    ],
    "Condition": {
        "StringLike": {
            "kms:ViaService": [
                "secretsmanager.{your-region}.amazonaws.com"
            ]
        }
    }
},
{
    "Effect": "Allow",
    "Action": [
        "ec2:CreateNetworkInterface",
        "ec2:DeleteNetworkInterface"
    ],
    "Resource": [
        "arn:aws:ec2:{your-region}:{{your-account-id}}:network-interface/*",
        "arn:aws:ec2:{your-region}:{{your-account-id}}:subnet/[[subnet-ids]]"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "ec2:DescribeSubnets",
        "ec2:DescribeNetworkInterfaces"
    ],
    "Resource": "*"
},
{
    "Effect": "Allow",
    "Action": [
        "ec2:CreateNetworkInterfacePermission"
    ],
    "Resource": "arn:aws:ec2:{your-region}:{{your-account-id}}:network-interface/*",
    "Condition": {
        "StringEquals": {
            "ec2:AuthorizedService": "kendra.*.amazonaws.com"
        },
        "ArnEquals": {
            "ec2:Subnet": [
                "arn:aws:ec2:{your-region}:{{your-account-id}}:subnet/[[subnet-ids]]"
            ]
        }
    }
},
{
    "Sid": "AllowsKendraToGetDomainNameOfActiveDirectory",
    "Effect": "Allow",
    "Action": "ds:DescribeDirectories",
    "Resource": "*"
},
{
    "Sid": "AllowsKendraToCallRequiredFsxAPIs",
    "Effect": "Allow",
    "Action": [
        "fsx:DescribeFileSystems"
    ],
    "Resource": "*"
},
{
    "Sid": "iamPassRole",
    "Effect": "Allow",
    "Action": "iam:PassRole",
    "Resource": "*",
}
```

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

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

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

IAMfunciones para las fuentes de datos de Amazon Kendra Web Crawler

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

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

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

```

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

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

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

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

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

IAMfunciones para fuentes Amazon WorkDocs de datos

Cuando lo usasAmazon WorkDocs, proporcionas un rol con las siguientes políticas

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

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Sid": "AllowsKendraToGetDomainNameOfActiveDirectory",  
            "Effect": "Allow",  
            "Action": "ds:DescribeDirectories",  
            "Resource": "*"  
        },  
        {  
            "Sid": "AllowsKendraToCallRequiredWorkDocsAPIs",  
            "Effect": "Allow",  
            "Action": [  
                "workdocs:GetDocumentPath",  
                "workdocs:GetGroup",  
                "workdocs:GetDocument",  
                "workdocs:DownloadDocumentVersions",  
                "workdocs:DescribeUsers",  
                "workdocs:DescribeFolderContents",  
                "workdocs:DescribeActivities",  
                "workdocs:DescribeComments",  
                "workdocs:GetFolder",  
                "workdocs:DescribeResourcePermissions",  
                "workdocs:GetFolderPath",  
                "workdocs:DescribeInstances"  
            ],  
            "Resource": "*"  
        },  
        {  
            "Sid": "iamPassRole",  
            "Effect": "Allow",  
            "Action": "iam:PassRole",  
            "Resource": "*",  
            "Condition": {  
                "StringEquals": {  
                    "iam:PassedToService": [  
                        "kendra.*.amazonaws.com"  
                    ]  
                }  
            }  
        },  
        {  
            "Sid": "AllowsKendraToCallBatchPutDeleteAPIs",  
            "Effect": "Allow",  
            "Action": [  
                "kendra:BatchPutDocument",  
                "kendra:BatchDeleteDocument"  
            ],  
            "Resource": [  
                "arn:aws:kendra:your-region:account-id:index/$index-id"  
            ]  
        }  
    ]  
}
```

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

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

IAMfunciones para las fuentes de datos de Box

Cuando usas Box, proporcionas un rol con las siguientes políticas.

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

```
{  
    "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"  
            ],  
            "Condition": {  
                "StringLike": {  
                    "kms:ViaService": [  
                        "secretsmanager.{your-region}.amazonaws.com"  
                    ]  
                }  
            }  
        }  
    ]  
}
```

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

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

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

IAMfunciones para las fuentes de datos de Confluence

IAMfunciones de Confluence Connector v1.0

Cuando utilizas un servidor de Confluence como fuente de datos, proporcionas un rol con las siguientes políticas:

- Permiso para acceder al AWS Secrets Manager secreto que contiene las credenciales necesarias para conectarse a Confluence. Para obtener más información sobre el contenido del secreto, consulta las [fuentes de datos de Confluence](#).
- Permite usar la clave maestra del AWS KMS cliente (CMK) para descifrar el secreto de nombre de usuario y contraseña almacenado por Secrets Manager
- Permite utilizar las BatchDeleteDocument operaciones BatchPutDocument y para actualizar el índice.

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

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

Si utilizas una VPC, proporciona una política que dé Amazon Kendra acceso a los recursos necesarios. Consulte las [IAMfunciones para las fuentes de datos y VPC](#) para ver la política requerida.

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

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

IAMfunciones de Confluence Connector v2.0

Para una fuente de datos de Confluence Connector v2.0, proporcionas un rol con las siguientes políticas.

- Permisos para acceder al AWS Secrets Manager secreto que contiene las credenciales de autenticación de Confluence. Para obtener más información sobre el contenido del secreto, consulta las [fuentes de datos de Confluence](#).
- Permisos para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el secreto de nombre de usuario y contraseña almacenado por AWS Secrets Manager
- Permisos para utilizar las BatchDeleteDocument operaciones BatchPutDocument y para actualizar el índice.

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

Una política de roles que permite conectarse Amazon Kendra a Confluence.

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

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

Una política de roles que permite conectarse Amazon Kendra a Confluence con la configuración de VPC.

```
{
    "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": [
                "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": [
    "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"
},
{
    "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_*"
        }
    }
},
{
    "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": "*"
}
]
```

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

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

IAMfunciones para las fuentes de datos de Dropbox

Cuando usas Dropbox, proporcionas un rol con las siguientes políticas.

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

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

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

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

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

IAMfunciones para fuentes GitHub de datos

Cuando lo usaGitHub, proporciona un rol con las siguientes políticas.

- Permiso para acceder a su AWS Secrets Manager secreto para autenticar suGitHub.
- Permiso para llamar a las API públicas necesarias para el GitHub conector.

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

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

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

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

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

IAMfunciones para fuentes de datos de Gmail

Cuando usas Gmail, proporcionas un rol con las siguientes políticas.

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

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

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

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

IAMfunciones para las fuentes de datos de Google Drive

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

- Permisos para obtener y descifrar el AWS Secrets Manager secreto que contiene el correo electrónico de la cuenta del cliente, el correo electrónico de la cuenta de administrador y la clave privada necesarios para conectarse al sitio de Google Drive. Para obtener más información sobre el contenido del secreto, consulta las [fuentes de datos de Google Drive](#).
- Permisos para usar las [BatchDeleteDocumentAPI](#) [BatchPutDocumenty](#).

La siguiente IAM política proporciona los permisos necesarios:

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

```
    }]
}
```

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

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

IAMfunciones para las fuentes de datos de Jira

Cuando usas Jira, proporcionas un rol con las siguientes políticas.

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

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

```

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

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

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

IAMfunciones para fuentes de datos de Microsoft Exchange

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

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

La siguiente IAM política proporciona los permisos necesarios:

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

```

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

Si almacena la lista de usuarios para indexarlos en un Amazon S3 bucket, también debe proporcionar permiso para usar la GetObject operación S3. La siguiente IAM política proporciona los permisos necesarios:

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

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

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

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

IAMfunciones para fuentes de OneDrive datos de Microsoft

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

- Permisos para obtener y descifrar el AWS Secrets Manager secreto que contiene el ID de la aplicación y la clave secreta necesarios para conectarse al OneDrive sitio. Para obtener más información sobre el contenido del secreto, consulte [Fuentes de OneDrive datos de Microsoft](#).
- Permisos para usar las [BatchDeleteDocument](#)[API BatchPutDocument](#).

La siguiente IAM política proporciona los permisos necesarios:

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

```

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

Si almacena la lista de usuarios para indexarlos en un Amazon S3 bucket, también debe proporcionar permiso para usar la GetObject operación S3. La siguiente IAM política proporciona los permisos necesarios:

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

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

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

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

IAMfunciones para fuentes de SharePoint datos de Microsoft

IAMfunciones para SharePoint Connector v1.0

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

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

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

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

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

Si ha cifrado el Amazon S3 bucket que contiene el certificado SSL utilizado para comunicarse con el SharePoint sitio, proporcione una política para dar Amazon Kendra acceso a la clave.

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

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

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

}

IAMfunciones para SharePoint Connector v2.0

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

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

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

```
{  
    "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",  
                "kendraDescribePrincipalMapping"  
            ],  
            "Resource": [  
                "arn:aws:kendra:your-region:your-account-id:index/index-id",  
                "arn:aws:kendra:your-region:your-account-id:index/index-id/data-source/*"  
            ]  
        },  
        {  
    ]  
}
```

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

```

```
        "ec2:DescribeNetworkInterfaces",
        "ec2:DescribeAvailabilityZones",
        "ec2:DescribeNetworkInterfaceAttribute",
        "ec2:DescribeVpcs",
        "ec2:DescribeRegions",
        "ec2:DescribeNetworkInterfacePermissions",
        "ec2:DescribeSubnets"
    ],
    "Resource": "*"
}
]
```

Si ha cifrado el Amazon S3 bucket que contiene el certificado SSL utilizado para comunicarse con el SharePoint sitio, proporcione una política para dar Amazon Kendra acceso a la clave.

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

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

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

IAMfunciones para las fuentes de datos de Microsoft Teams

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

- Permiso para obtener y descifrar el AWS Secrets Manager secreto que contiene el ID de cliente y el secreto de cliente necesarios para conectarse a Microsoft Teams. Para obtener más información sobre el contenido del secreto, consulte las [fuentes de datos de Microsoft Teams](#).

La siguiente IAM política proporciona los permisos necesarios:

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

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

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

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

IAMfunciones para las fuentes de datos de Microsoft Yammer

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

- Permiso para obtener y descifrar el AWS Secrets Manager secreto que contiene el ID de la aplicación y la clave secreta necesarios para conectarse al sitio de Microsoft Yammer. Para obtener más información sobre el contenido del secreto, consulte las [fuentes de datos de Microsoft Yammer](#).
- Permiso para usar las [BatchDeleteDocumentAPI](#) [BatchPutDocumenty](#).

La siguiente IAM política proporciona los permisos necesarios:

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

Si almacena la lista de usuarios para indexarlos en un Amazon S3 bucket, también debe proporcionar permiso para usar la GetObject operación S3. La siguiente IAM política proporciona los permisos necesarios:

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

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

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

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

IAMfunciones para las fuentes de datos de Quip

Cuando usa Quip, proporciona un rol con las siguientes políticas.

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

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "quip:BatchPutDocument"
      ]
    }
  ]
}
```

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

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

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

IAMfunciones para las fuentes de datos de Salesforce

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

- Permisos para acceder al AWS Secrets Manager secreto que contiene el nombre de usuario y la contraseña del sitio de Salesforce. Para obtener más información sobre el contenido del secreto, consulte las [fuentes de datos de Salesforce](#).
- Permisos para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el secreto de nombre de usuario y contraseña almacenado por Secrets Manager
- Permisos para utilizar las BatchDeleteDocument operaciones BatchPutDocument y para actualizar el índice.

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

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

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

IAMfunciones para fuentes ServiceNow de datos

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

- Permisos para acceder al Secrets Manager secreto que contiene el nombre de usuario y la contraseña del ServiceNow sitio. Para obtener más información sobre el contenido del secreto, consulte [fuentes ServiceNow de datos](#).
- Permisos para usar la clave maestra del AWS KMS cliente (CMK) para descifrar el secreto de nombre de usuario y contraseña almacenado por Secrets Manager
- Permisos para utilizar las BatchDeleteDocument operaciones BatchPutDocument y para actualizar el índice.

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

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

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

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

IAMfunciones para las fuentes de datos de Slack

Cuando usas Slack, proporcionas un rol con las siguientes políticas.

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

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

```
    }]
}
```

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

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

IAMfunciones para las fuentes de datos de Zendesk

Cuando usas Zendesk, proporcionas un rol con las siguientes políticas.

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

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

```

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

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

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

Función de nube privada virtual (VPC) IAM

Si utilizas una nube privada virtual (VPC) para conectarte a tu fuente de datos, debes proporcionar los siguientes permisos.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ec2>CreateNetworkInterface",
                "ec2>DescribeNetworkInterfaces",
                "ec2>DeleteNetworkInterface"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "ec2>CreateNetworkInterfacePermission"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "ec2:AuthorizedService": "kendra.*.amazonaws.com"
                },
                "ArnEquals": {

```

```
        "ec2:Subnet": [
            "arn:aws:ec2:your-region:your-account-id:subnet/subnet-ids"
        ]
    }
},
{
    "Effect": "Allow",
    "Action": [
        "ec2:DescribeSubnets"
    ],
    "Resource": "*"
},
{
    "Sid": "iamPassRole",
    "Effect": "Allow",
    "Action": "iam:PassRole",
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "iam:PassedToService": [
                "kendra.*.amazonaws.com"
            ]
        }
    }
}
]
```

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

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

IAMfunciones para preguntas frecuentes (FAQ)

Cuando utilice la [CreateFaq](#)API para cargar preguntas y respuestas en un índice, debe proporcionar Amazon Kendra un IAM rol con acceso al Amazon S3 bucket que contiene los archivos fuente. Si los archivos fuente están cifrados, debe proporcionar permiso para usar la clave maestra del AWS KMS cliente (CMK) para descifrar los archivos.

IAMroles para preguntas frecuentes

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

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

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

Una política de roles opcional que permite Amazon Kendra usar una clave maestra de AWS KMS cliente (CMK) para descifrar los archivos de un Amazon S3 bucket.

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

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

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

IAMfunciones para sugerencias de consultas

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

IAMfunciones para sugerencias de consultas

Una política de roles obligatoria que permita Amazon Kendra utilizar el Amazon S3 archivo como lista de bloqueo de sugerencias de consultas.

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

Una política de roles opcional que permite Amazon Kendra usar una clave maestra de AWS KMS cliente (CMK) para descifrar los documentos de un Amazon S3 bucket.

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

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

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

IAMfunciones para el mapeo principal de usuarios y grupos

Al utilizar la [PutPrincipalMapping API](#) para asignar usuarios a sus grupos a fin de filtrar los resultados de búsqueda por contexto de usuario, debe proporcionar una lista de los usuarios o subgrupos que

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

IAMfunciones para el mapeo principal

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

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

Una política de roles opcional que permite Amazon Kendra usar una clave maestra de AWS KMS cliente (CMK) para descifrar los documentos de un Amazon S3 bucket.

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

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

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

Se recomienda incluir aws:sourceAccount y aws:sourceArn en la política de confianza. Esto limita los permisos y comprueba de forma segura si aws:sourceAccount aws:sourceArn son los mismos que los que se proporcionan en la política de IAM roles para la sts:AssumeRole acción. Esto evita que

entidades no autorizadas accedan a sus IAM funciones y a sus permisos. Para obtener más información, consulte la AWS Identity and Access Management guía sobre el [problema del diputado confuso](#).

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

Roles de IAM para AWS IAM Identity Center (successor to AWS Single Sign-On)

Cuando utilice el [UserGroupResolutionConfiguration](#) objeto para obtener los niveles de acceso de grupos y usuarios de una fuente de AWS IAM Identity Center (successor to AWS Single Sign-On) identidad, debe proporcionar un rol que tenga permiso de acceso a IAM Identity Center.

Roles de IAM para AWS IAM Identity Center (successor to AWS Single Sign-On)

Una política de roles obligatoria para Amazon Kendra permitir el acceso a 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",  
            "Condition": {  
                "StringEquals": {  
                    "aws:SourceAccount": "your-account-id"  
                }  
            }  
        }  
    ]  
}
```

```
    "Resource": "*",
    "Condition": [
        "StringEquals": {
            "iam:PassedToService": [
                "kendra.*.amazonaws.com"
            ]
        }
    ]
}
```

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

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

IAMroles para Amazon Kendra experiencias

Cuando utilice las [UpdateExperienceAPI](#) [CreateExperience](#)o para crear o actualizar una aplicación de búsqueda, debe proporcionar un rol que tenga permiso para acceder a las operaciones necesarias y a IAM Identity Center.

IAMroles para la experiencia Amazon Kendra de búsqueda

Una política de roles obligatoria para permitir el acceso Amazon Kendra a Query operaciones, QuerySuggestions SubmitFeedback operaciones y al centro de identidades de IAM, que almacena la información de usuarios y grupos.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowsKendraSearchAppToCallKendraApi",
            "Effect": "Allow",
            "Action": [
                "kendra:GetQuerySuggestions",
                "kendra:Query",
                "kendra:DescribeIndex",
                "kendra>ListFaqs",
                "kendra:DescribeDataSource",
                "kendra>ListDataSources",
                "kendra:DescribeFaq",
                "kendra:SubmitFeedback"
            ],
            "Resource": [
                "arn:aws:kendra:your-region:your-account-id:index/index-id"
            ]
        }
    ]
},
```

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

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

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

Se recomienda incluir `aws:sourceAccount` y `aws:sourceArn` en la política de confianza. Esto limita los permisos y comprueba de forma segura si `aws:sourceAccount` `aws:sourceArn` son los mismos que los que se proporcionan en la política de IAM roles para la `sts:AssumeRole` acción. Esto evita que entidades no autorizadas accedan a sus IAM funciones y a sus permisos. Para obtener más información, consulte la AWS Identity and Access Management guía sobre el [problema del diputado confuso](#).

```
{
```

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

IAMfunciones para el enriquecimiento personalizado de documentos

Cuando utilice el [CustomDocumentEnrichmentConfiguration](#) objeto para aplicar modificaciones avanzadas a los metadatos y el contenido del documento, debe proporcionar un rol que tenga los permisos necesarios para ejecutarse `PreExtractionHookConfiguration` y/o `PostExtractionHookConfiguration`. Configura una función Lambda `PostExtractionHookConfiguration` para `PreExtractionHookConfiguration` o aplica alteraciones avanzadas de los metadatos y el contenido del documento durante el proceso de ingestión. Si decide activar el cifrado del lado del servidor para su Amazon S3 bucket, debe proporcionar permiso para usar la clave maestra del AWS KMS cliente (CMK) para cifrar y descifrar los objetos almacenados en su bucket. Amazon S3

IAMfunciones para el enriquecimiento personalizado de documentos

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

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

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

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

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:GetObject",
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name/*"
      ],
      "Effect": "Allow"
    },
    {
      "Action": [
        "s3>ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name"
      ],
      "Effect": "Allow"
    },
    {
      "Effect": "Allow",
      "Action": [
        "lambda:InvokeFunction"
      ],
      "Resource": "arn:aws:lambda:your-region:your-account-id:function:lambda-function"
    }
  ]
}
```

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

```
{
```

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

Se recomienda incluir `aws:sourceAccount` y `aws:sourceArn` en la política de confianza. Esto limita los permisos y comprueba de forma segura si `aws:sourceAccount` `aws:sourceArn` son los mismos que los que se proporcionan en la política de IAM roles para la `sts:AssumeRole` acción. Esto evita que entidades no autorizadas accedan a sus IAM funciones y a sus permisos. Para obtener más información, consulte la AWS Identity and Access Management guía sobre el [problema del diputado confuso](#).

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

Implementación de Amazon Kendra

Cuando llega el momento de la implementaciónAmazon Kendrabusque en su sitio web, le proporcionamos el código fuente que puede usar con React para comenzar con su aplicación. El código fuente se proporciona sin cargo bajo una licencia MIT modificada. Puede usarlo tal cual o cambiarlo según sus propias necesidades. La aplicación React proporcionada es un ejemplo que le ayudará a empezar. No es una aplicación lista para producción.

Para implementar una aplicación de búsqueda sin código y generar una URL de punto final para la página de búsqueda con control de acceso, consulte [Amazon KendraCreador de experiencias](#).

El siguiente código de ejemplo agregaAmazon Kendrabusque una aplicación web React existente:

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

Los ejemplos siguen el modelo de la página de búsqueda delAmazon Kendraconsola. Tienen las mismas funciones para buscar y mostrar los resultados de búsqueda. Puede utilizar todo el ejemplo o puede elegir solo una de las funciones para su propio uso.

Para ver los tres componentes de la página de búsqueda enAmazon Kendraconsola, selecciona el ícono de código (</>) en el menú de la derecha. Pase el puntero sobre cada sección para ver una breve descripción del componente y obtener la URL de la fuente del componente.

Temas

- [Información general \(p. 74\)](#)
- [Requisitos previos \(p. 75\)](#)
- [Configuración del ejemplo \(p. 75\)](#)
- [Página principal de búsqueda \(p. 76\)](#)
- [Componente de búsqueda \(p. 76\)](#)
- [Componente de resultados \(p. 76\)](#)
- [Componente de facetas \(p. 76\)](#)
- [Componente de paginación \(p. 76\)](#)
- [Creación de una experiencia de búsqueda sin código \(p. 77\)](#)

Información general

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

- Página principal de búsqueda (Search.tsx): esta es la página principal que contiene todos los componentes. Aquí es donde integras tu aplicación con laAmazon KendraAPI.
- Barra de búsqueda: es el componente en el que el usuario introduce un término de búsqueda y llama a la función de búsqueda.

- Resultados: este es el componente que muestra los resultados de Amazon Kendra. Tiene tres componentes: respuestas sugeridas, resultados de preguntas frecuentes y documentos recomendados.
- Facetas: este es el componente que muestra las facetas de los resultados de la búsqueda y permite elegir una faceta para restringir la búsqueda.
- Paginación: este es el componente que pagina la respuesta de Amazon Kendra.

Requisitos previos

Antes de comenzar, necesitará lo siguiente:

- Node.js y npm [instalada](#). Se requiere la versión 19 o anterior de Node.js.
- Python 3 o Python 2 [instalada](#).
- [SDK for Java](#) o [AWS SDK for JavaScript](#) para realizar llamadas a la API a Amazon Kendra.
- Una aplicación web React existente. El código de ejemplo incluye un archivo Readme con pasos sobre cómo configurar un nuevo entorno de desarrollo de React, incluido el uso de los marcos y bibliotecas necesarios. También puede seguir las instrucciones de inicio rápido de [Documentación de React sobre la creación de una aplicación web React](#).
- Las bibliotecas y dependencias necesarias configuradas en su entorno de desarrollo. El código de ejemplo incluye un archivo Readme que enumera las bibliotecas y dependencias de paquetes necesarias. Tenga en cuenta que `sass` es obligatorio, ya que `node-sass` está en desuso. Si ha instalado anteriormente `node-sass`, desinstala esto e instala `sass`.

Configuración del ejemplo

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

Para empezar a usar `kendrasamples-react-app.zip`

1. Asegúrese de haber completado el [Requisitos previos \(p. 75\)](#), incluida la descarga e instalación de Node.js y npm.
2. Descargue `kendrasamples-react-app.zip` y descomprima.
3. Abre tu terminal y ve a `aaws-kendra-example-react-app/src/services/`. Abrir `local-dev-credentials.json` y proporcione sus credenciales. No añada este archivo a ningún repositorio público.
4. Ir a `aaws-kendra-example-react-app` e instala las dependencias en `package.json`. Ejecute `npm install`.
5. Inicie una versión de demostración de la aplicación en su servidor local. Ejecute `npm start`. Puede detener el servidor local ingresando con el teclado Cmd/Ctrl + C.
6. Puede cambiar el puerto o el host (por ejemplo, la dirección IP) yendo a `package.json` y actualice el host y el puerto: "start": "HOST=[host] PORT=[port] react-scripts start". Si usa Windows: "start": "set HOST=[host] && set PORT=[port] && react-scripts start".
7. Si tiene un dominio de sitio web registrado, puede especificarlo en `package.json` después del nombre de la aplicación. Por ejemplo, "homepage": "<https://mywebsite.com>". Debes correr `npm install` nuevamente para actualizar las nuevas dependencias y, a continuación, ejecutar `npm start`.
8. Para crear la aplicación, ejecuta `npm build`. Sube el contenido del directorio de compilación a tu proveedor de alojamiento.

Warning

La aplicación React esnolisto para producción. Es un ejemplo de implementación de una aplicación paraAmazon Kendrabuscar.

Página principal de búsqueda

La página principal de búsqueda (`Search.tsx`) contiene todos los componentes de búsqueda del ejemplo. Incluye el componente de la barra de búsqueda para la salida y los componentes de resultados para mostrar la respuesta del[ConsultaAPI](#) y un componente de paginación para buscar la respuesta.

Componente de búsqueda

El componente de búsqueda proporciona un cuadro de texto para introducir el texto de la consulta. El `onSearch` función es un gancho que llama a la función principal `enSearch.tsx` para hacer el Amazon Kendra [Consulta](#) llamada a la API.

Componente de resultados

El componente de resultados muestra la respuesta del `QueryAPI`. Los resultados se muestran en tres áreas separadas.

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

Los componentes de resultados comparten un conjunto de componentes para funciones como el resaltado, los títulos, los enlaces y mucho más. Los componentes compartidos deben estar presentes para que los componentes del resultado funcionen.

Componente de facetas

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

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

Componente de paginación

El componente de paginación permite mostrar los resultados de la búsqueda desde `QueryAPI` en varias páginas. Llama al `QueryAPI` con `PageSize` y `PageNumber` parámetros para obtener una página específica de resultados.

Creación de una experiencia de búsqueda sin código

Puede crear e implementar un Amazon Kendra aplicación de búsqueda sin necesidad de ningún código de interfaz. Amazon Kendra Creador de experiencias le ayuda a crear e implementar una aplicación de búsqueda completamente funcional con unos pocos clics para que pueda empezar a buscar de inmediato. Puedes personalizar tu página de búsqueda y ajustar la búsqueda para adaptar la experiencia a las necesidades de tus usuarios. Amazon Kendra genera una URL de extremo único y completamente alojada de su página de búsqueda para empezar a buscar en sus documentos y preguntas frecuentes. Puedes crear rápidamente una prueba de concepto de tu experiencia de búsqueda y compartirla con otros usuarios.

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

Cómo funciona la búsqueda Experience Builder

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

1. Para crear su experiencia de búsqueda, debe asignarle un nombre, una descripción y elegir las fuentes de datos que desea utilizar para su experiencia de búsqueda.
2. La lista de usuarios y grupos se configura en AWS IAM Identity Center (successor to AWS Single Sign-On) y luego asigneles derechos de acceso a su experiencia de búsqueda. Te incluyes como propietario de la experiencia. Para obtener más información, consulte [the section called “Proporcionar acceso a tu página de búsqueda” \(p. 78\)](#).
3. Abre el Amazon Kendra Experience Builder para diseñar y ajustar tu página de búsqueda. Puedes compartir la URL de tu experiencia de búsqueda con otras personas a las que les asigne derechos de acceso propios para editar o para ver y buscar.

Tú llamas al [CreateExperience](#) API para crear y configurar tu experiencia de búsqueda. Si utilizas la consola, seleccionas tu índice y, a continuación, seleccionas Experiencias en el menú de navegación para configurar tu experiencia.

Diseña y ajusta tu experiencia de búsqueda

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

Las siguientes son formas de personalizar la experiencia de búsqueda.

Filtro

Añada una búsqueda por facetas o filtro por atributos del documento. Esto incluye atributos personalizados. Puede añadir un filtro mediante sus propios campos de metadatos configurados. Por

ejemplo, para buscar por facetas por categoría de ciudad, utilice un _category atributo de documento personalizado que contiene todas las categorías de ciudades.

Respuesta sugerida

Añada respuestas generadas por aprendizaje automático a las consultas de sus usuarios. Por ejemplo, «¿Qué tan difícil es este curso?». Amazon Kendrapuede recuperar el texto más relevante de todos los documentos relacionados con la dificultad de un curso y sugerir la respuesta más relevante.

Preguntas frecuentes

Añada un documento de preguntas frecuentes para responder a las preguntas más frecuentes. Por ejemplo, «¿Cuántas horas hay para completar este curso?». Amazon Kendrapuede utilizar el documento de preguntas frecuentes que contiene la respuesta a esta pregunta y dar la respuesta correcta.

Sort

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

Documentos de

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

Lenguaje

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

Cuadro de búsqueda

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

Ajuste de relevancia

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

Proporcionar acceso a tu página de búsqueda

El acceso a su experiencia de búsqueda se realiza a través de IAM Identity Center. Al configurar su experiencia de búsqueda, concede a las demás personas que figuran en el directorio de Identity Center acceso a suAmazon Kendrapágina de búsqueda. Reciben un correo electrónico que les indica que inicien sesión con sus credenciales en IAM Identity Center para acceder a la página de búsqueda. Debe configurar IAM Identity Center a nivel de organización o de titular de la cuenta enAWS Organizations. Para obtener más información sobre la configuración de IAM Identity Center, consulte [Introducción a IAM Identity Center](#).

Usted activa las identidades de usuario en IAM Identity Center con su experiencia de búsqueda y asignaVisor o Dueño permisos de acceso mediante la API o la consola.

- Visor: Se le permite realizar consultas, recibir sugerencias de respuestas relevantes para su búsqueda y aportar sus comentarios a Amazon Kendra para que siga mejorando la búsqueda.
- Dueño: Permite personalizar el diseño de la página de búsqueda, ajustar la búsqueda y utilizar la aplicación de búsqueda como Visor. Actualmente, no se permite deshabilitar el acceso a los espectadores en la consola.

Para asignar acceso a otras personas a tu experiencia de búsqueda, primero debes activar las identidades de usuario en IAM Identity Center con tu Amazon Kendra experiencia mediante el uso del [ExperienceConfiguration](#) objeto. Especifica el nombre del campo que contiene los identificadores de los usuarios, como el nombre de usuario o la dirección de correo electrónico. A continuación, concedes a tu lista de usuarios acceso a tu experiencia de búsqueda mediante el [AssociateEntitiesToExperience](#) API y define sus permisos como Visor o Dueño utilizando el [AssociatePersonsToEntities](#) API. Para especificar cada usuario o grupo, utilice la [EntityConfiguration](#) objeto y si ese usuario o grupo es un Visor o Dueño utilizando el [EntityPersonaConfiguration](#) objeto.

Para asignar acceso a otras personas a tu experiencia de búsqueda mediante la consola, primero tienes que crear una experiencia y confirmar tu identidad y que eres el propietario. A continuación, puedes asignar a otros usuarios o grupos como espectadores o propietarios. En la consola, selecciona tu índice y, a continuación, selecciona Experiencias en el menú de navegación. Despues de crear tu experiencia, puedes seleccionarla de la lista. Ir a Gestión de accesos para asignar usuarios o grupos como espectadores o propietarios.

Configuración de una experiencia de búsqueda

El siguiente es un ejemplo de configuración o creación de una experiencia de búsqueda.

Console

Para crear un Amazon Kendra experiencia de búsqueda

1. En el panel de navegación izquierdo, en Índices, selecciona Experiencias y luego selecciona Crea experiencia.
2. En el Configurar la experiencia página, introduzca un nombre y una descripción para su experiencia, elija las fuentes de contenido y elija el rol de IAM para su experiencia. Para obtener más información sobre las funciones de IAM, consulte [Funciones de IAM para Amazon Kendra experiencias](#).
3. En el Confirme su identidad desde el directorio del Centro de Identidad página, selecciona tu seudónimo, como tu correo electrónico. Si no tiene un directorio de Identity Center, simplemente introduzca su nombre completo y correo electrónico para crear un directorio de Identity Center. Esto lo incluye como usuario de la experiencia y le asigna automáticamente los derechos de acceso de propietario.
4. En el Revise para abrir Experience Builder página, revise los detalles de configuración y selecciona Cree experiencias y abra Experience Builder para empezar a editar la página de búsqueda.

CLI

Para crear un Amazon Kendra experiencia

```
aws kendra create-experience \
--name experience-name \
--description "experience description" \
```

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

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

Python

Para crear un Amazon Kendra experiencia

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

kendra = boto3.client("kendra")

print("Create an experience.")

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

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

    pprint.pprint(experience_response)

    experience_endpoints = experience_response["Endpoints"]

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

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

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

Java

Para crear un Amazon Kendra

```
package com.amazonaws.kendra;  
  
import java.util.concurrent.TimeUnit;  
import software.amazon.awssdk.services.kendra.KendraClient;  
import software.amazon.awssdk.services.kendra.model.CreateExperienceRequest;  
import software.amazon.awssdk.services.kendra.model.CreateExperienceResponse;  
import software.amazon.awssdk.services.kendra.model.DescribeExperienceRequest;  
import software.amazon.awssdk.services.kendra.model.DescribeExperienceResponse;  
import software.amazon.awssdk.services.kendra.model.ExperienceStatus;  
  
public class CreateExperienceExample {  
  
    public static void main(String[] args) throws InterruptedException {  
        System.out.println("Create an experience");  
  
        String experienceName = "experience-name";  
        String experienceDescription = "experience description";  
        String indexId = "index-id";  
        String experienceRoleArn = "arn:aws:iam::account-id:role/role-name";  
  
        KendraClient kendra = KendraClient.builder().build();  
  
        CreateExperienceRequest createExperienceRequest = CreateExperienceRequest  
            .builder()  
            .name(experienceName)  
            .description(experienceDescription)  
            .roleArn(experienceRoleArn)  
            .configuration(  
                ExperienceConfiguration  
                    .builder()  
                    .contentSourceConfiguration(  
                        ContentSourceConfiguration(  
                            .builder()  
                            .dataSourceIds("data-source-1","data-source-2")  
                            .build()  
                        )  
                    )  
            )  
            .userIdentityConfiguration(  
                UserIdentityConfiguration(  
                    .builder()  
                    .identityAttributeName("identity-attribute-name")  
                    .build()  
                )  
            ).build()  
        ).build();  
  
        CreateExperienceResponse createExperienceResponse =  
kendra.createExperience(createExperienceRequest);  
        System.out.println(String.format("Experience response %s",  
createExperienceResponse));  
  
        String experienceEndpoints = createExperienceResponse.endpoints();  
  
        System.out.println(String.format("Wait for Kendra to create the experience.",  
experienceEndpoints));
```

```
        while (true) {
            DescribeExperienceRequest describeExperienceRequest =
DescribeExperienceRequest.builder().endpoints(experienceEndpoints).build();
            DescribeExperienceResponse describeExperienceResponse =
kendra.describeExperience(describeExperienceRequest);
            ExperienceStatus status = describeExperienceResponse.status();
            TimeUnit.SECONDS.sleep(60);
            if (status != ExperienceStatus.CREATING) {
                break;
            }
        }
        System.out.println("Experience creation is complete.");
    }
}
```

Capacidad de ajuste

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

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

- 100 000 documentos o 30 GB de almacenamiento.

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

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

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

Puede añadir hasta 100 unidades de capacidad adicionales a su almacenamiento y consultar los recursos para obtener un índice. Si necesitas más unidades, ponte en [contacto con el servicio de asistencia](#).

Puede ajustar las unidades de capacidad hasta 5 veces al día para que se ajusten a sus requisitos de uso. No puedes reducir la capacidad de almacenamiento de documentos por debajo del número de documentos almacenados en tu índice. Por ejemplo, si almacena 150 000 documentos, no puedes reducir la capacidad de almacenamiento a menos de 1 unidad adicional.

Puede ver los recursos que utiliza un índice en la consola seleccionando el nombre del índice para abrir la configuración del índice y otra información, o bien puede utilizar la [DescribeIndex](#) API. Amazon Kendra también devuelve excepciones cuando se supera la capacidad de un índice. Se obtiene un [ServiceQuotaExceededException](#) cuando el tamaño total extraído de todos los documentos supera el límite de un índice. Se obtiene una [InvalidRequest](#) para cada documento cuando el número de documentos supera el límite de un índice. Se obtiene un [ThrottlingException](#) cuando el número de consultas por segundo supera el límite. Para obtener más información sobre los límites, consulte [Cuotas para Amazon Kendra](#).

Capacidad de visualización

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

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

1. Inicie sesión AWS Management Console y abra la Amazon Kendra consola en <https://console.aws.amazon.com/kendra/home>.
2. En la lista de índices, selecciona el índice al que deseas acceder.

3. Desplázate hasta la sección de configuración para ver la capacidad total actual de almacenamiento de documentos y consultas.

Para ver la capacidad mediante la Amazon Kendra API, utilice el CapacityUnits parámetro de la [DescribelIndexAPI](#).

Añadir y eliminar capacidad

Si necesitas capacidad adicional para tu índice, puedes añadirla mediante la consola o la Amazon Kendra API.

Para añadir o quitar capacidad de almacenamiento o consultas (consola)

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

Tras actualizar la capacidad del índice, los cambios pueden tardar varios minutos en surtir efecto.

Para añadir o eliminar capacidad mediante la Amazon Kendra API, utilice el CapacityUnits parámetro de la [UpdateIndexAPI](#).

Amazon KendraCapacidad de clasificación inteligente

Una unidad de capacidad proporciona las siguientes solicitudes de rescore adicionales por segundo para un plan de ejecución de rescore. Un plan de ejecución de rescore es un recurso que se utiliza para aprovisionar la API [Rescore](#).

- 0.01 solicitudes por segundo.

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

Puede añadir hasta 1000 unidades de capacidad adicionales para un plan de ejecución de rescore. Si necesitas más unidades, ponte en [contacto con el servicio de asistencia](#).

Capacidad de sugerencias de consultas

Cuando se utilizan [sugerencias de consulta](#), hay una capacidad de consulta base de 2,5 [GetQuerySuggestions](#) llamadas por segundo. La GetQuerySuggestions capacidad es cinco veces

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

Amazon Kendracapacidad de experiencia

Capacidad de experiencia de búsqueda

Amazon Kendra comienza a acelerarse QuerySuggestions, SubmitFeedback según su Amazon Kendra experiencia, con 15 solicitudes por segundo y 40 solicitudes por segundo por ráfaga de consultas. Para un índice con más de 150 unidades de capacidad de consulta, estos límites se siguen aplicando.

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

Ráfaga de consultas adaptable

Amazon Kendra tiene una capacidad base aprovisionada de 1 unidad de capacidad de consulta. Puede utilizar hasta 8 000 consultas al día con un rendimiento mínimo de 0,1 consultas por segundo (por unidad de capacidad de consulta). Las consultas acumuladas durarán hasta 24 horas y pueden albergar ráfagas de tráfico. La cantidad de ráfaga permitida varía porque depende de la carga del clúster en un momento dado. Aprovisione suficientes unidades de capacidad de consulta para gestionar sus niveles de carga máxima.

La ráfaga de consultas adaptativa integrada es un enfoque adaptable para gestionar ráfagas inesperadas de tráfico más allá Amazon Kendra del rendimiento aprovisionado. La fragmentación adaptativa de consultas está disponible en la edición empresarial de Amazon Kendra.

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

Introducción

En esta sección se muestra cómo crear una fuente de datos y añadir los documentos a un Amazon Kendraíndice. Se proporcionan instrucciones para AWSconsola, la AWS CLI, un programa de Python que utilizaAWS SDK for Python (Boto3), y un programa Java que utiliza elAWS SDK for Java.

Temas

- [Requisitos previos \(p. 86\)](#)
- [Empezar con elAmazon Kendraconsola \(p. 91\)](#)
- [Introducción \(AWS CLI\) \(p. 92\)](#)
- [Introducción \(AWS SDK for Python \(Boto3\)\) \(p. 93\)](#)
- [Introducción \(AWS SDK for Java\) \(p. 96\)](#)
- [Introducción a una fuente Amazon S3 de datos \(consola\) \(p. 98\)](#)
- [Introducción a una fuente de datos de base de datos MySQL \(consola\) \(p. 99\)](#)
- [Empezar con unAWS IAM Identity Center \(successor to AWS Single Sign-On\)fuentede identidad \(consola\) \(p. 101\)](#)

Requisitos previos

Los siguientes pasos son requisitos previos para los ejercicios de iniciación. Los pasos le muestran cómo configurar su cuenta, crear unalAMrol que daAmazon Kendrapermiso para realizar llamadas en su nombre e indexar documentos de unAmazon S3balde. Se utiliza un bucket de S3 como ejemplo, pero puede utilizar otras fuentes de datos queAmazon Kendrasoportes. Consulte[Fuentes de datos](#).

Registro en una Cuenta de AWS

Si no dispone de una Cuenta de AWS, siga los pasos que figuran a continuación para crear una.

Para registrarse en Cuenta de AWS

1. Abra <https://portal.aws.amazon.com/billing/signup>.
2. Siga las instrucciones en línea.

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

Al registrarse en una Cuenta de AWS, se crea un Usuario raíz de la cuenta de AWS. El usuario raíz tiene acceso a todos los recursos y Servicios de AWS de esa cuenta. Como práctica recomendada de seguridad, [asigne acceso administrativo a un usuario administrativo](#) y utilice únicamente el usuario raíz para realizar la ejecución [tareas que requieren acceso de usuario raíz](#).

AWS le enviará un email de confirmación luego de completar el proceso de registro. Puede ver la actividad de la cuenta y administrar la cuenta en cualquier momento entrando en <https://aws.amazon.com/> y seleccionando My Account (Mi cuenta).

Crear un usuario administrativo

Después de registrarse para obtener una Cuenta de AWS, cree un usuario administrativo para que no utilice el usuario raíz en las tareas cotidianas.

Proteger su Usuario raíz de la cuenta de AWS

1. Inicie sesión en [AWS Management Console](#) como propietario de cuenta eligiendo Usuario raíz e ingrese el email de su Cuenta de AWS. En la siguiente página, escriba su contraseña.

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

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

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

Crear un usuario administrativo

- Para las tareas administrativas diarias, conceda acceso administrativo a un usuario administrativo en AWS IAM Identity Center (successor to AWS Single Sign-On).
Para obtener instrucciones, consulte [Introducción](#) en la Guía del usuario de AWS IAM Identity Center (successor to AWS Single Sign-On).

Iniciar sesión como usuario administrativo

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

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

- Si está utilizando un bucket de S3 que contiene documentos para realizar la pruebaAmazon Kendra, cree un bucket de S3 en la misma región que está utilizandoAmazon Kendra. Para obtener instrucciones, consulte [Creación y configuración de un bucket S3](#) en elGuía del usuario de Amazon Simple Storage Service.

Sube tus documentos a tu bucket de S3. Para obtener instrucciones, consulte [Carga, descarga y administración de objetos](#) en elGuía del usuario de Amazon Simple Storage Service.

Si utiliza otra fuente de datos, debe tener un sitio activo y credenciales para conectarse a la fuente de datos.

Si está utilizando la consola para empezar, comience con [Empezar con elAmazon Kendraconsola \(p. 91\)](#).

Amazon Kendra recursos: AWS CLI, SDK, consola

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

Para usarAmazon Kendrapara la CLI, el SDK o la consola, debe tener permisos para permitirAmazon Kendrapara crear y administrar recursos en su nombre. Según su caso de uso, estos permisos incluyen el acceso aAmazon KendraLa propia API,AWS KMS keyssi desea cifrar sus datos a través de un directorio

CMK personalizado, Identity Center si desea integrarlos con AWS IAM Identity Center (successor to AWS Single Sign-On) o [crear una experiencia de búsqueda](#). Para obtener una lista completa de los permisos para diferentes casos de uso, consulte [IAM papeles](#).

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

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Sid": "Stmt1644430853544",  
            "Action": [  
                "kms>CreateGrant",  
                "kms>DescribeKey"  
            ],  
            "Effect": "Allow",  
            "Resource": "*"  
        },  
        {  
            "Sid": "Stmt1644430878150",  
            "Action": "kendra:*",  
            "Effect": "Allow",  
            "Resource": "*"  
        },  
        {  
            "Sid": "Stmt1644430973706",  
            "Action": [  
                "sso:AssociateProfile",  
                "sso>CreateManagedApplicationInstance",  
                "sso>DeleteManagedApplicationInstance",  
                "sso:DisassociateProfile",  
                "sso:GetManagedApplicationInstance",  
                "sso:GetProfile",  
                "sso>ListDirectoryAssociations",  
                "sso>ListProfileAssociations",  
                "sso>ListProfiles"  
            ],  
            "Effect": "Allow",  
            "Resource": "*"  
        },  
        {  
            "Sid": "Stmt1644430999558",  
            "Action": [  
                "sso-directory:DescribeGroup",  
                "sso-directory:DescribeGroups",  
                "sso-directory:DescribeUser",  
                "sso-directory:DescribeUsers"  
            ],  
            "Effect": "Allow",  
            "Resource": "*"  
        },  
        {  
            "Sid": "Stmt1644431025960",  
            "Action": [  
                "identitystore:DescribeGroup",  
                "identitystore:DescribeUser",  
                "identitystore>ListGroups",  
                "identitystore>ListUsers"  
            ],  
            "Effect": "Allow",  
            "Resource": "*"  
        }  
    ]  
}
```

En segundo lugar, si usa la CLI o el SDK, también debe crear un IAMrol y política de acceso Amazon CloudWatch Logs. Si está utilizando la consola, no necesita crear un IAMpapel y política para ello. Esto se crea como parte del procedimiento de la consola.

Para crear un IAMfunción y política de la AWS CLy un SDK que permite Amazon Kendra para acceder a tu Amazon CloudWatch Logs.

1. Inicie sesión en la AWS Management Console y abra la consola de IAM en <https://console.aws.amazon.com/iam/>.
2. En el menú de la izquierda, selecciona Políticas y luego elige Crear política.
3. Elige JSONy, a continuación, sustituya la política predeterminada por la siguiente:

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

4. Elija Review policy (Revisar política).
5. Ponle un nombre a la política "KendraPolicyForGettingStartedIndex" y luego elige Crear política.
6. En el menú de la izquierda, selecciona Funciones y luego elige Crear rol.

7. Elije Otro AWS cuenta y, a continuación, escriba el ID de su cuenta en ID de cuenta. Elija Next: Permissions (Siguiente: permisos).
8. Elija la política que creó anteriormente y, a continuación, elija Siguiente: Etiquetas.
9. No añada ninguna etiqueta. Elija Next: Review (Siguiente: revisar).
10. Asigne un nombre al rol "KendraRoleForGettingStartedIndex" y luego elige Crear rol.
11. Busca el rol que acabas de crear. Elija el nombre del rol para abrir el resumen. Elije Relaciones de confianza y luego elige Editar relación de confianza.
12. Sustituya la relación de confianza existente por la siguiente:

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

13. Elija Update trust policy (Actualizar política de confianza).

En tercer lugar, si usa un Amazon S3 para almacenar sus documentos o está utilizando S3 para probar Amazon Kendra, también debes crear un IAM rol y política para acceder a tu bucket. Si utiliza otra fuente de datos, consulte [IAM funciones para fuentes de datos](#).

Para crear un IAM rol y política que permiten Amazon Kendra para acceder e indexar sus Amazon S3 balde.

1. Inicie sesión en la AWS Management Console y abra la consola de IAM en <https://console.aws.amazon.com/iam/>.
2. En el menú de la izquierda, selecciona Políticas y luego elige Crear política.
3. Elije JSON, a continuación, sustituya la política predeterminada por la siguiente:

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

```
        "Action": [
            "kendra:BatchPutDocument",
            "kendra:BatchDeleteDocument"
        ],
        "Resource": "arn:aws:kendra:region:account ID:index/*"
    }
}
```

4. Elija Review policy (Revisar política).
5. Asigne un nombre a la política»KendraPolicyForGettingStartedDataSource«y luego eligeCrear política.
6. En el menú de la izquierda, seleccionaFuncionesy luego eligeCrear rol.
7. EligeOtroAWScuentay, a continuación, escriba el ID de su cuenta enID de cuenta. Elija Next: Permissions (Siguiente: permisos).
8. Elija la política que creó anteriormente y, a continuación, elijaSiguiente: Etiquetas
9. No añada ninguna etiqueta. Elija Next: Review (Siguiente: revisar).
10. Asigne un nombre al rol»KendraRoleForGettingStartedDataSource«y luego eligeCrear rol.
11. Busca el rol que acabas de crear. Elija el nombre del rol para abrir el resumen. EligeRelaciones de confianzay luego eligeEditar relación de confianza.
12. Sustituya la relación de confianza existente por la siguiente:

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

13. Elija Update trust policy (Actualizar política de confianza).

Dependiendo de cómo desee utilizar elAmazon KendraAPI, realice una de las siguientes acciones.

- [Introducción \(AWS CLI\) \(p. 92\)](#)
- [Introducción \(AWS SDK for Java\) \(p. 96\)](#)
- [Introducción \(AWS SDK for Python \(Boto3\)\) \(p. 93\)](#)

Empezar con elAmazon Kendraconsola

Los siguientes procedimientos muestran cómo crear y probar unAmazon Kendraindexar mediante elAWSconsola. En los procedimientos, se crea un índice y una fuente de datos para un índice. Por último, prueba el índice realizando una solicitud de búsqueda.

Paso 1: Para crear un índice (consola)

1. Inicia sesión enAWSConsola de administración y abraAmazon Kendraconsola en<https://console.aws.amazon.com/kendra/>.
2. Seleccione Create index (Crear índice) en la sección Indexes (Índices).
3. En elEspecificar los detalles del índicepagina, asigne un nombre y una descripción a su índice.

4. En IAMpapel, eligeCrear un nuevo roly, a continuación, asigne un nombre al rol. ElIAMel rol tendrá el prefijo»AmazonKendra-».
5. Deje todos los demás campos con sus valores predeterminados. Elija Siguiente.
6. En elConfigurar el control de acceso de usuariospágina, eligeSiguiente.
7. En elDetalles de aprovisionamientoopágina, eligeEdición para desarrolladores.
8. EligeCrearpara crear tu índice.
9. Espere a que se cree el índice.Amazon Kendraaprovisiona el hardware para su índice. Esta operación puede llevar algún tiempo.

Paso 2: Para añadir una fuente de datos a un índice (consola)

1. Ver los disponibles[fuentes de datos](#)para conectarAmazon Kendrapara indexar sus documentos.
2. En el panel de navegación, seleccioneFuentes de datosy, a continuación, seleccionaAregar fuente de datospara la fuente de datos elegida.
3. Siga los pasos para configurar la fuente de datos.

Paso 3: Para buscar en un índice (consola)

1. En el panel de navegación, elija la opción para buscar en el índice.
2. Introduce un término de búsqueda que sea adecuado para tu índice. Elmejores resultadosydocumento principalse muestran los resultados.

Introducción (AWS CLI)

El siguiente procedimiento muestra cómo crear unAmazon Kendraíndice medianteAWS CLI. El procedimiento crea una fuente de datos, un índice y ejecuta una consulta en el índice.

Para crear unAmazon Kendraíndice (CLI)

1. Haz el[Requisitos previos \(p. 86\)](#).
2. Introduzca el siguiente comando para crear un índice.

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

3. Espera aAmazon Kendrapara crear el índice. Compruebe el progreso mediante el siguiente comando. Cuando el campo de estado esACTIVE, vaya al paso siguiente.

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

4. En la línea de comandos, introduzca el siguiente comando para crear una fuente de datos.

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

Si se conecta a la fuente de datos mediante un esquema de plantilla, configure el esquema de plantilla.

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

5. Se necesitará Amazon Kendra un tiempo para crear la fuente de datos. Introduzca el siguiente comando para comprobar el progreso. Cuando el estado esACTIVE, vaya al paso siguiente.

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

6. Introduzca el siguiente comando para sincronizar la fuente de datos.

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

7. Amazon Kendra indexará su fuente de datos. La cantidad de tiempo que lleve depende de la cantidad de documentos. Puede comprobar el estado del trabajo de sincronización mediante el siguiente comando. Cuando el estado esACTIVE, vaya al paso siguiente.

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

8. Introduzca el siguiente comando para realizar una consulta.

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

Los resultados de la búsqueda se muestran en formato JSON.

Introducción (AWS SDK for Python (Boto3))

El siguiente programa es un ejemplo del uso de Amazon Kendra en un programa de Python. El programa realiza las siguientes acciones:

1. Crea un índice nuevo mediante la operación [CreateIndex](#).
2. Espera a que se complete la creación del índice. Utiliza la operación [DescribeIndex](#) para supervisar el estado del índice.
3. Una vez que el índice está activo, crea una fuente de datos mediante la operación [CreateDataSource](#).
4. Espera a que se complete la creación de la fuente de datos. Utiliza la operación [DescribeDataSource](#) para supervisar el estado de la fuente de datos.
5. Cuando la fuente de datos está activa, sincroniza el índice con el contenido de la fuente de datos mediante la operación [StartDataSourceSyncJob](#).

```
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 description 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 = description,
    RoleArn = data_source_role_arn,
    Type = data_source_type,
    Configuration = configuration,
    IndexId = index_id
)

pprint.pprint(data_source_response)

data_source_id = data_source_response["Id"]

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

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

print("Synchronize the data source.")

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

pprint.pprint(sync_response)

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

while True:

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

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

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

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

print("Program ends.")
```

Introducción (AWS SDK for Java)

El siguiente programa es un ejemplo del uso de Amazon Kendra en un programa Java. El programa realiza las siguientes acciones:

1. Crea un índice nuevo mediante la operación [CreateIndex](#).
2. Espera a que se complete la creación del índice. Utiliza la operación [DescribeIndex](#) para supervisar el estado del índice.
3. Una vez que el índice está activo, crea una fuente de datos mediante la operación [CreateDataSource](#).
4. Espera a que se complete la creación de la fuente de datos. Utiliza la operación [DescribeDataSource](#) para supervisar el estado de la fuente de datos.
5. Cuando la fuente de datos está activa, sincroniza el índice con el contenido de la fuente de datos mediante la operación [StartDataSourceSyncJob](#).

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.DataSourceStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJob;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJobStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.S3DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;

public class CreateIndexAndDataSourceExample {

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

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

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

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

```
String indexId = createIndexResponse.id();

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

    TimeUnit.SECONDS.sleep(60);
}

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

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

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

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

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

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

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

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

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

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

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

        TimeUnit.SECONDS.sleep(60);
    }

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

Introducción a una fuente Amazon S3 de datos (consola)

Puede utilizar la Amazon Kendra consola para empezar a utilizar un Amazon S3 bucket como almacén de datos. Cuando usas la consola, especificas toda la información de conexión que necesitas para indexar el contenido del bucket. Para obtener más información, consulte [Amazon S3 \(p. 303\)](#).

Utilice el siguiente procedimiento para crear una fuente de datos básica de bucket de S3 mediante la configuración predeterminada. El procedimiento supone que ha creado un índice siguiendo los pasos del paso 1 de [Empezar con el Amazon Kendra consola \(p. 91\)](#).

Para crear una fuente de datos de bucket de S3 mediante la Amazon Kendra consola

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala en <https://console.aws.amazon.com/kendra/home>.
2. En la lista de índices, elija el índice al que desee añadir la fuente de datos.

3. Elija Agregar fuentes de datos.
4. En la lista de conectores de fuentes de datos, seleccione Amazon S3.
5. En la página Definir atributos, asigne un nombre a la fuente de datos y, si lo desea, una descripción. Deje el campo Etiquetas en blanco. Elija Next (Siguiente) para continuar.
6. En el campo Introduzca la ubicación de la fuente de datos, introduzca el nombre del bucket de S3 que contiene sus documentos. Puede introducir el nombre directamente o buscar el nombre seleccionando Examinar. El bucket debe estar en la misma región que el índice.
7. En IAMrol, elija Crear un rol nuevo y, a continuación, escriba un nombre de rol. Para obtener más información, consulte las [IAMfunciones de las fuentes Amazon S3 de datos](#).
8. En la sección Establecer la programación de ejecución de la sincronización, selecciona Ejecutar bajo demanda.
9. Elija Next (Siguiente) para continuar.
10. En la página Revisar y crear, revise los detalles de su fuente de datos de S3. Si desea realizar cambios, pulse el botón Editar situado junto al elemento que desee cambiar. Cuando esté satisfecho con sus opciones, elija Crear para crear su fuente de datos S3.

Después de elegir Crear, Amazon Kendra comienza a crear la fuente de datos. La creación de la fuente de datos puede tardar varios minutos. Cuando termine, el estado de la fuente de datos cambia de Creando a Activa.

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

Introducción a una fuente de datos de base de datos MySQL (consola)

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

Primero debe crear una base de datos MySQL y, a continuación, puede crear una fuente de datos para la base de datos.

Utilice el siguiente procedimiento para crear una base de datos MySQL básica. El procedimiento supone que ya ha creado un índice siguiendo el paso 1 de [Empezar con el Amazon Kendraconsola \(p. 91\)](#).

Para crear una base de datos MySQL

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon RDS en <https://console.aws.amazon.com/rds/>.
2. En el panel de navegación, elija Grupos de subredes y, a continuación, elija Crear grupo de subredes de base de datos.
3. Asigne un nombre al grupo y elija su nube privada virtual (VPC). Para obtener más información sobre la configuración de una VPC, consulte [Configuración Amazon Kendra para usar una VPC](#).
4. Agrega las subredes privadas de tu VPC. Las subredes privadas son las que no están conectadas a la NAT. Seleccione Crear.
5. En el panel de navegación, elija Bases de datos y, a continuación, elija Crear base de datos.
6. Utilice los siguientes parámetros para crear la base de datos. Deje todos los demás parámetros en sus valores predeterminados.

- Opciones del motor: MySQL
 - Plantillas: capa gratuita
 - Configuración de credenciales: introduzca y confirme una contraseña
 - En Conectividad, seleccione Configuración de conectividad adicional. Realice las siguientes elecciones.
 - Grupo de subredes: elija el grupo de subredes que creó en el paso 4.
 - Grupo de seguridad de VPC: elija el grupo que contenga las reglas entrantes y salientes que creó en la VPC. Por ejemplo, **DataSourceSecurityGroup**. Para obtener más información sobre la configuración de una VPC, consulte [Configuración Amazon Kendra para usar una VPC](#).
 - En Configuración adicional, defina el nombre de la base de datos inicial en**content**.
7. Elija Create database (Crear base de datos).
 8. En la lista de bases de datos, selecciona tu nueva base de datos. Tome nota del punto final de la base de datos.
 9. Después de crear la base de datos, debe crear una tabla para guardar los documentos. La creación de una tabla está fuera del alcance de estas instrucciones. Al crear la tabla, tenga en cuenta lo siguiente:
 - Nombre de base de datos— **content**
 - Nombre de la tabla— **documents**
 - Columnas: **ID**,**Title**,**Body**, y**LastUpdate**. Puede incluir columnas adicionales si lo desea.

Ahora que ha creado su base de datos MySQL, puede crear una fuente de datos para la base de datos.

Para crear una fuente de datos MySQL

1. Inicie sesión en la Amazon Kendra consola AWS Management Console y ábrala en <https://console.aws.amazon.com/kendra/home>.
2. En el panel de navegación, elija Índices y, a continuación, elija su índice.
3. Elija Agregar fuentes de datos y, a continuación, elija Amazon RDS.
4. Escriba un nombre y una descripción para la fuente de datos y, a continuación, elija Siguiente.
5. Elige MySQL.
6. En Acceso a la conexión, introduzca la siguiente información:
 - Punto final: el punto final de la base de datos que creó anteriormente.
 - Puerto: el número de puerto de la base de datos. Para MySQL, el valor predeterminado es 3306.
 - Tipo de autenticación: elija Nueva.
 - Nuevo nombre de contenedor secreto: un nombre para el Secrets Manager contenedor de las credenciales de la base de datos.
 - Nombre de usuario: el nombre de un usuario con acceso administrativo a la base de datos.
 - Contraseña: la contraseña del usuario y, a continuación, seleccione Guardar autenticación.
 - Nombre de base de datos —**content**.
 - Nombre de la tabla —**documents**.
 - Función de IAM: elija Crear una nueva función y, a continuación, escriba un nombre para la función.
7. En Configuración de columnas, introduzca lo siguiente:
 - Nombre de la columna de ID del documento — **ID**
 - Nombre de la columna del título del documento — **Title**
 - Nombre de la columna de datos del documento — **Body**
8. En Detección de cambios de columna, introduzca lo siguiente:

- Columnas de detección de cambios — **LastUpdate**
9. En el grupo Configurar VPC y seguridad, proporcione lo siguiente:
- En Virtual Private Cloud (VPC), elige tu VPC.
 - En Subredes, elige las subredes privadas que creaste en tu VPC.
 - En los grupos de seguridad de VPC, elija el grupo de seguridad que contenga las reglas entrantes y salientes que creó en su VPC para bases de datos MySQL. Por ejemplo, **DataSourceSecurityGroup**.
10. En Definir el horario de ejecución de la sincronización, selecciona Ejecutar bajo demanda y, a continuación, selecciona Siguiente.
11. En Mapeo de campos de fuente de datos, elija Siguiente.
12. Revise la configuración de la fuente de datos para asegurarse de que es correcta. Cuando estés seguro de que todo está correcto, selecciona Crear.

Empezar con unAWS IAM Identity Center (successor to AWS Single Sign-On)fuente de identidad (consola)

UnAWS IAM Identity Center (successor to AWS Single Sign-On)la fuente de identidad contiene información sobre sus usuarios y grupos. Esto es útil para configurar el filtrado de contexto de usuario, dondeAmazon Kendra filtra los resultados de búsqueda para diferentes usuarios en función del acceso del usuario o de su grupo a los documentos.

Para crear una fuente de identidad de IAM Identity Center, debe activar IAM Identity Center y crear una organización enAWS Organizations. Al activar IAM Identity Center y crear una organización por primera vez, se establece automáticamente el directorio del Centro de identidades como fuente de identidad. Puede cambiar a Active Directory (gestionado o autogestionado por Amazon) o a un proveedor de identidad externo como fuente de identidad. Para ello, debe seguir las instrucciones correctas; consulte [Cambiar la fuente de identidad de IAM Identity Center](#). Solo puede tener una fuente de identidad por organización.

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

También debe conceder los permisos necesarios para usar IAM Identity Center conAmazon Kendra. Para obtener más información, consulte [IAMfunciones de IAM Identity Center](#).

Para configurar una fuente de identidad de IAM Identity Center

1. Abra el[Consola de IAM Identity Center](#).
2. EscojaHabilitar IAM Identity Centery, a continuación, elijaCrearAWSorganización.

El directorio de Identity Center se crea de forma predeterminada y se le envía un correo electrónico para verificar la dirección de correo electrónico asociada a la organización.

3. Para añadir un grupo a tuAWSorganización, en el panel de navegación, elijaGrupos.
4. En elPágina de grupos, eligeCrear grupo e introduzca un nombre y una descripción del grupo en el cuadro de diálogo. Seleccione Crear.

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

El directorio Identity Center es la fuente de identidad predeterminada y requiere que agregue manualmente sus usuarios y grupos mediante esta fuente si no tiene su propia lista administrada por un proveedor. Para cambiar la fuente de identidad, debe seguir las instrucciones correctas para ello: consulte [Cambiar la fuente de identidad de IAM Identity Center](#).

Note

Si utiliza Active Directory o un proveedor de identidad externo como fuente de identidad, debe asignar las direcciones de correo electrónico de sus usuarios a los nombres de usuario de IAM Identity Center al especificar el protocolo System for Cross-Domain Identity Management (SCIM). Para obtener más información, consulte la [Guía de IAM Identity Center sobre SCIM para habilitar IAM Identity Center](#).

Una vez que haya configurado su fuente de identidad de IAM Identity Center, podrá activarla en la consola al crear o editar el índice. Ir a Control de acceso de usuarios en la configuración del índice y edítela para permitir obtener información de grupos de usuarios desde IAM Identity Center. También puede activar IAM Identity Center mediante el [UserGroupResolutionConfiguration](#) objeto. Usted proporciona el `UserGroupResolutionMode` en `AWS_SSO` y crea un IAMrol que da permiso para llamar a `ListDirectoryAssociations`, `sso-directory:SearchUsers`, `sso-directory>ListGroupsForUser`, `sso-directory:DescribeGroups`.

Warning

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

A continuación se muestra una descripción general de cómo configurar una fuente de datos con `UserGroupResolutionConfiguration` y control de acceso de usuario para filtrar los resultados de búsqueda según el contexto del usuario. Esto supone que ya ha creado un índice y un IAM función para los índices. Crea un índice y proporciona el IAMrol mediante el [CreateIndex API](#).

Configuración de una fuente de datos con `UserGroupResolutionConfiguration` y filtrado de contexto de usuario

1. Crea un IAM papel que le da permiso para acceder a su fuente de identidad de IAM Identity Center.
2. Configurar `UserGroupResolutionConfiguration` configurando el modo en `AWS_SSO` y llama `UpdateIndex` para actualizar su índice para usar IAM Identity Center.
3. Si desea utilizar el control de acceso de usuario basado en tokens para filtrar los resultados de búsqueda según el contexto del usuario, defina `UserContextPolicy` a `USER_TOKEN` cuando llamas `UpdateIndex`. De lo contrario, Amazon Kendra rastrea la lista de control de acceso de cada uno de los documentos para la mayoría de los conectores de fuentes de datos. También puede filtrar los resultados de la búsqueda según el contexto del usuario en `Consulta API` al proporcionar información de usuarios y grupos en `UserContext`. También puede asignar usuarios a sus grupos

mediante[PutPrincipalMapping](#)de modo que solo tendrá que proporcionar el ID de usuario cuando emita la consulta.

4. Crea un[IAMpapelque da permiso para acceder a su fuente de datos.](#)
5. [Configurarsu fuente de datos. Debe proporcionar la información de conexión necesaria para conectarse a la fuente de datos.](#)
6. Cree una fuente de datos mediante el[CreateDataSourceAPI](#). Proporcione elDataSourceConfigurationobjeto, que incluyeTemplateConfiguration, el identificador de su índice, elIAMrol para la fuente de datos, el tipo de fuente de datos y asigne un nombre a la fuente de datos. También puede actualizar la fuente de datos.

Cambiar la fuente de identidad de IAM Identity Center

Warning

Cambiar la fuente de identidad en IAM Identity CenterAjustespodría afectar a la conservación de la información del usuario y del grupo. Para hacerlo de forma segura, se recomienda que revise[Consideraciones para cambiar la fuente de identidad](#). Cuando cambias la fuente de identidad, se genera una nueva ID de fuente de identidad. Compruebe que está utilizando el identificador correcto antes de configurar el modo enAWS_SSOn[UserGroupResolutionConfiguration](#).

Para cambiar la fuente de identidad de IAM Identity Center

1. Abra el[IAM Identity Center> consola.](#)
2. Elija Settings.
3. En elAjustespágina, enFuente de identidad, eligeCambiar.
4. En elCambiar fuente de identidadpágina, selecciona tu fuente de identidad preferida y, a continuación, eligeSiguiente.

Creación de un índice

Puede crear un índice mediante la consola o llamando a la [CreateIndexAPI](#). Puedes usar el AWS Command Line Interface (AWS CLI) o el SDK con la API. En este capítulo se describe cómo crear un índice utilizando cualquiera de estos métodos. Después de crear el índice, puede agregar documentos directamente a él o desde un origen de datos.

Para crear un índice, debe proporcionar el nombre de recurso de Amazon (ARN) de un rol AWS Identity and Access Management (IAM) para que los índices puedan acceder CloudWatch a él. Para obtener más información, consulte las [IAMfunciones de los índices](#).

Para crear un índice (consola)

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

Para crear un índice (AWS CLI).

1. Utilice el siguiente comando para crear un índice. El role-arn debe ser el nombre de recurso de Amazon (ARN) de un rol de IAM que pueda ejecutar acciones de Amazon Kendra. Para obtener más información, consulte [IAMroles](#).

El comando está formateado para Linux y macOS. Si utiliza Windows, sustituya el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

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

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

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

Para crea un índice (SDK).

1. Proporcione valores para las siguientes variables:

- `description`: una descripción del índice que está creando. Es opcional.
- `index_name`: el nombre del índice que está creando.
- `role_arn`—El nombre del recurso de Amazon (ARN) de un rol que puede ejecutar las Amazon Kendra API. Para obtener más información, consulte [IAMroles](#).

2. En los siguientes ejemplos, se crea un Amazon Kendra índice.

Python

```
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

```
package com.amazonaws.kendra;
```

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

public class CreateIndexExample {

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

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

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

        String indexId = createIndexResponse.id();

        System.out.println(String.format("Waiting until the index with ID %s is
created.", indexId));
        while (true) {
            DescribeIndexRequest describeIndexRequest =
DescribeIndexRequest.builder().id(indexId).build();
            DescribeIndexResponse describeIndexResponse =
kendra.describeIndex(describeIndexRequest);
            IndexStatus status = describeIndexResponse.status();
            if (status != IndexStatus.CREATING) {
                break;
            }
            TimeUnit.SECONDS.sleep(60);
        }
        System.out.println("Index creation is complete.");
    }
}
```

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

Temas

- [Añadir documentos directamente a un índice con carga por lotes \(p. 107\)](#)
- [Añadir preguntas frecuentes \(FAQs\) a un índice \(p. 111\)](#)
- [Creación de campos de documentos personalizados \(p. 116\)](#)
- [Controlar el acceso de los usuarios a los documentos con identificadores \(p. 117\)](#)

Añadir documentos directamente a un índice con carga por lotes

Puede agregar documentos directamente a un índice mediante la API [BatchPutDocument](#). No puede añadir documentos directamente con la consola. Si usa la consola, se conecta a una fuente de datos para añadir documentos al índice.

Los documentos se pueden agregar desde un bucket de S3 o suministrarse como datos binarios.

Para obtener una lista de los tipos de documentos admitidos por, Amazon Kendra consulte [Tipos de documentos](#).

Agregar documentos a un índice mediante BatchPutDocument una operación asíncrona. Después de llamar a la API BatchPutDocument, utilice la API [BatchGetDocumentStatus](#) para supervisar el progreso de la indexación de los documentos. Cuando se llama a la API BatchGetDocumentStatus con una lista de identificadores de documento, devuelve el estado del documento. Cuando el estado del documento sea INDEXED o FAILED, se habrá completado el procesamiento del documento. Cuando el estado sea FAILED, la API BatchGetDocumentStatus devolverá el motivo por el que el documento no se haya podido indexar.

Si desea modificar los campos o atributos de los metadatos del contenido y del documento durante el proceso de ingesta del documento, consulte [Enriquecimiento Amazon Kendra personalizado de documentos](#).

Si quieras usar una fuente de datos personalizada, cada documento que envíes mediante la BatchPutDocument API requiere un ID de fuente de datos y un ID de ejecución como atributos o campos. Para obtener más información, consulte [Atributos obligatorios para orígenes de datos personalizados](#).

Tenga en cuenta que cada ID de documento debe ser único por índice. No se puede crear un origen de datos para indexar los documentos con sus ID exclusivos y, a continuación, utilizar la API BatchPutDocument para indexar los mismos documentos o viceversa. No se puede crear un origen de datos para indexar los documentos con sus ID exclusivos y, a continuación, utilizar la API BatchPutDocument para indexar los mismos documentos o viceversa.

Los siguientes documentos de la guía para desarrolladores muestran cómo añadir documentos directamente a un índice.

Temas

- [Añadir documentos con la BatchPutDocument API \(p. 107\)](#)
- [Añadir documentos desde un bucket de S3 \(p. 109\)](#)

Añadir documentos con la BatchPutDocument API

En el siguiente ejemplo, se agrega una masa de texto a un índice mediante una llamada [BatchPutDocument](#). También puedes [añadir documentos desde un bucket](#) de S3 y llamar a la BatchPutDocument API.

Puedes usar la BatchPutDocument API para añadir documentos a tu índice. Para obtener una lista de los tipos de documentos admitidos por, Amazon Kendra consulte [Tipos de documentos](#).

Los archivos añadidos al índice deben estar en un flujo de bytes codificado en UTF-8.

Para ver un ejemplo de cómo crear un índice mediante los SDK AWS CLI y, consulta [Crear un índice](#). Para configurar la CLI y los SDK, consulte [Configuración](#). Amazon Kendra

En los ejemplos siguientes, se añade al índice texto con codificación UTF-8.

CLI

En la AWS Command Line Interface utilice el siguiente comando. El comando está formateado para Linux y macOS. Si utiliza Windows, sustituya el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

```
aws kendra batch-put-document \
    --index-id index-id \
    --documents '[{"Id":"doc-id-1", "Blob":"Amazon.com is an online retailer.", "ContentType":"PLAIN_TEXT", "Title";"Information about Amazon.com"}'
```

Python

```
import boto3

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"

# Provide the title and text
title = "Information about Amazon.com"
text = "Amazon.com is an online retailer."

document = {
    "Id": "1",
    "Blob": text,
    "ContentType": "PLAIN_TEXT",
    "Title": title
}

documents = [
    document
]

result = kendra.batch_put_document(
    IndexId = index_id,
    Documents = documents
)

print(result)
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.core.SdkBytes;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentRequest;
import software.amazon.awssdk.services.kendra.model.BatchPutDocumentResponse;
import software.amazon.awssdk.services.kendra.model.ContentType;
import software.amazon.awssdk.services.kendra.model.Document;

public class AddDocumentsViaAPIExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indexId = "yourIndexId";

        Document testDoc = Document
            .builder()
            .title("The title of your document")
```

```
.id("a_doc_id")
.blob(SdkBytes.fromUtf8String("your text content"))
.contentType(ContentType.PLAIN_TEXT)
.build();

BatchPutDocumentRequest batchPutDocumentRequest = BatchPutDocumentRequest
.builder()
.indexId(indexId)
.documents(testDoc)
.build();

BatchPutDocumentResponse result =
kendra.batchPutDocument(batchPutDocumentRequest);

System.out.println(String.format("BatchPutDocument Result: %s", result));
}
```

Añadir documentos desde un bucket de S3

Puede añadir documentos directamente a su índice desde un Amazon S3 depósito. Puede añadir hasta 10 documentos en la misma llamada. Cuando utilice un bucket de S3, debe proporcionar un rol de IAM con permiso para acceder al bucket que contenga los documentos. Especifique el rol en el parámetro `RoleArn`.

El uso de la [BatchPutDocument](#)API para añadir documentos desde un Amazon S3 bucket es una operación que se realiza una sola vez. Para mantener un índice sincronizado con el contenido de un bucket, cree una fuente de Amazon S3 datos. Para obtener más información, consulte [fuente Amazon S3 de datos](#).

Para ver un ejemplo de cómo crear un índice mediante los SDK AWS CLI y, consulta [Crear un índice](#). Para configurar la CLI y los SDK, consulte [Configuración](#). Amazon Kendra Para obtener información sobre la creación de un bucket de S3, consulte [Amazon Simple Storage Service documentación](#).

En el siguiente ejemplo, se agregan dos documentos de Microsoft Word al índice mediante la API `BatchPutDocument`.

Python

```
import boto3

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the IAM role ARN required to index documents in an S3 bucket
role_arn = "arn:aws:iam::${accountID}:policy/${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));
}
```

Añadir preguntas frecuentes (FAQs) a un índice

Puede agregar preguntas frecuentes directamente al índice mediante la consola o la API [CreateFaq](#). La adición de preguntas frecuentes a un índice es asíncrona. Coloque los datos de las preguntas frecuentes en un archivo almacenado en un bucket de Amazon Simple Storage Service. Puede utilizar archivos CSV o JSON como entrada para las preguntas frecuentes:

- CSV básico: un archivo CSV en el que cada fila contiene una pregunta, una respuesta y un URI de origen opcional.
- CSV personalizado: un archivo CSV que contiene preguntas, respuestas y encabezados para campos o atributos personalizados que puedes usar para facetar, mostrar u ordenar las respuestas a las preguntas frecuentes. También puede definir campos de control de acceso para limitar la respuesta a las preguntas frecuentes a ciertos usuarios y grupos que pueden ver la respuesta a las preguntas frecuentes.
- JSON: un archivo JSON que contiene preguntas, respuestas y campos/atributos personalizados que puedes usar para facetar, mostrar u ordenar las respuestas a las preguntas frecuentes. También puede definir campos de control de acceso para limitar la respuesta a las preguntas frecuentes a ciertos usuarios y grupos que pueden ver la respuesta a las preguntas frecuentes.

Por ejemplo, el siguiente es un archivo CSV básico que proporciona respuestas a preguntas sobre clínicas gratuitas en Spokane, Washington, EE. UU. y Mountain View, Missouri, EE. UU.

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

Cuando utilizas un archivo CSV o JSON personalizado para la entrada, puedes declarar campos personalizados para tus preguntas frecuentes. Por ejemplo, puedes crear un campo personalizado que asigne a cada pregunta de preguntas frecuentes un departamento comercial. Cuando las preguntas frecuentes aparezcan en una respuesta, puedes utilizar el departamento como una faceta para limitar la búsqueda únicamente a «Recursos Humanos» o «Finanzas», por ejemplo.

Un campo personalizado debe asignarse a un campo de índice. En la consola, utilice la página [Facet definition \(Definición de faceta\)](#) para crear un campo de índice. Al utilizar la API, primero debe crear un campo de índice mediante la API [UpdateIndex](#).

El tipo de campo/atributo del archivo de preguntas frecuentes debe coincidir con el tipo del campo de índice asociado. Por ejemplo, el campo «Departamento» es un campo STRING_LIST de tipo. Por lo tanto, debe proporcionar valores para el campo del departamento como una lista de cadenas en su archivo de preguntas frecuentes. Puede comprobar el tipo de campos de índice utilizando la página [Facet Definition \(Definición de faceta\)](#) en la consola o mediante la API [DescribeIndex](#).

Cuando se crea un campo de índice que se asigna a un atributo personalizado, se puede marcar como visualizable, facetable u ordenable. No se puede hacer que un atributo personalizado se pueda buscar.

Además de los atributos personalizados, también puedes usar los campos Amazon Kendra reservados o comunes de un archivo CSV o JSON personalizado. Para obtener más información, consulte [Atributos o campos del documento](#).

Archivo CSV básico

Utilice un archivo CSV básico cuando desee utilizar una estructura sencilla para sus preguntas frecuentes. En un archivo CSV básico, cada fila tiene dos o tres campos: una pregunta, una respuesta y un URI de origen opcional que apunta a un documento con más información.

El contenido del archivo debe seguir el [Formato común RFC 4180 y tipo MIME para archivos de valores separados por comas \(CSV\)](#).

A continuación se muestra un archivo de preguntas frecuentes en formato CSV básico.

```
How many free clinics are in Spokane WA?, 13, https://s3.region.company.com/bucket-name/
directory/faq.csv
How many free clinics are there in Mountain View Missouri?, 7, https://
s3.region.company.com/bucket-name/directory/faq.csv
```

Archivo CSV personalizado

Usa un archivo CSV personalizado cuando quieras añadir campos o atributos personalizados a tus preguntas frecuentes. Para un archivo CSV personalizado, utilice una fila de encabezado del archivo CSV para definir los atributos adicionales.

El archivo CSV debe contener los dos campos obligatorios siguientes:

- `_question`—La pregunta más frecuente
- `_answer`—La respuesta a la pregunta más frecuente

El archivo puede contener tanto campos Amazon Kendra reservados como campos personalizados. A continuación se muestra un ejemplo de un archivo CSV personalizado.

```
_question,_answer,_last_updated_at,custom_string
How many free clinics are in Spokane WA?, 13, 2012-03-25T12:30:10+01:00, Note: Some free
clinics require you to meet certain criteria in order to use their services
How many free clinics are there in Mountain View Missouri?, 7, 2012-03-25T12:30:10+01:00,
Note: Some free clinics require you to meet certain criteria in order to use their
services
```

El contenido del archivo personalizado debe seguir el [formato común y el tipo MIME de la RFC 4180 para archivos de valores separados por comas \(CSV\)](#).

A continuación se enumeran los tipos de campos personalizados:

- Fecha: valores de fecha y hora codificados en ISO 8601.

Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en la zona horaria de Europa Central.

- Largo: números, como 1234
- Cadena: valores de cadena. Si la cadena contiene comas, incluya todo el valor entre comillas dobles ("") (por ejemplo, "custom attribute, and more").
- Lista de cadenas: una lista de valores de cadena. Enumere los valores de una lista separada por comas incluidos entre comillas ("") (por ejemplo, "item1, item2, item3"). Si la lista contiene solo una entrada, puede omitir las comillas (por ejemplo, item1).

Un archivo CSV personalizado puede contener campos de control de acceso de usuario. Puede utilizar estos campos para limitar el acceso a las preguntas frecuentes a determinados usuarios y grupos. Para

filtrar el contexto de usuario, el usuario debe proporcionar en la consulta información de usuario y grupo. De lo contrario, se devuelven todas las preguntas frecuentes pertinentes. Para obtener más información, consulte [Filtrado de contextos de usuario](#).

A continuación se enumeran los filtros de contexto de usuario para las preguntas frecuentes:

- `_acl_user_allow`—Los usuarios de la lista de permitidos pueden ver las preguntas frecuentes en la respuesta a la consulta. Las preguntas frecuentes no se devuelven a otros usuarios.
- `_acl_user_deny`—Los usuarios de la lista de denegaciones no pueden ver las preguntas frecuentes en la respuesta a la consulta. Las preguntas frecuentes se devuelven a todos los demás usuarios cuando son relevantes para la consulta.
- `_acl_group_allow`—Los usuarios que son miembros de un grupo permitido pueden ver las preguntas frecuentes en la respuesta a la consulta. Las preguntas frecuentes no se devuelven a los usuarios que sean miembros de otro grupo.
- `_acl_group_deny`—Los usuarios que son miembros de un grupo denegado no pueden ver las preguntas frecuentes en la respuesta a la consulta. Las preguntas frecuentes se devuelven a todos los demás grupos cuando sea relevante para la consulta.

Proporcione los valores de las listas de permitir y denegar listas separadas por comas entre comillas (por ejemplo, `"user1,user2,user3"`). Puede incluir un usuario o un grupo en una lista de permitidos o en una lista de denegaciones, pero no en ambas en las que se permita al mismo usuario de forma individual, sino también en grupo. Si incluye un usuario o grupo en ambos, recibirá un error.

A continuación se muestra un ejemplo de un archivo CSV personalizado.

```
_question, _answer, _acl_user_allow, _acl_user_deny, _acl_group_allow, _acl_group_deny
How many free clinics are in Spokane WA?, 13, "userID6201(userID7552",
"userID1001(userID2020", groupBasicPlusRate, groupPremiumRate
```

Archivo JSON

Puedes usar un archivo JSON para proporcionar preguntas, respuestas y campos para tu índice. Puede añadir cualquiera de los campos Amazon Kendra reservados o campos personalizados a las preguntas frecuentes.

El siguiente es el esquema del archivo JSON.

```
{
  "SchemaVersion": 1,
  "FaqDocuments": [
    {
      "Question": string,
      "Answer": string,
      "Attributes": {
        string: object
        additional attributes
      },
      "AccessControlList": [
        {
          "Name": string,
          "Type": enum( "GROUP" | "USER" ),
          "Access": enum( "ALLOW" | "DENY" )
        },
        additional user context
      ]
    },
    additional FAQ documents
  ]
}
```

}

El siguiente archivo JSON de ejemplo muestra dos documentos de preguntas frecuentes. Uno de los documentos contiene únicamente la pregunta y la respuesta requeridas. El otro documento también incluye información adicional sobre el campo y el contexto del usuario o el control de acceso.

```
{  
    "SchemaVersion": 1,  
    "FaqDocuments": [  
        {  
            "Question": "How many free clinics are in Spokane WA?",  
            "Answer": "13"  
        },  
        {  
            "Question": "How many free clinics are there in Mountain View Missouri?",  
            "Answer": "7",  
            "Attributes": {  
                "_source_uri": "https://s3.region.company.com/bucket-name/directory/  
faq.csv",  
                "_category": "Charitable Clinics"  
            },  
            "AccessControlList": [  
                {  
                    "Name": "user@amazon.com",  
                    "Type": "USER",  
                    "Access": "ALLOW"  
                },  
                {  
                    "Name": "Admin",  
                    "Type": "GROUP",  
                    "Access": "ALLOW"  
                }  
            ]  
        }  
    ]  
}
```

A continuación se enumeran los tipos de campos personalizados:

- Fecha: un valor de cadena JSON con valores de fecha y hora codificados en ISO 8601. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en la zona horaria de Europa Central.
- Largo: un valor numérico JSON, por ejemplo1234.
- Cadena: un valor de cadena JSON (por ejemplo, "custom attribute").
- Lista de cadenas: una matriz JSON de valores de cadena (por ejemplo,["item1,item2,item3"]).

Un archivo CSV personalizado puede contener campos de control de acceso de usuario. Puede utilizar estos campos para limitar el acceso a las preguntas frecuentes a determinados usuarios y grupos. Para filtrar el contexto de usuario, el usuario debe proporcionar en la consulta información de usuario y grupo. De lo contrario, se devuelven todas las preguntas frecuentes pertinentes. Para obtener más información, consulte [Filtrado de contextos de usuario](#).

Puede incluir un usuario o un grupo en una lista de permitidos o en una lista de denegaciones, pero no en ambas en las que se permita al mismo usuario de forma individual, sino también en grupo. Si incluye un usuario o grupo en ambos, recibirá un error.

El siguiente es un ejemplo de cómo incluir el control de acceso de usuarios en las preguntas frecuentes de JSON.

```
"AccessControlList": [
```

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

Uso del archivo de preguntas frecuentes

Después de almacenar el archivo de entrada de preguntas frecuentes en un bucket de S3, utilice la consola o la API CreateFaq para incluir las preguntas y respuestas en el índice. Si quiere actualizar una pregunta frecuente, elimínela y vuelva a crearla. Utilice la API DeleteFaq para eliminar una pregunta frecuente.

Debe proporcionar un rol de IAM con acceso al bucket de S3 que contenga los archivos de origen. Puede especificar el rol en la consola o en el parámetro RoleArn. El siguiente es un ejemplo de cómo agregar un archivo de preguntas frecuentes a un índice.

Python

```
import boto3  
  
kendra = boto3.client("kendra")  
  
# Provide the index ID  
index_id = "index-id"  
# Provide the IAM role ARN required to index documents in an S3 bucket  
role_arn = "arn:aws:iam::${accountId}:role/${roleName}"  
  
# Provide the S3 bucket path information to the FAQ file  
faq_path = {  
    "Bucket": "bucket-name",  
    "Key": "FreeClinicsUSA.csv"  
}  
  
response = kendra.create_faq(  
    S3Path = faq_path,  
    Name = "FreeClinicsUSA",  
    IndexId = index_id,  
    RoleArn = role_arn  
)  
  
print(response)
```

Java

```
package com.amazonaws.kendra;  
  
import software.amazon.awssdk.services.kendra.KendraClient;  
import software.amazon.awssdk.services.kendra.model.CreateFaqRequest;  
import software.amazon.awssdk.services.kendra.model.CreateFaqResponse;  
import software.amazon.awssdk.services.kendra.model.S3Path;  
  
public class AddFaqExample {  
    public static void main(String[] args) {  
        KendraClient kendra = KendraClient.builder().build();  
  
        String indexId = "yourIndexId";  
        String roleArn = "your role for accessing S3 files";
```

```
CreateFaqRequest createFaqRequest = CreateFaqRequest
    .builder()
    .indexId(indexId)
    .name("FreeClinicsUSA")
    .roleArn(roleArn)
    .s3Path(
        S3Path
            .builder()
            .bucket("an-aws-kendra-test-bucket")
            .key("FreeClinicsUSA.csv")
            .build())
    .build();

CreateFaqResponse response = kendra.createFaq(createFaqRequest);

System.out.println(String.format("The result of creating FAQ: %s", response));

}
```

Archivos de preguntas frecuentes en idiomas distintos del inglés

Puede indexar las preguntas frecuentes en un idioma admitido. Amazon Kendra indexa las preguntas frecuentes en inglés de forma predeterminada si no especifica un idioma. Especifica el código de idioma al llamar a la [CreateFaq](#)operación o puede incluir el código de idioma de las preguntas frecuentes en los metadatos de las preguntas frecuentes como campo. Si una pregunta frecuente no contiene un código de idioma en sus metadatos especificado en un campo de metadatos, las preguntas frecuentes se indexan utilizando el código de idioma especificado al llamar a la operación CreateFAQ. Para indexar un documento de preguntas frecuentes en un idioma admitido en la consola, vaya a FAQs (Preguntas frecuentes) y seleccione Add FAQ (Añadir pregunta frecuente). Elija un idioma en el menú desplegable Language (Idioma).

Creación de campos de documentos personalizados

Puede aplicar atributos o campos personalizados para sus documentos específicos. Por ejemplo, puede crear un campo o atributo personalizado denominado «Departamento» con los valores de «Recursos Humanos», «Ventas» y «Fabricación». Puede utilizar estos campos o atributos para filtrar los resultados de la búsqueda a documentos del departamento de «Recursos Humanos», por ejemplo.

Puede crear hasta 500 campos o atributos personalizados.

Para la mayoría de las fuentes de datos, los campos de la fuente de datos externa se asignan a los campos correspondientes deAmazon Kendra. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos). Para las fuentes de datos de S3, puede crear campos o atributos personalizados mediante un archivo de metadatos JSON.

Antes de poder utilizar un campo o atributo personalizado, primero debe crear el campo en el índice. Utilice la consola para editar las asignaciones de campos de la fuente de datos para añadir un campo personalizado o utilice la [UpdateIndex](#)API para crear el campo de índice. No puede cambiar el tipo de datos del campo una vez creado el campo.

También puede utilizar Amazon Kendra los campos reservados o comunes. Para obtener más información, consulte [Atributos o campos del documento](#).

Con la `UpdateIndex` API, puedes añadir campos o atributos personalizados mediante el `DocumentMetadataConfigurationUpdates` parámetro.

En el siguiente ejemplo de JSON se utiliza `DocumentMetadataConfigurationUpdates` para agregar al índice un campo denominado «Department».

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

Añadir atributos o campos personalizados con la BatchPutDocument API

Cuando utilizas la [BatchPutDocument](#) API para añadir un documento a tu índice, especificas campos o atributos personalizados como parte de `Attributes`. Puedes añadir varios campos o atributos cuando llamas a la API. Puede crear hasta 500 campos o atributos personalizados. El siguiente ejemplo es un campo o atributo personalizado que agrega «Departamento» a un documento.

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

Agregar atributos o campos personalizados a una fuente Amazon S3 de datos

Cuando se utilice un bucket de S3 como origen de datos para el índice, se agregan metadatos a los documentos con archivos de metadatos complementarios. Los archivos JSON de metadatos se colocan en una estructura de directorios paralela a los documentos. Para obtener más información, consulte [Metadatos de documentos de S3](#).

Los campos o atributos personalizados se especifican en la estructura `Attributes` JSON. Puede crear hasta 500 campos o atributos personalizados. Por ejemplo, el siguiente ejemplo se utiliza `Attributes` para definir tres campos o atributos personalizados y un campo reservado.

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

Controlar el acceso de los usuarios a los documentos con identificadores

Puede controlar qué usuarios o grupos pueden acceder a ciertos documentos de su índice o ver determinados documentos en sus resultados de búsqueda. Esto se denomina filtrado de contexto de

usuario. Es un tipo de búsqueda personalizada con la ventaja de controlar el acceso a los documentos. Por ejemplo, no todos los equipos que buscan información en el portal de la empresa deben acceder a documentos de alto secreto de la empresa, ni estos documentos son relevantes para todos los usuarios. Solo los usuarios o grupos de equipos específicos que tengan acceso a documentos de alto secreto deberían ver estos documentos en los resultados de búsqueda.

Amazon Kendra admite el control de acceso de usuarios basado en tokens mediante los siguientes tipos de tokens:

- Abrir ID
- JWT con un secreto compartido
- JWT con clave pública
- JSON

Amazon Kendra ofrece una búsqueda empresarial altamente segura para sus aplicaciones de búsqueda. Los resultados de la búsqueda reflejan el modelo de seguridad de su organización. Los clientes son responsables de autenticar y autorizar a los usuarios a acceder a su aplicación de búsqueda. En el momento de la búsqueda, el Amazon Kendra servicio filtra los resultados de la búsqueda en función del identificador de usuario proporcionado por la aplicación de búsqueda del cliente y las listas de control de acceso a los documentos (ACL) recopiladas por el Amazon Kendra conectores durante el tiempo de rastreo/indexación. Los resultados de la búsqueda muestran URL que apuntan a los repositorios de documentos originales, además de breves extractos. El repositorio original sigue exigiendo el acceso al documento completo.

Temas

- [Uso de OpenID \(p. 118\)](#)
- [Uso de un token web JSON \(JWT\) con un secreto compartido \(p. 120\)](#)
- [Uso de un token web JSON \(JWT\) con una clave pública \(p. 122\)](#)
- [Uso de JSON \(p. 124\)](#)

Uso de OpenID

Para configurar un Amazon Kendra índice para usar un token de OpenID para el control de acceso, necesita la URL JWKS (conjunto de claves web JSON) del proveedor de OpenID. En la mayoría de los casos, la URL de JWKS tiene el siguiente formato (si siguen el descubrimiento de OpenID). <https://domain-name/.well-known/jwks.json>

Los siguientes ejemplos muestran cómo utilizar un token de OpenID para el control de acceso de los usuarios al crear un índice.

Console

1. Elija Crear índice para empezar a crear un índice nuevo.
2. En la página Especificar los detalles del índice, asigne un nombre y una descripción al índice.
3. Para IAM rol, seleccione un rol o seleccione Crear un nuevo rol para y especifique un nombre de rol para crear un rol nuevo. El rol de IAM tendrá el prefijo "AmazonKendra->".
4. Deje todos los demás campos con sus valores predeterminados. Elija Siguiente.
5. En la página Configurar el control de acceso de los usuarios, en Configuración del control de acceso, elija Sí para usar los identificadores para el control de acceso.
6. En Configuración del token, seleccione OpenID como tipo de token.
7. Especifique una URL de clave de firma. La URL debe apuntar a un conjunto de claves web JSON.

8. Opcional en Configuración avanzada:
 - a. Especifique un nombre de usuario para usarlo en la comprobación de ACL.
 - b. Especifique uno o más grupos para usarlos en la comprobación de ACL.
 - c. Especifique el emisor que validará el emisor del token.
 - d. Especifique los identificadores de cliente. Debe especificar una expresión regular que coincida con la audiencia del JWT.
9. En la página de detalles del aprovisionamiento, seleccione Edición para desarrolladores.
10. Elige Crear para crear tu índice.
11. Espere a que se cree el índice. Amazon Kendraaprovisiona el hardware para su índice. Esta operación puede llevar algún tiempo.

CLI

Para crear un índice con el AWS CLI uso de un archivo de entrada JSON, primero cree un archivo JSON con los parámetros que deseé:

```
{  
    "Name": "user-context",  
    "Edition": "ENTERPRISE_EDITION",  
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",  
    "UserTokenConfigurations": [  
        {  
            "JwtTokenTypeConfiguration": {  
                "KeyLocation": "URL",  
                "Issuer": "optional: specify the issuer url",  
                "ClaimRegex": "optional: regex to validate claims in the token",  
                "UserNameAttributeField": "optional: user",  
                "GroupAttributeField": "optional: group",  
                "URL": "https://example.com/.well-known/jwks.json"  
            }  
        }  
    ],  
    "UserContextPolicy": "USER_TOKEN"  
}
```

Puede anular los nombres de campo de usuario y grupo predeterminados. El valor predeterminado para UserNameAttributeField es «user». El valor predeterminado para GroupAttributeField es «groups».

A continuación, llame `create-index` con el archivo de entrada. Por ejemplo, si el nombre del archivo JSON es `create-index-openid.json`, puede utilizar lo siguiente:

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

Python

```
response = kendra.create_index(  
    Name='user-context',  
    Edition='ENTERPRISE_EDITION',  
    RoleArn='arn:aws:iam::account-id:role:/my-role',  
    UserTokenConfigurations=[  
        {  
            "JwtTokenTypeConfiguration": {  
                "KeyLocation": "URL",  
                "Issuer": "optional: specify the issuer url",  
                "ClaimRegex": "optional: regex to validate claims in the token",  
            }  
        }  
    ],  
    UserContextPolicy="USER_TOKEN")
```

```
        "UserNameAttributeField": "optional: user",
        "GroupAttributeField": "optional: group",
        "URL": "https://example.com/.well-known/jwks.json"
    }
],
UserContextPolicy='USER_TOKEN'
)
```

Uso de un token web JSON (JWT) con un secreto compartido

Los siguientes ejemplos muestran cómo utilizar un token web JSON (JWT) con un token secreto compartido para controlar el acceso de los usuarios al crear un índice.

Console

1. Escoja **Crear índice** para empezar a crear un índice nuevo.
2. En el **Especificación de detalles del índice** página, asigne un nombre y una descripción a su índice.
3. Para **Función de IAM**, seleccione un rol o seleccione **Crear un nuevo rol** y especifique un nombre de rol para crear un rol nuevo. El IAM rol tendrá el prefijo »AmazonKendra».
4. Deje todos los demás campos con sus valores predeterminados. Elija **Siguiente**.
5. En el **Configuración de control de acceso de usuarios** página, en **Ajustes de control de acceso**, elige **Sí** para usar tokens para el control de acceso.
6. Por debajo **Configuración del token**, seleccione **JWT con secreto compartido** como el **Tipo de token**.
7. Por debajo **Parámetros para firmar un secreto compartido**, elige el **Tipo de secreto**. Puede utilizar una existente **AWS Secrets Manager** secreto compartido o cree un secreto compartido nuevo.

Para crear un nuevo secreto compartido, elige **Nuevo**, a continuación, sigue estos pasos:

- a. Por debajo **Nuevo AWS Secrets Manager** secreto, especifique un **Nombre secreto**. El prefijo **AmazonKendra-** se agregará cuando guarde la clave pública.
 - b. Especifique un **ID de clave**. El identificador de clave es un indicio que indica qué clave se usó para proteger la firma web JSON del token.
 - c. Elige la **firmaAlgoritmo** para el token. Este es el algoritmo criptográfico que se utiliza para proteger el token de identificación. Para obtener más información sobre RSA, consulte [Criptografía de RSA](#).
 - d. Especifique un **secreto compartido** introduciendo un secreto codificado en la URL de base64. También puedes seleccionar **Generar secreto** para que se genere un secreto para ti. Debe asegurarse de que el secreto sea un secreto codificado en URL en base64.
 - e. (Opcional) Especifique cuándo es válido el secreto compartido. Puede especificar la fecha y la hora desde las que un secreto es válido, válido para o ambas cosas. El secreto será válido en el intervalo especificado.
 - f. Seleccione **Guardar secreto** para guardar el nuevo secreto.
8. (Opcional) Bajo **Configuración avanzada**:
 - a. Especifique un **Nombre de usuario** para usar en la comprobación de ACL.
 - b. Especifique uno o más **Grupos** para usar en la comprobación de ACL.
 - c. Especifique el **Emisor** o validará al emisor del token.
 - d. Especifique el **ID (s)** de reclamación. Debe especificar una expresión regular que coincida con la audiencia del JWT.
 9. En el **Detalles de aprovisionamiento** página, elige **Edición** para desarrolladores.

10. Escoja `Create` para crear tu índice.
11. Espere a que se cree el índice. Amazon Kendra proporciona el hardware para su índice. Esta operación puede llevar algún tiempo.

CLI

Puede usar el token JWT con un secreto compartido dentro de AWS Secrets Manager. El secreto debe ser un secreto codificado en la URL de base64. Necesitas el Secrets Manager ARN y su Amazon Kendra rol debe tener acceso a `GetSecretValue` en el Secrets Manager recurso. Si está cifrando el Secrets Manager recurso con AWS KMS, el rol también debe tener acceso a la acción de descifrado.

Para crear un índice con AWS CLI mediante un archivo de entrada JSON, primero cree un archivo JSON con los parámetros que deseas:

```
{  
    "Name": "user-context",  
    "Edition": "ENTERPRISE_EDITION",  
    "RoleArn": "arn:aws:iam::account-id:role:/my-role",  
    "UserTokenConfigurations": [  
        {  
            "JwtTokenTypeConfiguration": {  
                "KeyLocation": "SECRET_MANAGER",  
                "Issuer": "optional: specify the issuer url",  
                "ClaimRegex": "optional: regex to validate claims in the token",  
                "UserNameAttributeField": "optional: user",  
                "GroupAttributeField": "optional: group",  
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account  
id:secret:/my-user-context-secret"  
            }  
        }  
    ],  
    "UserContextPolicy": "USER_TOKEN"  
}
```

Puede anular los nombres de campo de usuario y grupo predeterminados. El valor predeterminado para `UserNameAttributeField` es «usuario». El valor predeterminado para `GroupAttributeField` es «grupos».

A continuación, llame `create-index` mediante el archivo de entrada. Por ejemplo, si el nombre del archivo JSON es `create-index-openid.json`, puede utilizar lo siguiente:

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

El secreto debe tener el siguiente formato en AWS Secrets Manager:

```
{  
    "keys": [  
        {  
            "kid": "key_id",  
            "alg": "HS256|HS384|HS512",  
            "kty": "OCT",  
            "use": "sig", //this value can be sig only for now  
            "k": "secret",  
            "nbf": "ISO1806 date format"  
            "exp": "ISO1806 date format"  
        }  
    ]  
}
```

Para obtener más información sobre JWT, consulte[jwt.io](#).

Python

Puede usar el token JWT con un secreto compartido dentro deAWS Secrets Manager. El secreto debe ser un secreto codificado en la URL de base64. Necesitas elSecrets ManagerARN y suAmazon Kendra rol debe tener acceso aGetSecretValueen elSecrets Managerrecurso. Si está cifrando elSecrets Managerrecurso conAWS KMS, el rol también debe tener acceso a la acción de descifrado.

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account-id:role:/my-role',
    UserTokenConfigurations=[
        {
            "JwtTokenTypeConfiguration": {
                "KeyLocation": "URL",
                "Issuer": "optional: specify the issuer url",
                "ClaimRegex": "optional: regex to validate claims in the token",
                "UserNameAttributeField": "optional: user",
                "GroupAttributeField": "optional: group",
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account
id:secret:/my-user-context-secret"
            }
        },
        ],
    UserContextPolicy='USER_TOKEN'
)
```

Uso de un token web JSON (JWT) con una clave pública

Los siguientes ejemplos muestran cómo utilizar un token web JSON (JWT) con un token de certificado para el control de acceso de los usuarios al crear un índice. Para obtener más información sobre JWT, consulte[jwt.io](#).

Console

1. EscojaCrear índicepara empezar a crear un índice nuevo.
2. En elEspecificarpaginade los detalles del índice, asigne un nombre y una descripción a su índice.
3. ParaFunción de IAM, seleccione un rol o seleccioneCrear un nuevo rola y especifique un nombre de rol para crear un rol nuevo. ElIAMel rol tendrá el prefijo»AmazonKendra».
4. Deje todos los demás campos con sus valores predeterminados. Elija Siguiente.
5. En elConfigurar el control de acceso de usuariospaginade, enAjustes de control de acceso, eligeSípara usar tokens para el control de acceso.
6. BajoConfiguración del token, seleccioneJWT con clave públicacomo elTipo de token.
7. BajoParámetros para firmar la clave pública, elige elTipo de secreto. Puede utilizar una existenteAWS Secrets Managersecreto o crea un secreto nuevo.

Para crear un secreto nuevo, eligeNuevoy, a continuación, sigue estos pasos:

- a. BajonuevoAWS Secrets Managersecreto, especifique unNombre secreto. El prefijoAmazonKendra-se agrega cuando guarde la clave pública.
- b. Especifique unID de clave. El identificador de clave es un indicio que indica qué clave se usó para proteger la firma web JSON del token.

- c. Elige la firmaAlgoritmopara el token. Este es el algoritmo criptográfico que se utiliza para proteger el token de identificación. Para obtener más información sobre RSA, consulte [Criptografía de RSA](#).
 - d. BajoAtributos del certificado, especifique unopcional Cadena de certificados. La cadena de certificados se compone de una lista de certificados. Comienza con el certificado de un servidor y termina con el certificado raíz.
 - e. OpcionalEspecifique elHuella digital o huella digital. Debe ser un hash de un certificado, calculado sobre todos los datos del certificado y su firma.
 - f. Especifique elExponente. Este es el valor exponente de la clave pública RSA. Se representa como un valor codificado en Base64urlUInt.
 - g. Especifique elMódulo. Este es el valor exponente de la clave pública RSA. Se representa como un valor codificado en Base64urlUInt.
 - h. SeleccioneGuardar clavepara guardar la nueva clave.
8. OpcionalBajoConfiguración avanzada:
 - a. Especifique unNombre de usuariopara usar en la comprobación de ACL.
 - b. Especifique uno o másGrupospara usar en la comprobación de ACL.
 - c. Especifique elEmisoreso validará al emisor del token.
 - d. Especifique elID (s) de cliente. Debe especificar una expresión regular que coincida con la audiencia del JWT.
 9. En elDetalles de aprovisionamiento página, eligeEdición para desarrolladores.
 10. EscojaCrearpara crear tu índice.
 11. Espere a que se cree el índice.Amazon Kendraaprovisiona el hardware para su índice. Esta operación puede llevar algún tiempo.

CLI

Puede usar JWT con una clave pública dentro de unAWS Secrets Manager. Necesitas elSecrets ManagerARN y suAmazon Kendrael rol debe tener acceso aGetSecretValueen elSecrets Managerrecurso. Si está cifrando elSecrets Managerrecurso conAWS KMS, el rol también debe tener acceso a la acción de descifrado.

Para crear un índice conAWS CLImediante un archivo de entrada JSON, primero cree un archivo JSON con los parámetros que desee:

```
{  
    "Name": "user-context",  
    "Edition": "ENTERPRISE_EDITION",  
    "RoleArn": "arn:aws:iam::account id:role:/my-role",  
    "UserTokenConfigurationList": [  
        {  
            "JwtTokenTypeConfiguration": {  
                "KeyLocation": "SECRET_MANAGER",  
                "Issuer": "optional: specify the issuer url",  
                "ClaimRegex": "optional: regex to validate claims in the token",  
                "UserNameAttributeField": "optional: user",  
                "GroupAttributeField": "optional: group",  
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account  
id:secret:/my-user-context-secret"  
            }  
        },  
        {"UserContextPolicy": "USER_TOKEN"}  
    ]  
}
```

Puede anular los nombres de campo de usuario y grupo predeterminados. El valor predeterminado para `UserNameAttributeField` es «usuario». El valor predeterminado para `GroupAttributeField` es «grupos».

A continuación, llame `create-index` mediante el archivo de entrada. Por ejemplo, si el nombre del archivo JSON es `create-index-openid.json`, puede utilizar lo siguiente:

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

El secreto debe tener el siguiente formato en 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 thumbprint for x.509 cert",  
            "x5c": [  
                "certificate chain"  
            ]  
        }  
    ]  
}
```

Para obtener más información sobre JWT, consulte [jwt.io](#).

Python

```
response = kendra.create_index(  
    Name='user-context',  
    Edition='ENTERPRISE_EDITION',  
    RoleArn='arn:aws:iam::account id:role:/my-role',  
    UserTokenConfigurationList=[  
        {  
            "JwtTokenTypeConfiguration": {  
                "KeyLocation": "URL",  
                "Issuer": "optional: specify the issuer url",  
                "ClaimRegex": "optional: regex to validate claims in the token",  
                "UserNameAttributeField": "optional: user",  
                "GroupAttributeField": "optional: group",  
                "SecretManagerArn": "arn:aws:secretsmanager:us-west-2:account  
id:secret:/my-user-context-secret"  
            }  
        }  
    ],  
    UserContextPolicy='USER_TOKEN'  
)
```

Uso de JSON

Los siguientes ejemplos muestran cómo utilizar un JWT con un token de certificado para el control de acceso de los usuarios al crear un índice.

Warning

El token JSON es una carga no validada. Solo debe usarse cuando las solicitudes Amazon Kendra provengan de un servidor de confianza y nunca de un navegador.

Console

1. Elija Crear índice para empezar a crear un índice nuevo.
2. En la página Especificar los detalles del índice, asigne un nombre y una descripción al índice.
3. Para IAM rol, seleccione un rol o seleccione Crear un nuevo rol para y especifique un nombre de rol para crear un rol nuevo. El IAM rol tendrá el prefijo "AmazonKendra->".
4. Deje todos los demás campos con sus valores predeterminados. Elija Siguiente.
5. En la página Configurar el control de acceso de los usuarios, en Configuración del control de acceso, elija Sí para usar los identificadores para el control de acceso.
6. En Configuración de token, seleccione JSON como tipo de token.
7. Especifique un nombre de usuario para usarlo en la comprobación de ACL.
8. Especifique uno o más grupos para usarlos en la comprobación de ACL.
9. Elija Siguiente.
10. En la página de detalles del aprovisionamiento, seleccione Edición para desarrolladores.
11. Elije Crear para crear tu índice.
12. Espere a que se cree el índice. Amazon Kendraaprovisiona el hardware para su índice. Esta operación puede llevar algún tiempo.

CLI

Para crear un índice con el AWS CLI uso de un archivo de entrada JSON, primero cree un archivo JSON con los parámetros que deseas:

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

A continuación, llame `create-index` con el archivo de entrada. Por ejemplo, si el nombre del archivo JSON es `create-index-openid.json`, puede utilizar lo siguiente:

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

Si no está utilizando Open ID paraAWS IAM Identity Center (successor to AWS Single Sign-On), puede enviarnos el token en formato JSON. Si lo hace, debe especificar qué campo del token JSON contiene el nombre de usuario y qué campo contiene los grupos. Los valores de los campos del grupo deben ser una matriz de cadenas JSON. Por ejemplo, si utilizas SAML, tu token sería similar al siguiente:

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

```
}
```

TokenConfigurationEspecificaría el nombre de usuario y los nombres de los campos del grupo:

```
{
    "UserNameAttributeField": "username",
    "GroupAttributeField": "groups"
}
```

Python

```
response = kendra.create_index(
    Name='user-context',
    Edition='ENTERPRISE_EDITION',
    RoleArn='arn:aws:iam::account-id:role:/my-role',
    UserTokenConfigurationList=[
        {
            "JwtTokenTypeConfiguration": {
                "UserNameAttributeField": "user",
                "GroupAttributeField": "group",
            }
        }
    ],
    UserContextPolicy='USER_TOKEN'
)
```

Creación de un conector de fuente de datos

Puede crear un conector de fuente de datos para Amazon Kendra para conectarse a sus documentos e indexarlos. Amazon Kendra puede conectarse a Microsoft SharePoint, Google Drive y muchos otros proveedores. Cuando crea un conector de fuente de datos, proporciona Amazon Kendra la información de configuración necesaria para conectarse al repositorio de origen. A diferencia de agregar documentos directamente a un índice, puede escanear periódicamente la fuente de datos para actualizar el índice.

Por ejemplo, supongamos que tiene un repositorio de documentos fiscales almacenado en un Amazon S3 balde. De vez en cuando, los documentos existentes se modifican y se añaden nuevos documentos al repositorio. Si agregas el repositorio a Amazon Kendra como fuente de datos, puedes mantener el índice actualizado mediante la configuración de sincronizaciones periódicas entre la fuente de datos y el índice.

Puede optar por actualizar un índice manualmente mediante la consola o el [StartDataSourceSyncJob API](#). De lo contrario, configura un cronograma para actualizar un índice y sincronizarlo con la fuente de datos.

Un índice puede tener más de una fuente de datos. Cada fuente de datos puede tener su propio programa de actualización. Por ejemplo, puedes actualizar el índice de sus documentos de trabajo a diario, o incluso cada hora, y actualizar los documentos archivados de forma manual cada vez que cambie el archivo.

Si deseas modificar los metadatos o atributos del documento y el contenido durante el proceso de ingestión de documentos, consulta [Custom Document Enrichment de Amazon Kendra](#).

Tenga en cuenta que cada ID de documento debe ser único por índice. No se puede crear un origen de datos para indexar los documentos con sus ID exclusivos y, a continuación, utilizar la API BatchPutDocument para indexar los mismos documentos o viceversa. No se puede crear un origen de datos para indexar los documentos con sus ID exclusivos y, a continuación, utilizar la API BatchPutDocument para indexar los mismos documentos o viceversa.

Establecer un cronograma de actualizaciones

Configure la fuente de datos para que se actualice periódicamente con la consola o mediante la `schedule` parámetro al crear o actualizar una fuente de datos. El contenido del parámetro es una cadena que contiene un `cron`-formato o una cadena de programación o una cadena vacía para indicar que el índice se actualiza a pedido. Para conocer el formato de una expresión cron, consulta [Programar expresiones para reglas en el Amazon CloudWatch Events Guía del usuario](#). Amazon Kendra solo admite expresiones cron. No admite expresiones de tarifas.

Configuración de un idioma

Puedes indexar todos los documentos de una fuente de datos en un idioma compatible. Usted especifica el código de idioma de todos sus documentos en su fuente de datos cuando llama [CreateDataSource](#). Si un documento no tiene un código de idioma especificado en un campo de metadatos, el documento se indexa con el código de idioma que se especifica para todos los documentos a nivel de fuente de datos. Si no especificas un idioma, Amazon Kendra indexa los documentos de una fuente de datos en inglés de forma predeterminada. Para obtener más información sobre los idiomas admitidos, incluidos sus códigos, consulta [Añadir documentos en idiomas distintos del inglés](#).

Puedes indexar todos los documentos de una fuente de datos en un idioma compatible mediante la consola. Ir a Fuentes de datos y edita tu fuente de datos o Agregar fuente de datos si va a añadir una nueva.

fuente de datos. En el Especificar los detalles de la fuente de datos página, elige un idioma del menú desplegable Idioma. Tú seleccionas Actualizar o continúa introduciendo la información de configuración para conectarse a la fuente de datos.

Conectores de fuentes de datos

En esta sección se muestra cómo conectarse a Amazon Kendra a bases de datos y repositorios de fuentes de datos compatibles mediante la Amazon Kendra en la AWS Management Console y los Amazon Kendra APIs.

Temas

- [Esquemas de plantillas de fuentes de datos \(p. 128\)](#)
- [Adobe Experience Manager \(p. 283\)](#)
- [Alfresco \(p. 289\)](#)
- [Amazon RDS/Aurora \(p. 294\)](#)
- [Amazon FSx \(p. 299\)](#)
- [Amazon S3 \(p. 303\)](#)
- [Amazon Kendra Rastreador web \(p. 312\)](#)
- [Amazon WorkDocs \(p. 325\)](#)
- [Box \(p. 328\)](#)
- [Confluence \(p. 333\)](#)
- [Conector de fuente de datos personalizado \(p. 344\)](#)
- [Dropbox \(p. 350\)](#)
- [GitHub \(p. 355\)](#)
- [Gmail \(p. 360\)](#)
- [Google Drive \(p. 365\)](#)
- [Jira \(p. 376\)](#)
- [Microsoft Exchange \(p. 380\)](#)
- [Microsoft OneDrive \(p. 384\)](#)
- [Microsoft SharePoint \(p. 393\)](#)
- [Equipos de Microsoft \(p. 414\)](#)
- [Microsoft Yammer \(p. 420\)](#)
- [Quip \(p. 424\)](#)
- [Salesforce \(p. 427\)](#)
- [ServiceNow \(p. 438\)](#)
- [Slack \(p. 450\)](#)
- [Zendesk \(p. 455\)](#)

Esquemas de plantillas de fuentes de datos

Los siguientes son esquemas de plantillas para fuentes de datos en las que se admiten plantillas.

Temas

- [Adobe Experience Manager esquema de plantilla \(p. 129\)](#)
- [Alfresco esquema de plantilla \(p. 147\)](#)
- [Amazon S3 esquema de plantilla \(p. 153\)](#)
- [Amazon Kendra Esquema de plantillas de Web Crawler \(p. 156\)](#)
- [Esquema de plantillas de Confluence \(p. 165\)](#)

- [Esquema de plantillas de Dropbox \(p. 175\)](#)
- [Esquema de plantillas de Gmail \(p. 181\)](#)
- [Esquema de plantillas de Google Drive \(p. 187\)](#)
- [Esquema de plantillas de Microsoft Exchange \(p. 193\)](#)
- [MicrosoftOneDriveesquema de plantilla \(p. 201\)](#)
- [MicrosoftSharePointesquema de plantilla \(p. 206\)](#)
- [Esquema de plantillas de Microsoft Teams \(p. 217\)](#)
- [Esquema de plantillas de Microsoft Yammer \(p. 227\)](#)
- [Esquema de plantillas de Salesforce \(p. 233\)](#)
- [ServiceNowesquema de plantilla \(p. 267\)](#)
- [Esquema de plantillas de Zendesk \(p. 275\)](#)

Adobe Experience Manageresquema de plantilla

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. Usted proporciona el Adobe Experience Manager la URL del servidor, el tipo de autenticación y si utiliza Adobe Experience Manager (AEM) como servicio en la nube o AEM On-Premise como parte de la configuración de la conexión o los detalles del punto final del repositorio. Además, especifique el tipo de fuente de datos como AEM, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como el Type cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Para obtener más información, consulte [Adobe Experience Manager Esquema JSON \(p. 132\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
URL de AEM	El Adobe Experience Manager URL del servidor. Por ejemplo, si usa AEM On-Premise, incluye el nombre de host y el puerto: https://hostname:port. O bien, si usa AEM como servicio en la nube, puede usar la URL del autor: https://author-xxxxxx-xxxxxx.adobeaecloud.com.
authType	El tipo de autenticación que utiliza, ya sea Basico o Auth2.
Tipo de despliegue	El tipo de Adobe Experience Manager que usas, ya sea CLOUD o ON_PREMISE.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
• page • recurso	Una lista de objetos que mapean los atributos o nombres de campo de su Adobe Experience Manager páginas y recursos para Amazon

Configuración	Descripción
	Kendranombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos.
timeZoneId	<p>Si utiliza AEM On-Premise y la zona horaria de su servidor es diferente a la zona horaria delAmazon KendraConector o índice de AEM, puede especificar la zona horaria del servidor para alinearla con el conector o índice de AEM.</p> <p>La zona horaria predeterminada de AEM On-Premise es la zona horaria delAmazon KendraConector o índice AEM. La zona horaria predeterminada para AEM como servicio en la nube es la hora del meridiano de Greenwich.</p>
<ul style="list-style-type: none"> • pageRootPaths • assetRootPaths 	Una lista de las rutas raíz de las páginas y los recursos. Por ejemplo, la ruta raíz de una página podría ser/content/suby la ruta raíz de un activo podría ser/content/sub/asset1.
CrawAssets	truepara rastrear activos.
Rastrear páginas	truepara rastrear páginas.
<ul style="list-style-type: none"> • pagePathInclusionPatrones • pageNameInclusionPatrones • assetPathInclusionPatrones • assetTypeInclusionPatrones • assetNameInclusionPatrones 	<p>Una lista de patrones de expresiones regulares para incluir determinadas páginas y recursos en suAdobe Experience Managerfuente de datos. Las páginas y los recursos que coinciden con los patrones se incluyen en el índice. Las páginas y los recursos que no coincidan con los patrones se excluyen del índice. Si una página o un recurso coinciden con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.</p>
<ul style="list-style-type: none"> • pagePathExclusionPatrones • pageNameExclusionPatrones • assetPathExclusionPatrones • assetTypeInclusionPatrones • assetNameInclusionPatrones 	<p>Una lista de patrones de expresiones regulares para excluir determinadas páginas y recursos de suAdobe Experience Managerfuente de datos. Las páginas y los recursos que coincidan con los patrones se excluyen del índice. Las páginas y los recursos que no coincidan con los patrones se incluyen en el índice. Si una página o un recurso coinciden con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el contenido no se incluye en el índice.</p>
Componentes de página	Una lista de nombres de los componentes de página específicos que desea indexar.

Configuración	Descripción
contentFragmentVariations	Una lista de nombres para las variaciones guardadas específicas de Adobe Experience Manager Fragmentos de contenido que desea indexar.
type	El tipo de fuente de datos. Especificar AEM como tipo de fuente de datos.
enableIdentityCrawler	true para usar Amazon Kendra como un rastreador de identidades para sincronizar la información principal o de identidad de los usuarios y grupos con acceso a ciertos documentos. Si optas por desactivar el rastreador de identidades, debes cargar la información de identidad o principal mediante el PutPrincipalMapping API .
Modo de sincronización	Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puede elegir entre las siguientes opciones: <ul style="list-style-type: none"> FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice. FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice. CHANGE_LOG para rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice.

Configuración	Descripción
SecretarioN	<p>El nombre de recurso de Amazon (ARN) de unAWS Secrets Managersecreto que contiene los pares clave-valor necesarios para conectarse a suAdobe Experience Manager. El secreto debe contener una estructura JSON con las siguientes claves:</p> <p>Si utiliza la autenticación básica para AEM On-Premise o Cloud:</p> <pre>{ "aemUrl": "Adobe Experience Manager On-Premise host URL", "username": "user name with admin permissions", "password": "password with admin permissions" }</pre> <p>Si utiliza la autenticación OAuth 2.0 para AEM On-Premise:</p> <pre>{ "aemUrl": "Adobe Experience Manager host URL", "clientId": "client ID", "clientSecret": "client secret", "privateKey": "private key" }</pre> <p>Si utiliza la autenticación OAuth 2.0 para AEM como servicio en la nube:</p> <pre>{ "clientId": "client ID", "clientSecret": "client secret", "privateKey": "private key", "orgId": "organization ID", "technicalAccountId": "technical account ID", "imsHost": "Adobe Identity Management System (IMS) host" }</pre>
versión	La versión de esta plantilla que se admite actualmente.

Adobe Experience ManagerEsquema JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties":
  {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "host": {
          "type": "string"
        }
      }
    }
  }
}
```

```
"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":

```


"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"
]
}

```

Alfrescoesquema de plantilla

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. Usted proporciona el Alfrescoel identificador del sitio, la URL del repositorio, la URL de la interfaz de usuario, el tipo de autenticación, si utiliza la nube o de forma local y el tipo de contenido que desea rastrear. Esto se proporciona como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como ALFRESCO, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como el Typecuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [AlfrescoEsquema JSON \(p. 150\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
ID del sitio	El identificador del sitio de Alfresco.
URL del repositorio	La URL de tuAlfrescorepositorio. Puede obtener la URL del repositorio en suAlfrescoadministrador. Por ejemplo, si usaAlfrescoCloud (PaaS), la URL del repositorio podría serhttps://company.alfrescocloud.com. O bien, si usaAlfrescoEn las instalaciones, la URL del

Configuración	Descripción
	repositorio podría serhttps://company-alfresco-instance.company-domain.suffix:port.
webAppUrl	La URL de tuAlfrescointerfaz de usuario. Puede obtener elAlfrescoURL de la interfaz de usuario de suAlfrescoadministrador. Por ejemplo, la URL de la interfaz de usuario podría serhttps://example.com.
repositoryAdditionalProperties	Propiedades adicionales para conectarse con el punto final del repositorio/fuente de datos.
authType	El tipo de autenticación que utiliza, ya seaOAuth2oBasic.
tipo (despliegue)	El tipo deAlfrescoque utilizas, ya seaPAASoON-PREM.
Tipo de rastreo	El tipo de contenido que quieras rastrear, ya seaASPECT(contenido marcado con «Aspectos» enAlfresco),SITE_ID(contenido dentro de un determinadoAlfrescositio), oALL_SITES(contenido en todos tusAlfrescositios).
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none"> documento comment 	Una lista de objetos que asignan los atributos o nombres de campo de sus documentos y comentarios de Alfresco aAmazon Kendranombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos.
Nombre de aspecto	El nombre de un «aspecto» específico que desea indexar.
Propiedades de aspecto	Una lista de propiedades específicas del contenido de «Aspect» que desea indexar.
enableFineGrainedControlar	truepara rastrear «Aspectos».
isCrawlComment	truepara indexar comentarios.
<ul style="list-style-type: none"> inclusionFileNamePatrones inclusionFileTypePatrones inclusionFilePathPatrones 	Una lista de patrones de expresiones regulares para incluir ciertos archivos en suAlfrescofuente de datos. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.

Configuración	Descripción
<ul style="list-style-type: none"> • exclusionFileNamePatrones • exclusionFileTypePatrones • exclusionFilePathPatrones 	Una lista de patrones de expresiones regulares para excluir ciertos archivos de suAlfresco fuente de datos. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
type	El tipo de fuente de datos. Especificar ALFRESCO como tipo de fuente de datos.
SecretarioN	<p>El nombre de recurso de Amazon (ARN) de unAWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a suAlfresco. El secreto debe contener una estructura JSON con las siguientes claves:</p> <p>Si utiliza la autenticación básica:</p> <pre>{ "username": "user name", "password": "password" }</pre> <p>Si usas la autenticación OAuth 2.0:</p> <pre>{ "clientId": "client ID", "clientSecret": "client secret", "tokenUrl": "token URL" }</pre>
Modo de sincronización	<p>Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:</p> <ul style="list-style-type: none"> • FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice. • FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice.
enableIdentityCrawler	true para utilizar el Amazon Kendra rastreador de identidad para sincronizar la información principal o de identidad de los usuarios y grupos con acceso a ciertos documentos. Si optas por desactivar el rastreador de identidades, debes cargar la información de identidad o principal mediante la API PutPrincipalMapping .

Configuración	Descripción
versión	La versión de esta plantilla que se admite actualmente.

AlfrescoEsquema JSON

```
{
  "$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"
                        ]
                    }
                ]
            }
        },
        "comment": {
            "type": "object",
            "properties": {
                "fieldMappings": {
                    "type": "array",
                    "items": {
                        "anyOf": [
                            {
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": [
                                            "STRING",
                                            "DATE",
                                            "STRING_LIST",
                                            "LONG"
                                        ]
                                    }
                                }
                            }
                        ]
                    }
                }
            }
        }
    }
},  
"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"
]
}
```

Amazon S3esquema de plantilla

Incluye un JSON que contiene el esquema de la fuente de datos como parte de la configuración de la plantilla. Debe proporcionar el nombre del bucket de S3 como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como S3y otras configuraciones necesarias. A continuación, especifique TEMPLATEcomo elTypecuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Esquema JSON S3 \(p. 154\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
BucketName	Nombre del bucket de Amazon S3.
Configuraciones de repositorios	Información de configuración para el contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos
• Patrones de inclusión	Una lista de patrones de expresiones regulares para incluir o excluir archivos específicos en

Configuración	Descripción
<ul style="list-style-type: none"> • Patrones de exclusión • Prefijos de inclusión • Prefijos de exclusión 	su Amazon S3 fuente de datos. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
aclConfigurationFileRuta	La ruta del archivo que controla el acceso a los documentos de un Amazon Kendra índice.
metadataFilesPrefix	La ubicación dentro de tu bucket para los archivos de metadatos.
Modo de sincronización	Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir <ul style="list-style-type: none"> • FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice • FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice
type	El tipo de fuente de datos. Especificar S3 como tipo de fuente de datos.
versión	La versión de la plantilla que se admite.

Esquema JSON S3

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "BucketName": {
              "type": "string"
            }
          },
          "required": [
            "BucketName"
          ]
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    }
  }
}
```

```
"repositoryConfigurations": {
    "type": "object",
    "properties": {
        "document": {
            "type": "object",
            "properties": {
                "fieldMappings": {
                    "type": "array",
                    "items": [
                        {
                            "type": "object",
                            "properties": {
                                "indexFieldName": {
                                    "type": "string"
                                },
                                "indexFieldType": {
                                    "type": "string",
                                    "enum": [
                                        "STRING"
                                    ]
                                },
                                "dataSourceFieldName": {
                                    "type": "string"
                                }
                            },
                            "required": [
                                "indexFieldName",
                                "indexFieldType",
                                "dataSourceFieldName"
                            ]
                        }
                    ]
                },
                "required": [
                    "fieldMappings"
                ]
            }
        },
        "required": [
            "document"
        ]
    },
    "additionalProperties": {
        "type": "object",
        "properties": {
            "inclusionPatterns": {
                "type": "array"
            },
            "exclusionPatterns": {
                "type": "array"
            },
            "inclusionPrefixes": {
                "type": "array"
            },
            "exclusionPrefixes": {
                "type": "array"
            },
            "aclConfigurationFilePath": {
                "type": "string"
            },
            "metadataFilesPrefix": {
                "type": "string"
            }
        }
    }
},
```

```
"syncMode": {  
    "type": "string",  
    "enum": [  
        "FULL_CRAWL",  
        "FORCED_FULL_CRAWL"  
    ]  
},  
"type": {  
    "type": "string",  
    "pattern": "S3"  
},  
"version": {  
    "type": "string",  
    "anyOf": [  
        {  
            "pattern": "1.0.0"  
        }  
    ]  
}  
},  
"required": [  
    "connectionConfiguration",  
    "type",  
    "syncMode",  
    "repositoryConfigurations"  
]  
}
```

Amazon KendraEsquema de plantillas de Web Crawler

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. Usted proporciona las URL iniciales o de inicio, o puede proporcionar las URL del mapa del sitio, como parte de la configuración de la conexión o de los detalles del punto final del repositorio. En lugar de enumerar manualmente todas tus URL, puedes proporcionar la ruta a la Amazon S3bucket que almacena un archivo de texto para tu lista de URL iniciales o archivos XML de mapa del sitio, que puedes agrupar en un archivo ZIP en S3.

Especifique también el tipo de fuente de datos como WEBCRAWLERV2, las credenciales de autenticación del sitio web y el tipo de autenticación si sus sitios web requieren autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como el tipo cuando llamas [CreateDataSource](#).

Al seleccionar los sitios web que se van a indexar, se debe respetar la [Política de uso aceptable de Amazon](#) y todas las demás condiciones de Amazon. Recuerda que solo debes usar Amazon Kendra Web Crawler para indexar sus propias páginas web o páginas web que tenga autorización para indexar. Para aprender a parar Amazon Kendra Web Crawler para indexar sus sitios web, consulte [Configuración del robots.txt archivo para Amazon Kendra Web Crawler \(p. 324\)](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Amazon KendraEsquema JSON de Web Crawler \(p. 160\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.

Configuración	Descripción
siteMapUrls	La lista de URL del mapa del sitio web que quieres rastrear. Puedes publicar hasta tres URL de mapa del sitio.
s3SeedUrl	La ruta de S3 al archivo de texto que almacena la lista de URL iniciales o de punto de inicio. Por ejemplo, s3://bucket-name/directory/. Cada URL del archivo de texto debe formatearse en una línea independiente. Puedes incluir hasta 100 URL iniciales en un archivo.
s3SiteMapUrl	La ruta S3 a los archivos XML del mapa del sitio. Por ejemplo, s3://bucket-name/directory/. Puedes listar hasta tres archivos XML de mapa del sitio. Puede agrupar varios archivos de mapa del sitio en un archivo ZIP y almacenar el archivo ZIP en su Amazon S3balde.
seedUrlConnections	La lista de URL iniciales o de punto de partida de los sitios web que quieres rastrear. Puedes incluir hasta 100 URL iniciales.
Ver URL	La URL inicial o del punto de partida.
autenticación	El tipo de autenticación si sus sitios web requieren la misma autenticación; de lo contrario, especifiqueNoAuthentication.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none"> • página web • attachment 	Una lista de objetos que asignan los atributos o nombres de campo de las páginas web y los archivos de páginas web aAmazon Kendranombres de campos de índice. Por ejemplo, la etiqueta de título de la página web HTML se puede asignar a_document_titlecampo de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
Modo de sincronización	<p>Especifique siAmazon Kendradebe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:</p> <ul style="list-style-type: none"> • FORCED_FULL_CRAWLpara volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice. • FULL_CRAWLpara rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice.

Configuración	Descripción
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos.
Límite de tarifa	El número de direcciones URL rastreadas por host de sitio web por minuto.
maxFileSize	El tamaño máximo (en MB) de una página web o un archivo adjunto que se puede rastrear.
Profundidad de rastreo	El número de niveles desde la URL inicial hasta el rastreo. Por ejemplo, la página URL inicial tiene la profundidad 1 y todos los hipervínculos de esta página que también se rastrean tienen la profundidad 2.
maxLinksPerURL	El número máximo de URL de una página web que se debe incluir al rastrear un sitio web. Este número es por página web. A medida que se rastrean las páginas web de un sitio web, también se rastrean las URL a las que se vinculan las páginas web. Las URL de una página web se rastrean por orden de aparición.
crawlSubDomain	truepara rastrear los dominios del sitio web únicamente con subdominios. Por ejemplo, si la URL inicial es»abc.example.com«, entonces»a.abc.example.com«y»b.abc.example.com«también se arrastran. Si no configuras crawlSubDomain nocrawlAllDomain true, entoncesAmazon Kendrasolo rastrea los dominios de los sitios web que deseas rastrear.
crawlAllDomain	truepara rastrear los dominios del sitio web con subdominios y otros dominios a los que se enlazan las páginas web. Si no configuras crawlSubDomain nocrawlAllDomain true, entoncesAmazon Kendrasolo rastrea los dominios de los sitios web que deseas rastrear.
Honra a los robots	truepara respetar las directivas robots.txt de los sitios web que desea rastrear. Estas directivas controlan cómoAmazon KendraWeb Crawler rastrea los sitios web, ya seaAmazon Kendrapuede rastrear solo contenido específico o no rastrear ningún contenido.
Rastrea archivos adjuntos	truepara rastrear los archivos a los que enlazan las páginas web.

Configuración	Descripción
<ul style="list-style-type: none"> • URL de inclusiónCrawlPatterns • URL de inclusiónIndexPatterns 	Una lista de patrones de expresiones regulares para incluir rastrear ciertas URL e indexar los hipervínculos de estas páginas web de URL. Las URL que coinciden con los patrones se incluyen en el índice. Las URL que no coinciden con los patrones se excluyen del índice. Si una URL coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y la URL o las páginas web del sitio web no se incluyen en el índice.
<ul style="list-style-type: none"> • URL de exclusiónCrawlPatterns • URL de exclusiónIndexPatterns 	Una lista de patrones de expresiones regulares para excluir rastrear ciertas URL e indexar los hipervínculos de estas páginas web de URL. Las URL que coinciden con los patrones se excluyen del índice. Las URL que no coinciden con los patrones se incluyen en el índice. Si una URL coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y la URL o las páginas web del sitio web no se incluyen en el índice.
inclusionFileIndexPatrones	Una lista de patrones de expresiones regulares para incluir ciertos archivos de páginas web. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coinciden con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
exclusionFileIndexPatrones	Una lista de patrones de expresiones regulares para excluir ciertos archivos de páginas web. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
proxy	Información de configuración necesaria para conectarse a sus sitios web internos a través de un proxy web.
host	El nombre de host del servidor proxy que desea utilizar para conectarse a sitios web internos. Por ejemplo, el nombre de host de https://a.example.com/page1.htmls»a.example.com».
puerto	El número de puerto del servidor proxy que desea utilizar para conectarse a sitios web internos. Por ejemplo, 443 es el puerto estándar para HTTPS.

Configuración	Descripción
Secretario N (apoderado)	Si se requieren credenciales de proxy web para conectarse a un proveedor de alojamiento web, puede crear unAWS Secrets Managersecreto que almacena las credenciales. Proporcione el nombre de recurso de Amazon (ARN) del secreto.
type	El tipo de fuente de datos. EspecificarWEBCRAWLERV2como tipo de fuente de datos.
SecretarioN	<p>El nombre de recurso de Amazon (ARN) de unAWS Secrets Managersecreto que se utiliza si sus sitios web requieren autenticación para acceder a los sitios web. Las credenciales de autenticación del sitio web se almacenan en el secreto que contiene los pares clave-valor de JSON.</p> <p>Si usa Basic o NTLM/Kerberos, introduzca el nombre de usuario y la contraseña. Las claves JSON del secreto deben seruserName y password. El protocolo de autenticación NTLM incluye el cifrado de contraseñas y el protocolo de autenticación Kerberos incluye el cifrado de contraseñas.</p> <p>Si utiliza la autenticación mediante SAML o mediante formularios, introduzca el nombre de usuario y la contraseña, XPath para el campo de nombre de usuario (y el botón de nombre de usuario si utiliza SAML), XPaths para el campo y el botón de contraseña y la URL de la página de inicio de sesión. Las claves JSON del secreto deben seruserName, password, userNameFieldxpath, userNameButtonLoginPageUrl. Puedes encontrar los XPaths (XML Path Language) de los elementos utilizando las herramientas de desarrollo de tu navegador web. Los XPaths suelen seguir este formato://tagname[@attribute='Value'].</p> <p>Amazon Kendra también comprueba si la información del punto final (URL de inicio) incluida en el secreto es la misma que la información del punto final especificada en los detalles de configuración del extremo de la fuente de datos.</p>
versión	La versión de esta plantilla que se admite actualmente.

Amazon KendraEsquema JSON de Web Crawler

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "version": {
      "type": "string"
    }
  }
}
```

```
"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": {
                                "name": "string",
                                "path": "string"
                            }
                        }
                    ]
                }
            }
        }
    }
}
```

```
"type": "object",
"properties": {
    "indexFieldName": {
        "type": "string"
    },
    "indexFieldType": {
        "type": "string",
        "enum": [
            "STRING",
            "DATE",
            "LONG"
        ]
    },
    "dataSourceFieldName": {
        "type": "string"
    },
    "dateFieldFormat": {
        "type": "string",
        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
    }
},
"required": [
    "indexFieldName",
    "indexFieldType",
    "dataSourceFieldName"
]
},
"required": [
    "fieldMappings"
]
},
"attachment": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": [
                {
                    "type": "object",
                    "properties": {
                        "indexFieldName": {
                            "type": "string"
                        },
                        "indexFieldType": {
                            "type": "string",
                            "enum": [
                                "STRING",
                                "DATE",
                                "LONG"
                            ]
                        },
                        "dataSourceFieldName": {
                            "type": "string"
                        },
                        "dateFieldFormat": {
                            "type": "string",
                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                        }
                    },
                    "required": [
                        "indexFieldName",
                        "indexFieldType",
                        "dataSourceFieldName"
                    ]
                }
            ]
        }
    }
}
```

```
        ]
      }
    ]
  },
  "required": [
    "fieldMappings"
  ]
}
},
{
  "syncMode": {
    "type": "string",
    "enum": [
      "FORCED_FULL_CRAWL",
      "FULL_CRAWL"
    ]
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "rateLimit": {
        "type": "string",
        "default": "300"
      },
      "maxFileSize": {
        "type": "string",
        "default": "50"
      },
      "crawlDepth": {
        "type": "string",
        "default": "2"
      },
      "maxLinksPerUrl": {
        "type": "string",
        "default": "100"
      },
      "crawlSubDomain": {
        "type": "boolean",
        "default": false
      },
      "crawlAllDomain": {
        "type": "boolean",
        "default": false
      },
      "honorRobots": {
        "type": "boolean",
        "default": false
      },
      "crawlAttachments": {
        "type": "boolean",
        "default": false
      },
      "inclusionURLCrawlPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "exclusionURLCrawlPatterns": {
        "type": "array",
        "items": {
          "type": "string"
        }
      },
      "inclusionURLIndexPatterns": {
```

```
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionURLIndexPatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "inclusionFileIndexPatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionFileIndexPatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "proxy": {
        "type": "object",
        "properties": {
            "host": {
                "type": "string"
            },
            "port": {
                "type": "string"
            },
            "secretArn": {
                "type": "string",
                "minLength": 20,
                "maxLength": 2048
            }
        }
    }
},
"required": [
    "rateLimit",
    "maxFileSize",
    "crawlDepth",
    "crawlSubDomain",
    "crawlAllDomain",
    "maxLinksPerUrl",
    "honorRobots"
]
},
"type": {
    "type": "string",
    "pattern": "WEBCRAWLERV2"
},
"secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
},
"version": {
    "type": "string",
    "anyOf": [
        {
            "pattern": "1.0.0"
        }
    ]
}
```

```
        ],
      "required": [
        "connectionConfiguration",
        "repositoryConfigurations",
        "syncMode",
        "type",
        "additionalProperties"
      ]
}
```

Esquema de plantillas de Confluence

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. Proporcionas la URL del host de Confluence, el método de alojamiento y el tipo de autenticación como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como CONFLUENCEV2, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como elType cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Esquema JSON de Confluence \(p. 169\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
URL del servidor	La URL de tu instancia de Confluence. Por ejemplo, https://example.confluence.com .
type	El método de alojamiento de tu instancia de Confluence, ya sea SAASyON_PREM.
authType	El método de autenticación de tu instancia de Confluence, ya sea Basic, OAuth2, o Personal-token.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none">• espacio• page• blog• comment• attachment	Una lista de objetos que asignan los atributos o nombres de campo de tus espacios, páginas, blogs, comentarios y archivos adjuntos de Confluence a Amazon Kendra nombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos). Los nombres de los campos del origen de datos de Confluence deben existir en sus metadatos personalizados de Confluence.
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos.

Configuración	Descripción
<ul style="list-style-type: none"> • inclusionSpaceKeyFiltrar • exclusionSpaceKeyFiltrar • pageTitleRegEX • blogTitleRegEX • commentTitleRegEX • attachmentTitleRegEX • inclusionFileTypePatrones • exclusionFileTypePatrones • inclusionUrlPatterns • exclusionUrlPatterns • fieldForUserID 	Una lista de patrones de expresiones regulares para incluir o excluir ciertos archivos de tu fuente de datos de Confluence. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
Servidor proxy	El nombre de host del proxy web que está utilizando, sin la <code>http://</code> o <code>https://</code> protocolo.
Puerto proxy	El número de publicación utilizado por el protocolo de transporte de URL del host. Debe ser un valor numérico comprendido entre 0 y 65535.
<ul style="list-style-type: none"> • isCrawlPersonalEspacio • isCrawlArchivedEspacio • isCrawlArchivedpágina • isCrawlPage • isCrawlBlog • isCrawlPageComentario • isCrawlPageAdjuntivo • isCrawlBlogComentario • isCrawlBlogAdjuntivo 	true para indexar archivos en tus espacios personales, páginas, blogs, comentarios de página, adjuntos de página, comentarios de blog y archivos adjuntos de blog de Confluence.
type	El tipo de fuente de datos. Especificar <code>CONFLUENCEV2</code> como tipo de fuente de datos.
enableIdentityCrawler	true para activar el rastreador de identidades. El rastreador de identidades está activado de forma predeterminada. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el PutPrincipalMapping API . Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte Filtrado en contexto de usuario .

Configuración	Descripción
Modo de sincronización	<p>Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:</p> <ul style="list-style-type: none">• FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice• FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice

Configuración	Descripción
SecretarioN	<p>El nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene los pares clave-valor necesarios para conectarte a tu instancia de Confluence.</p> <p>Si utilizas la autenticación básica, el secreto debe contener una estructura JSON con las siguientes claves:</p> <pre>{ "username": "Confluence account user name", "password": "Confluence API token" }</pre> <p>Si usas la autenticación OAuth 2.0, el secreto debe contener una estructura JSON con las siguientes claves:</p> <pre>{ "confluenceAppKey": "app key for your Confluence account", "confluenceAppSecret": "app secret from your Confluence token", "confluenceAccessToken": "access token created in Confluence", "confluenceRefreshToken": "refresh token created in Confluence" }</pre> <p>(Solo para Confluence Server) Si usas la autenticación básica, el secreto se almacena en una estructura JSON con las siguientes claves:</p> <pre>{ "hostUrl": "Confluence Server host URL", "username": "Confluence Server user name", "password": "Confluence Server password" }</pre> <p>(Solo para Confluence Server) Si usas la autenticación con token de acceso personal, el secreto se almacena en una estructura JSON con las siguientes claves:</p> <pre>{ "hostUrl": "Confluence Server host URL", "patToken": "Confluence token" }</pre>
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Confluence

```
{  
    "$schema": "http://json-schema.org/draft-04/schema#",  
    "type": "object",  
    "properties": {  
        "connectionConfiguration": {  
            "type": "object",  
            "properties": {  
                "repositoryEndpointMetadata": {  
                    "type": "object",  
                    "properties": {  
                        "hostUrl": {  
                            "type": "string",  
                            "pattern": "https:.*"  
                        },  
                        "type": {  
                            "type": "string",  
                            "enum": [  
                                "SAAS",  
                                "ON_PREM"  
                            ]  
                        },  
                        "authType": {  
                            "type": "string",  
                            "enum": [  
                                "Basic",  
                                "OAuth2",  
                                "Personal-token"  
                            ]  
                        }  
                    }  
                },  
                "required": [  
                    "hostUrl",  
                    "type",  
                    "authType"  
                ]  
            }  
        },  
        "required": [  
            "repositoryEndpointMetadata"  
        ]  
    },  
    "repositoryConfigurations": {  
        "type": "object",  
        "properties": {  
            "space": {  
                "type": "object",  
                "properties": {  
                    "fieldMappings": {  
                        "type": "array",  
                        "items": [  
                            {  
                                "type": "object",  
                                "properties": {  
                                    "indexFieldName": {  
                                        "type": "string"  
                                    },  
                                    "indexFieldType": {  
                                        "type": "string",  
                                        "enum": [  
                                            "STRING",  
                                            "STRING_LIST",  
                                            "DATE"  
                                        ]  
                                    }  
                                }  
                            }  
                        ]  
                    }  
                }  
            }  
        }  
    }  
}
```

```
        },
        "dataSourceFieldName": {
            "type": "string"
        },
        "dateFieldFormat": {
            "type": "string",
            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
        }
    },
    "required": [
        "indexFieldName",
        "indexFieldType",
        "dataSourceFieldName"
    ]
}
],
"required": [
    "fieldMappings"
]
},
"page": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": [
                {
                    "type": "object",
                    "properties": {
                        "indexFieldName": {
                            "type": "string"
                        },
                        "indexFieldType": {
                            "type": "string",
                            "enum": [
                                "STRING",
                                "STRING_LIST",
                                "DATE",
                                "LONG"
                            ]
                        },
                        "dataSourceFieldName": {
                            "type": "string"
                        },
                        "dateFieldFormat": {
                            "type": "string",
                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                        }
                    },
                    "required": [
                        "indexFieldName",
                        "indexFieldType",
                        "dataSourceFieldName"
                    ]
                }
            ]
        }
    }
},
"required": [
    "fieldMappings"
]
},
"blog": {
    "type": "object",

```

```
"properties": {
    "fieldMappings": {
        "type": "array",
        "items": [
            {
                "type": "object",
                "properties": {
                    "indexFieldName": {
                        "type": "string"
                    },
                    "indexFieldType": {
                        "type": "string",
                        "enum": [
                            "STRING",
                            "STRING_LIST",
                            "DATE",
                            "LONG"
                        ]
                    },
                    "dataSourceFieldName": {
                        "type": "string"
                    },
                    "dateFieldFormat": {
                        "type": "string",
                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                },
                "required": [
                    "indexFieldName",
                    "indexFieldType",
                    "dataSourceFieldName"
                ]
            }
        ]
    },
    "required": [
        "fieldMappings"
    ]
},
"comment": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": [
                {
                    "type": "object",
                    "properties": {
                        "indexFieldName": {
                            "type": "string"
                        },
                        "indexFieldType": {
                            "type": "string",
                            "enum": [
                                "STRING",
                                "STRING_LIST",
                                "DATE",
                                "LONG"
                            ]
                        },
                        "dataSourceFieldName": {
                            "type": "string"
                        },
                        "dateFieldFormat": {
                            "type": "string",
                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                        }
                    }
                }
            ]
        }
    }
}
```

```
        "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": {
        "fieldForUserId": {
            "type": "string"
        }
    }
},
```



```
        "type": "string"
    }
},
"exclusionFileTypePatterns": {
    "type": "array",
    "items": {
        "type": "string"
    }
},
"inclusionUrlPatterns": {
    "type": "array",
    "items": {
        "type": "string"
    }
},
"exclusionUrlPatterns": {
    "type": "array",
    "items": {
        "type": "string"
    }
},
"proxyHost": {
    "type": "string"
},
"proxyPort": {
    "type": "string"
}
},
"required": []
},
"type": {
    "type": "string",
    "pattern": "CONFLUENCEV2"
},
"enableIdentityCrawler": {
    "type": "boolean"
},
"syncMode": {
    "type": "string",
    "enum": [
        "FULL_CRAWL",
        "FORCED_FULL_CRAWL"
    ]
},
"secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
}
},
"version": {
    "type": "string",
    "anyOf": [
        {
            "pattern": "1.0.0"
        }
    ]
},
"required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
]
```

}

Esquema de plantillas de Dropbox

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. Proporcionas la clave de la aplicación de Dropbox, el secreto de la aplicación y el token de acceso como parte del secreto que almacena tus credenciales de autenticación. Especifique también el tipo de fuente de datos como `DROPBOX`, el tipo de token de acceso que desea utilizar (temporal o permanente) y otras configuraciones necesarias. A continuación, especifique `TEMPLATE` como el `Type` cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Esquema JSON de Dropbox \(p. 176\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos. Esta fuente de datos no especifica un punto final en <code>repositoryEndpointMetadata</code> . Más bien, la información de conexión se incluye en un AWS Secrets Manager secreto que proporcionas en <code>secretArn</code> .
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none">• file• artículo• papel• atajo	Una lista de objetos que asignan los atributos o nombres de campo de tus archivos de Dropbox, Dropbox Paper y los accesos directos a Amazon Kendra a los nombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
SecretarioN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarte a tu Dropbox. El secreto debe contener una estructura JSON con las siguientes claves: <pre>{ "appKey": "Dropbox app key", "appSecret": "Dropbox app secret", "accesstoken": "temporary access token or refresh access token" }</pre>
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos.
<ul style="list-style-type: none">• inclusionFileNamePatrones	Una lista de patrones de expresiones regulares para incluir ciertos nombres y tipos de archivos de

Configuración	Descripción
• inclusionFileTypePatrones	tu fuente de datos de Dropbox. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
• exclusionFileNamePatrones • exclusionFileTypePatrones	Una lista de patrones de expresiones regulares paraexcluirciertos nombres y tipos de archivos de tu fuente de datos de Dropbox. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y otro de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
• Archivo de rastreo • Papel de rastreo • Crawl Papert • Atajo de rastreo	truepara indexar los archivos de Dropbox, los documentos de Dropbox Paper, las plantillas de Dropbox Paper y los accesos directos a páginas web almacenados en tu Dropbox.
type	El tipo de fuente de datos. Especificardropboxcomo tipo de fuente de datos.
useChangeLog	truepara usar el registro de cambios de Dropbox para determinar qué documentos deben agregarse, actualizarse o eliminarse en el índice. Según el tamaño del registro de cambios, es posible que tarde másAmazon Kendrapara usar el registro de cambios en lugar de escanear todos tus documentos de Dropbox.
Tipo de token	Especifique el tipo de token de acceso: identificador de acceso permanente o temporal. Se recomienda crear un token de acceso de actualización que nunca caduque en Dropbox, en lugar de utilizar un token de acceso único que caduque a las 4 horas. Creas una aplicación y un token de acceso de actualización en la consola para desarrolladores de Dropbox y proporcionas el token de acceso en tu secreto.
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Dropbox

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "type": "string"
      }
    }
  }
}
```

```
"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"
]
},
"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"
                            }
                        }
                    }
                ]
            }
        }
    }
},
"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"
                        ]
                    }
                ]
            }
        }
    }
},
"secretArn": {
    "type": "string"
},
"additionalProperties": {
    "type": "object",
    "properties": {
```

```
        "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"
},
"useChangeLog": {
    "type": "string",
    "enum": [
        "true",
        "false"
    ]
},
"tokenType": {
    "type": "string",
    "enum": [
        "PERMANENT",
        "TEMPORARY"
    ]
},
"version": {
    "type": "string",
    "anyOf": [
        {
            "pattern": "1.0.0"
        }
    ]
},
"additionalProperties": false,
"required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "useChangeLog",
    "secretArn",
    "type",
    "tokenType"
]
}
```

Esquema de plantillas de Gmail

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. Especifique el tipo de fuente de datos como GMAIL, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como el Type cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Esquema JSON de Gmail \(p. 183\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos. Especifique el tipo de fuente de datos y el ARN secreto.
<ul style="list-style-type: none">• message• adjuntos	Una lista de objetos que asignan los atributos o nombres de campo de tus mensajes y archivos adjuntos de Gmail a Amazon Kendra nombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos.
<ul style="list-style-type: none">• inclusionLabelNamePatrones• exclusionLabelNamePatrones• inclusionAttachmentTypePatrones• exclusionAttachmentTypePatrones• inclusionAttachmentNamePatrones• exclusionAttachmentNamePatrones• inclusionSubjectFilter• exclusionSubjectFilter• isSubjectAnd• inclusionFromFilter• exclusionFromFilter• inclusionToFilter• exclusionToFilter• inclusionCcFilter• exclusionCcFilter• inclusionBccFilter• exclusionBccFilter	Una lista de patrones de expresiones regulares para incluir o excluir mensajes con nombres de asunto específicos en tu fuente de datos de Gmail. Los archivos que coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
beforeDateFilter	Especifique los mensajes y archivos adjuntos que se incluirán antes de una fecha determinada.

Configuración	Descripción
afterDateFilter	Especifique los mensajes y archivos adjuntos que se incluirán después de una fecha determinada.
isCrawlAttachment	Un valor booleano para elegir si desea rastrear los archivos adjuntos. Los mensajes se rastrean automáticamente.
type	El tipo de fuente de datos. Especificar GMAIL como tipo de fuente de datos.
shouldCrawlDraftMensajes	Un valor booleano para elegir si desea rastrear los borradores de mensajes.
Modo de sincronización	<p>Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir</p> <ul style="list-style-type: none"> • FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice • FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice <p>Important</p> <p>Como no existe una API para actualizar los mensajes de Gmail eliminados permanentemente, una sincronización de contenido nuevo, modificado o eliminado:</p> <ul style="list-style-type: none"> • No eliminará de tu cuenta los mensajes que se eliminaron permanentemente de Gmail Amazon Kendra índice • No sincronizará los cambios en las etiquetas de correo electrónico de Gmail <p>Para sincronizar tu fuente de datos de Gmail, los cambios en las etiquetas y los mensajes de correo electrónico eliminados permanentemente con tu Amazon Kendra índice, debe ejecutar rastreos completos periódicamente.</p>

Configuración	Descripción
SecretarioN	<p>El nombre de recurso de Amazon (ARN) de un secreto de Secrets Manager que contiene los pares clave-valor necesarios para conectarte a Gmail. El secreto debe contener una estructura JSON con las siguientes claves:</p> <pre>{ "adminAccountEmailId": "service account email", "clientEmailId": "user account email", "privateKey": "private key" }</pre>
versión	La versión de la plantilla que se admite actualmente.

Esquema JSON de Gmail

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "type": "object"
      }
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "message": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string"
                    }
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              ]
            }
          }
        }
      }
    }
  }
}
```

```
        ]
    }
},
"attachments": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": [
                {
                    "type": "object",
                    "properties": {
                        "indexFieldName": {
                            "type": "string"
                        },
                        "indexFieldType": {
                            "type": "string",
                            "enum": ["STRING"]
                        },
                        "dataSourceFieldName": {
                            "type": "string"
                        }
                    },
                    "required": [
                        "indexFieldName",
                        "indexFieldType",
                        "dataSourceFieldName"
                    ]
                }
            ]
        }
    }
},
"required": []
},
"additionalProperties": {
    "type": "object",
    "properties": {
        "inclusionLabelNamePatterns": {
            "type": "array",
            "items": [
                "type": "string"
            ]
        },
        "exclusionLabelNamePatterns": {
            "type": "array",
            "items": [
                "type": "string"
            ]
        },
        "inclusionAttachmentTypePatterns": {
            "type": "array",
            "items": [
                "type": "string"
            ]
        },
        "exclusionAttachmentTypePatterns": {
            "type": "array",
            "items": [
                "type": "string"
            ]
        },
        "inclusionAttachmentNamePatterns": {
            "type": "array",

```

```
        "items": {
            "type": "string"
        }
    },
    "exclusionAttachmentNamePatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "inclusionSubjectFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionSubjectFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "isSubjectAnd": {
        "type": "boolean"
    },
    "inclusionFromFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionFromFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "inclusionToFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionToFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "inclusionCcFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionCcFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "inclusionBccFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    }
```

```
        },
        "exclusionBccFilter": {
            "type": "array",
            "items": {
                "type": "string"
            }
        },
        "beforeDateFilter": {
            "anyOf": [
                {
                    "type": "string",
                    "pattern": "[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
                },
                {
                    "type": "string",
                    "pattern": ""
                }
            ]
        },
        "afterDateFilter": {
            "anyOf": [
                {
                    "type": "string",
                    "pattern": "[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}Z$"
                },
                {
                    "type": "string",
                    "pattern": ""
                }
            ]
        },
        "isCrawlAttachment": {
            "type": "boolean"
        },
        "shouldCrawlDraftMessages": {
            "type": "boolean"
        }
    },
    "required": [
        "isCrawlAttachment",
        "shouldCrawlDraftMessages"
    ]
},
"type" : {
    "type" : "string",
    "pattern": "GMAIL"
},
"syncMode": {
    "type": "string",
    "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL"
    ]
},
"secretArn": {
    "type": "string"
},
"version": {
    "type": "string",
    "anyOf": [
        {
            "pattern": "1.0.0"
        }
    ]
}
},
```

```
    "required": [
        "connectionConfiguration",
        "repositoryConfigurations",
        "additionalProperties",
        "syncMode",
        "secretArn",
        "type"
    ]
}
```

Esquema de plantillas de Google Drive

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. Especifique el tipo de fuente de datos como GOOGLEDRIVE2, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como elType cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Esquema JSON de Google Drive \(p. 189\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos. Esta fuente de datos no especifica un punto final. Tú eliges el tipo de autenticación: serviceAccountOAuth2. La información de conexión se incluye en un AWS Secrets Manager secreto que proporcionas en secretArn.
authType	Elige entre serviceAccountOAuth2 según su caso de uso.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
• file • comment	Una lista de objetos que asignan los atributos o nombres de campo de tu Google Drive a Amazon Kendra nombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos
• maxFileSizeInMegabytes	Especifique un límite de tamaño de archivo en MB que Amazon Kendra debería arrastrarse.
• Comentario IsCrawl	true para indexar comentarios en tu fuente de datos de Google Drive.
• isCrawlMyDriveAndSharedWithMe	true para indexar MyDrive y Shared With Me Drives en tu fuente de datos de Google Drive.

Configuración	Descripción
<ul style="list-style-type: none"> • isCrawlSharedConduce 	truepara indexar Shared Drives en tu fuente de datos de Google Drive.
<ul style="list-style-type: none"> • isCrawlAcl 	truepara rastrear la información de ACL de tu fuente de datos de Google Drive.
<ul style="list-style-type: none"> • excludeUserAccounts • excludeSharedDrives • excludeMimeTypes • exclusionFileTypePatrones • exclusionFileNamePatrones • exclusionFilePathFiltrar 	Una lista de patrones de expresiones regulares paraexcluir ciertos archivos de tu fuente de datos de Google Drive. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y otro de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
<ul style="list-style-type: none"> • includeUserAccounts • includeSharedDrives • includeMimeTypes • inclusionFileTypePatrones • inclusionFileNamePatrones • inclusionFilePathFiltrar 	Una lista de patrones de expresiones regulares para incluir ciertos archivos de tu fuente de datos de Google Drive. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coinciden con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
type	El tipo de fuente de datos. Especificar G000GLEDRIVEV2 como tipo de fuente de datos.
enableIdentityCrawler	truepara activar el rastreador de identidades. El rastreador de identidades está activado de forma predeterminada. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante la API PutPrincipalMapping . Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte Filtrado en contexto de usuario .

Configuración	Descripción
Modo de sincronización	<p>Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:</p> <ul style="list-style-type: none"> • FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice • FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice • CHANGE_LOG para rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice
SecretarioN	<p>El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a Google Drive. El secreto debe contener una estructura JSON con las siguientes claves:</p> <p>Si utilizas la autenticación de la cuenta de servicio de Google:</p> <pre>{ "clientEmail": "user account email", "adminAccountEmail": "service account email", "privateKey": "private key" }</pre> <p>Si usas la autenticación OAuth 2.0:</p> <pre>{ "clientID": "OAuth client ID", "clientSecret": "client secret", "refreshToken": "refresh token" }</pre>
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Google Drive

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object"
        }
      }
    }
  }
}
```

```
"type": "object",
"properties": {
    "authType": {
        "type": "string",
        "enum": [
            "serviceAccount",
            "OAuth2"
        ]
    }
},
"required": [
    "authType"
]
},
"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": {
            "name": {
                "type": "string"
            }
        }
    }
},
"required": [
    "repositoryEndpointMetadata"
]
}
},
```

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

```
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "excludeMimeTypes": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "includeUserAccounts": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "includeSharedDrives": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "includeMimeTypes": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "includeTargetAudienceGroup": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "inclusionFileTypePatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "inclusionFileNamePatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionFileTypePatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionFileNamePatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "inclusionFilePathFilter": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionFilePathFilter": {
```

```
        "type": "array",
        "items": {
            "type": "string"
        }
    }
},
"type": {
    "type": "string",
    "pattern": "GOOGLEDRIVEV2"
},
"enableIdentityCrawler": {
    "type": "boolean"
},
"syncMode": {
    "type": "string",
    "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
    ]
},
"secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
},
"version": {
    "type": "string",
    "anyOf": [
        {
            "pattern": "1.0.0"
        }
    ]
},
"required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
]
}
```

Esquema de plantillas de Microsoft Exchange

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. El ID de inquilino se proporciona como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como `MSEXCHANGE`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique `TEMPLATE` como el `Type` cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Esquema JSON de Microsoft Exchange \(p. 196\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
ID de inquilino	El ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none"> • email • attachment • calendario • contacts • notas 	Una lista de objetos que asignan los atributos o nombres de campo de la fuente de datos de Microsoft Exchange a Amazon Kendra campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos
Patrones de inclusión	Una lista de patrones de expresiones regulares para incluir ciertos archivos de la fuente de datos de Microsoft Exchange. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coinciden con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
Patrones de exclusión	Una lista de patrones de expresiones regulares para excluir ciertos archivos de la fuente de datos de Microsoft Exchange. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y otro de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
<ul style="list-style-type: none"> • inclusionUsersList • inclusionUsersFileName • inclusionDomainUsers 	Una lista de patrones de expresiones regulares para incluir ciertos usuarios y archivos de usuario de su fuente de datos de Microsoft Exchange. Los usuarios que coinciden con los patrones se incluyen en el índice. Los usuarios que no coinciden con los patrones se excluyen del índice. Si un usuario coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el usuario no se incluye en el índice.

Configuración	Descripción
<ul style="list-style-type: none"> • exclusionUsersList • exclusionUsersFileName • exclusionDomainUsers 	Una lista de patrones de expresiones regulares para excluir ciertos usuarios y archivos de usuario de su fuente de datos de Microsoft Exchange. Los usuarios que coincidan con los patrones se excluyen del índice. Los usuarios que no coinciden con los patrones se incluyen en el índice. Si un usuario coincide con un patrón de exclusión y otro de inclusión, el patrón de exclusión tiene prioridad y el usuario no se incluye en el índice.
Nombre del bucket S3	El nombre de su bucket de S3, si lo desea utilizar.
<ul style="list-style-type: none"> • Calendario de rastreo • Notas de rastreo • crawlFolderAcl • Rastrear contactos • crawlFolderAcl 	truepara indexar este contenido en su fuente de datos de Microsoft Exchange.
startCalendarDateHora	Puede configurar una fecha y hora de inicio específicas para el contenido de su calendario.
endCalendarDateHora	Puede configurar una fecha y hora de finalización específicas para el contenido del calendario.
subject	Puedes configurar una línea de asunto específica para el contenido de tu correo.
Correo electrónico de	Puede configurar un correo electrónico específico para el contenido del correo «De» o del remitente.
Enviar correo electrónico a	Puede configurar un correo electrónico específico para el contenido del correo «Para» o del destinatario.
Modo de sincronización	Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre: <ul style="list-style-type: none"> • FORCED_FULL_CRAWLpara volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice • FULL_CRAWLpara rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice • CHANGE_LOGpara rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice
type	El tipo de fuente de datos. Especificarmsexchangecomo tipo de fuente de datos.

Configuración	Descripción
SecretarioN	El nombre de recurso de Amazon (ARN) de unAWS Secrets Managersecreto que contiene los pares clave-valor necesarios para conectarse a Microsoft Exchange. Esto incluye su ID de cliente y su secreto de cliente que se generan al crear una aplicación OAuth en el portal de Azure.
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Microsoft Exchange

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "tenantId": {
              "type": "string",
              "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}$",
              "minLength": 36,
              "maxLength": 36
            }
          },
          "required": ["tenantId"]
        }
      }
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "email": {
          "type": "object",
          "properties": {
            "fieldMappings": {
              "type": "array",
              "items": [
                {
                  "type": "object",
                  "properties": {
                    "indexFieldName": {
                      "type": "string"
                    },
                    "indexFieldType": {
                      "type": "string",
                      "enum": ["STRING", "STRING_LIST", "DATE"]
                    },
                    "dataSourceFieldName": {
                      "type": "string"
                    },
                    "dateFieldFormat": {
                      "type": "string",
                      "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                    }
                  }
                }
              ]
            }
          }
        }
      }
    }
  }
}
```

```
        "required": [
            "indexFieldName",
            "indexFieldType",
            "dataSourceFieldName"
        ]
    }
}
],
"required": [
    "fieldMappings"
]
},
"attachment": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": [
                {
                    "type": "object",
                    "properties": {
                        "indexFieldName": {
                            "type": "string"
                        },
                        "indexFieldType": {
                            "type": "string",
                            "enum": ["STRING", "DATE", "LONG"]
                        },
                        "dataSourceFieldName": {
                            "type": "string"
                        },
                        "dateFieldFormat": {
                            "type": "string",
                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                        }
                    },
                    "required": [
                        "indexFieldName",
                        "indexFieldType",
                        "dataSourceFieldName"
                    ]
                }
            ]
        }
    }
},
"required": [
    "fieldMappings"
]
},
"calendar": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": [
                {
                    "type": "object",
                    "properties": {
                        "indexFieldName": {
                            "type": "string"
                        },
                        "indexFieldType": {
                            "type": "string",
                            "enum": ["STRING", "STRING_LIST", "DATE"]
                        },
                        "dataSourceFieldName": {
                            "type": "string"
                        }
                    }
                }
            ]
        }
    }
}
```

```
        "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", "STRING_LIST", "DATE"]
                        },
                        "dataSourceFieldName": {
                            "type": "string"
                        },
                        "dateFieldFormat": {
                            "type": "string",
                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                        }
                    },
                    "required": [
                        "indexFieldName",
                        "indexFieldType",
                        "dataSourceFieldName"
                    ]
                }
            ]
        }
    }
},
"required": [
    "fieldMappings"
]
}
```

```
        "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"
    ]
}

```

MicrosoftOneDriveesquema de plantilla

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. El ID de inquilino se proporciona como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como ONEDRIVEV2, y un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como el Type cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [MicrosoftOneDriveEsquema JSON \(p. 203\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.

Configuración	Descripción
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
ID de inquilino	El ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
file	Una lista de objetos que mapean los atributos o nombres de campo de su MicrosoftOneDrivearchivos aAmazon Kendranombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos
<ul style="list-style-type: none"> • userNameFilter • userFilterPath • inclusionFileTypePatrones • exclusionFileTypePatrones • inclusionFileNamePatrones • exclusionFileNamePatrones • inclusionFilePathPatrones • exclusionFilePathPatrones • inclusionOneNoteSectionNamePatterns • exclusionOneNoteSectionNamePatterns • inclusionOneNotePageNamePatterns • exclusionOneNotePageNamePatterns 	Puede optar por indexar archivos específicos, OneNotes secciones, OneNote páginas y filtra por nombre de usuario.
isUserNameEn S3	truepara proporcionar una lista de nombres de usuario en un archivo almacenado en unAmazon S3.
type	El tipo de fuente de datos. Especificar ONEDRIVEV2 como tipo de fuente de datos.

Configuración	Descripción
enableIdentityCrawler	truepara activar el rastreador de identidades. El rastreador de identidades está activado de forma predeterminada. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el PutPrincipalMappingAPI . Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte Filtrado en contexto de usuario .
type	El tipo de fuente de datos. EspecificarONEDRIVEV2como tipo de fuente de datos.
Modo de sincronización	Especifique siAmazon Kendradebe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre: <ul style="list-style-type: none"> FORCED_FULL_CRAWLpara volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice FULL_CRAWLpara rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice CHANGE_LOGpara rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice
SecretarioN	El nombre de recurso de Amazon (ARN) de unAWS Secrets Managersecreto que contiene los pares clave-valor necesarios para conectarse a su MicrosoftOneDrive. El secreto debe contener una estructura JSON con las siguientes claves: <pre>{ "clientId": "client ID", "clientSecret": "client secret" }</pre>
versión	La versión de esta plantilla que se admite actualmente.

MicrosoftOneDriveEsquema JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
```

```
"type": "object",
"properties": {
  "connectionConfiguration": {
    "type": "object",
    "properties": {
      "repositoryEndpointMetadata": {
        "type": "object",
        "properties": {
          "tenantId": {
            "type": "string",
            "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}$",
            "minLength": 36,
            "maxLength": 36
          }
        },
        "required": [
          "tenantId"
        ]
      }
    },
    "required": [
      "repositoryEndpointMetadata"
    ]
  },
  "repositoryConfigurations": {
    "type": "object",
    "properties": {
      "file": {
        "type": "object",
        "properties": {
          "fieldMappings": {
            "type": "array",
            "items": [
              {
                "type": "object",
                "properties": {
                  "indexFieldName": {
                    "type": "string"
                  },
                  "indexFieldType": {
                    "type": "string",
                    "enum": [
                      "STRING",
                      "STRING_LIST",
                      "DATE",
                      "LONG"
                    ]
                  },
                  "dataSourceFieldName": {
                    "type": "string"
                  },
                  "dateFieldFormat": {
                    "type": "string",
                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                  }
                },
                "required": [
                  "indexFieldName",
                  "indexFieldType",
                  "dataSourceFieldName"
                ]
              }
            ]
          }
        }
      }
    },
    "required": [
      "file"
    ]
  }
}
```

```
        "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"
]
}
```

MicrosoftSharePointesquema de plantilla

Incluye un JSON que contiene el esquema de la fuente de datos como parte de [TemplateConfiguration](#) objeto. Usted proporciona el SharePoint la URL del host, el dominio y también un identificador de inquilino, si es necesario, como parte de la configuración de la conexión o los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como **SHAREPOINTV2**,

un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique `TEMPLATE` como el `Tecleacuando llamas`[CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [SharePointEsquema JSON \(p. 209\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos
repositoryEndpointMetadata	La información del punto final de la fuente de datos
ID de inquilino	El ID de inquilino de su SharePointcuenta.
dominio	El dominio de tu SharePointcuenta.
URL del sitio	Las URL de host de tu SharePointcuenta.
repositoryAdditionalProperties	Propiedades adicionales para conectarse con el punto final del repositorio/fuente de datos.
Nombre del bucket S3	El nombre del Amazon S3bucket que almacena su certificado X.509 autofirmado de Azure AD.
Nombre del certificado S3	El nombre del certificado X.509 autofirmado de Azure AD almacenado en su Amazon S3balde.
authType	El tipo de autenticación que utiliza, ya sea <code>OAuth2</code> , <code>OAuth2Certificate</code> , <code>OAuth2App</code> , <code>Basic</code> .
versión	El SharePointversión que está utilizando, ya sea <code>ServerOnline</code> .
onPremVersion	El SharePointLa versión de servidor que está utilizando, ya sea <code>2013</code> , <code>2016</code> , <code>2019</code> , o <code>SubscriptionEdition</code> .
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos.
<ul style="list-style-type: none"> • <code>eventTitleFilterRegEx</code> • <code>pageTitleFilterRegEx</code> • <code>linkTitleFilterRegEx</code> • <code>inclusionFilePath</code> • <code>exclusionFilePath</code> • <code>inclusionFileTypePatrones</code> • <code>exclusionFileTypePatrones</code> • <code>inclusionFileNamePatrones</code> • <code>exclusionFileNamePatrones</code> • <code>inclusionOneNoteSectionNamePatterns</code> • <code>exclusionOneNoteSectionNamePatterns</code> 	Una lista de patrones de expresiones regulares para incluir/excluir ciertos archivos de la fuente de datos de Sharepoint. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.

Configuración	Descripción
<ul style="list-style-type: none"> inclusionOneNotePageNamePatterns exclusionOneNotePageNamePatterns Configuración ACL Dominio de correo electrónico Servidor proxy Puerto proxy 	
<ul style="list-style-type: none"> Rastrea archivos Rastrear páginas Eventos de rastreo Rastrear comentarios Enlaces de rastreo Rastrea archivos adjuntos crawlListData Rastrea la CL isCrawlLocalGroupMapping isCrawlAdGroupMapping 	Entrada TRUE para indexar.
type	Especificar SHAREPOINTV2 como tipo de fuente de datos
enableIdentityCrawler	<p>true para activar el rastreador de identidades. El rastreador de identidades está activado de forma predeterminada. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el PutPrincipalMapping API.</p> <p>Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte Filtrado en contexto de usuario.</p>
Modo de sincronización	<p>Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:</p> <ul style="list-style-type: none"> FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice CHANGE_LOG para rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice

Configuración	Descripción
SecretarioN	El nombre de recurso de Amazon (ARN) de unAWS Secrets Managersecreto que contiene los pares clave-valor necesarios para conectarse a suSharePoint. Para obtener información sobre estos pares clave-valor, consulte Instrucciones de conexión para SharePoint En línea y SharePoint Servidor.
versión	La versión de esta plantilla que se admite actualmente.

SharePointEsquema JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "tenantId": {
              "type": "string",
              "pattern": "[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}",
              "minLength": 36,
              "maxLength": 36
            },
            "domain": {
              "type": "string"
            },
            "siteUrls": {
              "type": "array",
              "items": {
                "type": "string",
                "pattern": "https://.*"
              }
            }
          }
        },
        "repositoryAdditionalProperties": {
          "type": "object",
          "properties": {
            "s3bucketName": {
              "type": "string"
            },
            "s3certificateName": {
              "type": "string"
            },
            "authType": {
              "type": "string",
              "enum": [
                "OAuth2",
                "OAuth2Certificate",
                "OAuth2App",
                "Basic",
                "NTLM",
                "Kerberos"
              ]
            },
            "version": {
              "type": "string"
            }
          }
        }
      }
    }
  }
}
```

```
        "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": [
    "indexFieldName",
    "indexFieldType",
    "dataSourceFieldName"
]
```

```
        "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"  
},  
"aclConfiguration": {  
    "type": "string",  
    "enum": [  
        "ACLWithLDAPEmailFmt",  
        "ACLWithManualEmailFmt",  
        "ACLWithUsernameFmt"  
    ]  
}
```

```
        ],
      },
      "emailDomain": {
        "type": "string"
      },
      "isCrawlLocalGroupMapping": {
        "type": "boolean"
      },
      "isCrawlAdGroupMapping": {
        "type": "boolean"
      },
      "proxyHost": {
        "type": "string"
      },
      "proxyPort": {
        "type": "string"
      }
    },
    "required": [
    ]
  },
  "type": {
    "type": "string",
    "pattern": "SHAREPOINTV2"
  },
  "enableIdentityCrawler": {
    "type": "boolean"
  },
  "syncMode": {
    "type": "string",
    "enum": [
      "FULL_CRAWL",
      "FORCED_FULL_CRAWL",
      "CHANGE_LOG"
    ]
  },
  "secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
  }
},
"version": {
  "type": "string",
  "anyOf": [
    {
      "pattern": "1.0.0"
    }
  ]
},
"required": [
  "connectionConfiguration",
  "repositoryConfigurations",
  "enableIdentityCrawler",
  "syncMode",
  "additionalProperties",
  "secretArn",
  "type"
]
}
```

Esquema de plantillas de Microsoft Teams

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. El ID de inquilino se proporciona como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como MSTEAMS, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como el Type cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Esquema JSON de Microsoft Teams \(p. 218\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
ID de inquilino	El ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none">• Mensaje de chat• Archivo adjunto al chat• Publicación del canal• Canal Wiki• Conexión al canal• Chat de reunión• Nota de reunión• Reunión del calendario	Una lista de objetos que asignan los atributos o nombres de campo del contenido de Microsoft Teams a Amazon Kendra nombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos.
<ul style="list-style-type: none">• isCrawlChatMensaje• isCrawlChatAdjuntivo• isCrawlChannelPublicar• isCrawlChannelAdjuntivo• isCrawlChannelWiki• isCrawlCalendarReunión• isCrawlMeetingCharla• isCrawlMeetingExpediente• isCrawlMeetingNota	true para indexar este contenido en su fuente de datos de Microsoft Teams.
Modelo de pago	Especifica qué tipo de modelo de pago se va a utilizar con la fuente de datos de Teams. Los modelos de pago del modelo A están restringidos

Configuración	Descripción
	a los modelos de licencias y pago que requieren el cumplimiento de las normas de seguridad. Los modelos de pago del modelo B son adecuados para los modelos de licencias y pago que no requieren el cumplimiento de las normas de seguridad.
Modo de sincronización	<p>Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:</p> <ul style="list-style-type: none"> • FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice • FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice • CHANGE_LOG para rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice
SecretarioN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a Microsoft Teams. Esto incluye el ID de cliente y el secreto de cliente que se generan al crear una aplicación OAuth en el portal de Azure.
type	El tipo de fuente de datos. Especificando MSTEAMS como tipo de fuente de datos.
enableIdentityCrawler	true para activar el rastreador de identidades. El rastreador de identidades está activado de forma predeterminada. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el PutPrincipalMapping API . Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte Filtrado en contexto de usuario .
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Microsoft Teams

```
{
```

```
"$schema": "http://json-schema.org/draft-04/schema#",
"type": "object",
"properties": {
    "connectionConfiguration": {
        "type": "object",
        "properties": {
            "repositoryEndpointMetadata": {
                "type": "object",
                "properties": {
                    "tenantId": {
                        "type": "string",
                        "pattern": "^[0-9a-f]{8}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{4}-[0-9a-f]{12}$",
                        "minLength": 36,
                        "maxLength": 36
                    }
                },
                "required": [
                    "tenantId"
                ]
            }
        },
        "required": [
            "repositoryEndpointMetadata"
        ]
    },
    "repositoryConfigurations": {
        "type": "object",
        "properties": {
            "chatMessage": {
                "type": "object",
                "properties": {
                    "fieldMappings": {
                        "type": "array",
                        "items": [
                            {
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": [
                                            "STRING",
                                            "STRING_LIST",
                                            "DATE"
                                        ]
                                    },
                                    "dataSourceFieldName": {
                                        "type": "string"
                                    },
                                    "dateFieldFormat": {
                                        "type": "string",
                                        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                    }
                                },
                                "required": [
                                    "indexFieldName",
                                    "indexFieldType",
                                    "dataSourceFieldName"
                                ]
                            }
                        ]
                    }
                }
            },
            "required": [
                "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",  
                        "dateFieldFormat"  
                    ]  
                }  
            ]  
        }  
    }  
},  
"required": [  
    "indexFieldName",  
    "indexFieldType",  
    "dataSourceFieldName",  
    "dateFieldFormat"  
]
```

```
        "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"  
            }  
        }  
    }  
}
```

```
        }
    },
    "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": [
        "Sync"
    ]
}
```

```

        "FORCED_FULL_CRAWL",
        "FULL_CRAWL",
        "CHANGE_LOG"
    ],
},
"secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
}
},
"version": {
    "type": "string",
    "anyOf": [
        {
            "pattern": "1.0.0"
        }
    ]
},
"required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
]
}
}

```

Esquema de plantillas de Microsoft Yammer

Incluye un JSON que contiene el esquema de la fuente de datos como parte de [TemplateConfiguration](#) objeto. Especifique el tipo de fuente de datos como YAMMER, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como el Tercleacuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores.

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos. Esta fuente de datos no especifica un punto final en repositoryEndpointMetadata. Más bien, la información de conexión se incluye en un AWS Secrets Manager secreto que proporcionas en secretArn.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none"> • comunidad • user • message • attachment 	Lista de objetos que asignan atributos o nombres de campo del contenido de Microsoft Yammer a los nombres de campos de índice de Amazon Kendra. Para obtener más información, consulte

Configuración	Descripción
	Mapping data source fields (Asignación de campos de origen de datos).
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos
Patrones de inclusión	Una lista de patrones de expresiones regulares para incluir ciertos archivos de la fuente de datos de Microsoft Yammer. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coinciden con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
Patrones de exclusión	Una lista de patrones de expresiones regulares para excluir ciertos archivos de la fuente de datos de Microsoft Yammer. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y otro de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
Desde la fecha	Puede optar por configurar <code>unsinceDate</code> parámetro para que el conector Yammer de Microsoft rastree el contenido en función de un parámetro específicosinceDate.
communityNameFilter	Puedes elegir indexar el contenido específico de la comunidad.
<ul style="list-style-type: none"> • <code>isCrawlMessage</code> • <code>isCrawlAttachment</code> • <code>isCrawlPrivateMensaje</code> 	true para indexar mensajes, archivos adjuntos y mensajes privados.
type	Especificar YAMMER como tipo de fuente de datos.
SecretarioN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a Microsoft Yammer. Esto incluye su nombre de usuario y contraseña de Microsoft Yammer, así como el ID de cliente y el secreto de cliente que se generan al crear una aplicación de OAuth en el portal de Azure.
useChangeLog	true para usar el registro de cambios de Microsoft Yammer para determinar qué documentos deben agregarse, actualizarse o eliminarse en el índice.

Configuración	Descripción
Modo de sincronización	Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre: <ul style="list-style-type: none"> FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice CHANGE_LOG para rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice
enableIdentityCrawler	true para activar el rastreador de identidades. El rastreador de identidades está activado de forma predeterminada. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el PutPrincipalMapping API . Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte Filtrado en contexto de usuario .

Esquema JSON de Microsoft Yammer

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            ...
          }
        }
      },
      "required": [
        "repositoryEndpointMetadata"
      ]
    },
    "repositoryConfigurations": {
      "type": "object",
      "properties": {
        "community": {
          "type": "object",
          ...
        }
      }
    }
  }
}
```

```
"properties": {
    "fieldMappings": {
        "type": "array",
        "items": {
            "anyOf": [
                {
                    "type": "object",
                    "properties": {
                        "indexFieldName": {
                            "type": "string"
                        },
                        "indexFieldType": {
                            "type": "string",
                            "enum": [
                                "STRING",
                                "DATE"
                            ]
                        },
                        "dataSourceFieldName": {
                            "type": "string"
                        },
                        "dateFieldFormat": {
                            "type": "string",
                            "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                        }
                    },
                    "required": [
                        "indexFieldName",
                        "indexFieldType",
                        "dataSourceFieldName"
                    ]
                }
            ]
        }
    },
    "required": [
        "fieldMappings"
    ]
},
"user": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": {
                "anyOf": [
                    {
                        "type": "object",
                        "properties": {
                            "indexFieldName": {
                                "type": "string"
                            },
                            "indexFieldType": {
                                "type": "string",
                                "enum": [
                                    "STRING",
                                    "DATE"
                                ]
                            },
                            "dataSourceFieldName": {
                                "type": "string"
                            },
                            "dateFieldFormat": {
                                "type": "string",
                                "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                            }
                        }
                    }
                ]
            }
        }
    }
}
```

```
        }
    },
    "required": [
        "indexFieldName",
        "indexFieldType",
        "dataSourceFieldName"
    ]
}
]
}
},
"required": [
    "fieldMappings"
]
},
"message": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": {
                "anyOf": [
                    {
                        "type": "object",
                        "properties": {
                            "indexFieldName": {
                                "type": "string"
                            },
                            "indexFieldType": {
                                "type": "string",
                                "enum": [
                                    "STRING",
                                    "DATE"
                                ]
                            },
                            "dataSourceFieldName": {
                                "type": "string"
                            },
                            "dateFieldFormat": {
                                "type": "string",
                                "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                            }
                        },
                        "required": [
                            "indexFieldName",
                            "indexFieldType",
                            "dataSourceFieldName"
                        ]
                    }
                ]
            }
        },
        "required": [
            "fieldMappings"
        ]
    },
    "attachment": {
        "type": "object",
        "properties": {
            "fieldMappings": {
                "type": "array",
                "items": {
                    "anyOf": [
                        {
                            "type": "object",
                            "properties": {
                                "indexFieldName": {
                                    "type": "string"
                                },
                                "indexFieldType": {
                                    "type": "string",
                                    "enum": [
                                        "STRING",
                                        "DATE"
                                    ]
                                },
                                "dataSourceFieldName": {
                                    "type": "string"
                                },
                                "dateFieldFormat": {
                                    "type": "string",
                                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                }
                            },
                            "required": [
                                "indexFieldName",
                                "indexFieldType",
                                "dataSourceFieldName"
                            ]
                        }
                    ]
                }
            }
        }
    }
}
```

```
"type": "object",
"properties": {
    "indexFieldName": {
        "type": "string"
    },
    "indexFieldType": {
        "type": "string",
        "enum": [
            "STRING",
            "DATE"
        ]
    },
    "dataSourceFieldName": {
        "type": "string"
    },
    "dateFieldFormat": {
        "type": "string",
        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
    }
},
"required": [
    "indexFieldName",
    "indexFieldType",
    "dataSourceFieldName"
]
}
],
"required": [
    "fieldMappings"
]
}
},
"additionalProperties": {
    "type": "object",
    "properties": {
        "inclusionPatterns": {
            "type": "array"
        },
        "exclusionPatterns": {
            "type": "array"
        },
        "sinceDate": {
            "type": "string",
            "pattern": "^((19|2[0-9])[0-9]{2}-(0[1-9]|1[012])-(0[1-9]|[12][0-9]|3[01])T(0[0-9]|1[0-9]|2[0-3]):([0-5][0-9]):([0-5][0-9]):((\\+|-)(0[0-9]|1[0-9]|2[0-3]):([0-5][0-9]))?\\$"
        },
        "communityNameFilter": {
            "type": "array",
            "items": {
                "type": "string"
            }
        },
        "isCrawlMessage": {
            "type": "boolean"
        },
        "isCrawlAttachment": {
            "type": "boolean"
        },
        "isCrawlPrivateMessage": {
            "type": "boolean"
        }
}
```

```
        },
        "required": [
            "sinceDate"
        ]
    },
    "type": {
        "type": "string",
        "pattern": "YAMMER"
    },
    "secretArn": {
        "type": "string",
        "minLength": 20,
        "maxLength": 2048
    },
    "useChangeLog": {
        "type": "string",
        "enum": [
            "true",
            "false"
        ]
    },
    "syncMode": {
        "type": "string",
        "enum": [
            "FORCED_FULL_CRAWL",
            "FULL_CRAWL",
            "CHANGE_LOG"
        ]
    },
    "enableIdentityCrawler": {
        "type": "boolean"
    },
    "version": {
        "type": "string",
        "anyOf": [
            {
                "pattern": "1.0.0"
            }
        ]
    }
},
"required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "type",
    "secretArn",
    "syncMode"
]
}
```

Esquema de plantillas de Salesforce

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. Usted proporciona la URL del host de Salesforce como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como SALESFORCEV2, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como elType cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Esquema JSON de Salesforce \(p. 239\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
URL del servidor	La URL de la instancia de Salesforce que se va a indexar.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none"> • cuenta • contact • campaña • caso • producto • lead • contrato • compañero • profile • idea • libro de precios • tarea • solution • attachment • user • documento • Artículos de conocimiento • grupo • oportunidad • charla • Entidad personalizada 	<p>Una lista de objetos que asignan los atributos o nombres de campo de sus entidades de Salesforce a Amazon Kendra nombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).</p>
SecretarioN	<p>El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a su Salesforce. El secreto debe contener una estructura JSON con las siguientes claves:</p> <pre>{ "authenticationUrl": "<i>OAUTH endpoint that Amazon Kendra connects to get an OAUTH token</i>", "consumerKey": "<i>Application public key generated when you created your Salesforce application</i>", "consumerSecret": "<i>Application private key generated when you created your Salesforce application</i>",</pre>

Configuración	Descripción
	<pre>"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" }</pre>
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos
<ul style="list-style-type: none"> • Filtro de cuenta • Filtro de contacto • Filtro de estuche • Filtro de campaña • Filtro de contrato • Filtro de grupo • Filtro de plomo • Filtro de producto • Filtro de oportunidades • Filtro de socios • Filtro Pricebook • Filtro de ideas • Filtro de perfil • Filtro de tareas • Filtro de soluciones • Filtro de usuario • Filtro Chatter • Filtro de documentos • knowledgeArticleFilter • Entidades personalizadas 	Colección de cadenas que especifica las entidades que se van a filtrar.

Configuración	Descripción
<p>Patrones de inclusión</p> <ul style="list-style-type: none">• inclusionDocumentFileTypePatterns• inclusionDocumentFileNamePatterns• inclusionAccountFileTypePatterns• inclusionCampaignFileTypePatterns• inclusionDocumentFileNamePatterns• inclusionCampaignFileNamePatterns• inclusionCaseFileTypePatterns• inclusionCaseFileNamePatterns• inclusionContactFileTypePatterns• inclusionContractFileNamePatterns• inclusionLeadFileTypePatterns• inclusionLeadFileNamePatterns• inclusionOpportunityFileTypePatterns• inclusionOpportunityFileNamePatterns• inclusionSolutionFileTypePatterns• inclusionSolutionFileNamePatterns• inclusionTaskFileTypePatterns• inclusionTaskFileNamePatterns• inclusionGroupFileTypePatterns• inclusionGroupFileNamePatterns• inclusionChatterFileTypePatterns• inclusionChatterFileNamePatterns• inclusionCustomEntityFileTypePatterns• inclusionCustomEntityFileNamePatterns	<p>Una lista de patrones de expresiones regulares para incluir ciertos archivos de su fuente de datos de Salesforce. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.</p>

Configuración	Descripción
<p>Patrones de exclusión</p> <ul style="list-style-type: none">• exclusionDocumentFileTypePatterns• exclusionDocumentFileNamePatterns• exclusionAccountFileTypePatterns• exclusionCampaignFileTypePatterns• exclusionCampaignFileNamePatterns• exclusionCaseFileTypePatterns• exclusionCaseFileNamePatterns• exclusionContactFileTypePatterns• exclusionContractFileNamePatterns• exclusionLeadFileTypePatterns• exclusionLeadFileNamePatterns• exclusionOpportunityFileTypePatterns• exclusionOpportunityFileNamePatterns• exclusionSolutionFileTypePatterns• exclusionSolutionFileNamePatterns• exclusionTaskFileTypePatterns• exclusionTaskFileNamePatterns• exclusionGroupFileTypePatterns• exclusionGroupFileNamePatterns• exclusionChatterFileTypePatterns• exclusionChatterFileNamePatterns• exclusionCustomEntityFileTypePatterns• exclusionCustomEntityFileNamePatterns	<p>Una lista de patrones de expresiones regulares para excluir ciertos archivos de su fuente de datos de Salesforce. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y otro de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.</p>

Configuración	Descripción
<ul style="list-style-type: none"> • isCrawlAccount • isCrawlContact • isCrawlCase • isCrawlCampaign • isCrawlProduct • isCrawlLead • isCrawlContract • isCrawlPartner • isCrawlProfile • isCrawlIdea • isCrawlPricebook • isCrawlDocument • crawlSharedDocument • isCrawlGroup • isCrawlOpportunity • isCrawlChatter • isCrawlUser • isCrawlSolution • isCrawlTask • isCrawlAccountAdjuntos • isCrawlContactAdjuntos • isCrawlCaseAdjuntos • isCrawlCampaignAdjuntos • isCrawlLeadAdjuntos • isCrawlContractAdjuntos • isCrawlGroupAdjuntos • isCrawlOpportunityAdjuntos • isCrawlChatterAdjuntos • isCrawlSolutionAdjuntos • isCrawlTaskAdjuntos • isCrawlCustomEntityAttachments • isCrawlKnowledgeArtículos <ul style="list-style-type: none"> • isCrawlDraft • isCrawlPublish • isCrawlArchived 	truepara indexar los archivos correspondientes en su cuenta de Salesforce.
type	El tipo de fuente de datos. EspecificarSALESFORCEV2como tipo de fuente de datos.

Configuración	Descripción
enableIdentityCrawler	truepara activar el rastreador de identidades. El rastreador de identidades está activado de forma predeterminada. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el PutPrincipalMapping API. Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte Filtrado en contexto de usuario .
Modo de sincronización	Especifique siAmazon Kendradebe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre: <ul style="list-style-type: none"> • FORCED_FULL_CRAWLpara volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice • FULL_CRAWLpara rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice • CHANGE_LOGpara rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice
versión	La versión de esta plantilla que se admite actualmente.

Esquema JSON de Salesforce

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties":
  {
    "connectionConfiguration": {
      "type": "object",
      "properties":
      {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties":
          {
            "hostUrl":
            {
              "type": "string",
              "pattern": "https:.*"
            }
          }
        }
      }
    }
  }
}
```

```
        },
        "required":
        [
            "hostUrl"
        ]
    }
},
"required":
[
    "repositoryEndpointMetadata"
]
},
"repositoryConfigurations": {
    "type": "object",
    "properties":
    {
        "account":
        {
            "type": "object",
            "properties":
            {
                "fieldMappings":
                {
                    "type": "array",
                    "items":
                    [
                        {
                            "type": "object",
                            "properties":
                            {
                                "indexFieldName":
                                {
                                    "type": "string"
                                },
                                "indexFieldType":
                                {
                                    "type": "string",
                                    "enum":
                                    [
                                        "STRING",
                                        "STRING_LIST",
                                        "DATE",
                                        "LONG"
                                    ]
                                },
                                "dataSourceFieldName":
                                {
                                    "type": "string"
                                },
                                "dateFieldFormat":
                                {
                                    "type": "string",
                                    "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
                                }
                            },
                            "required":
                            [
                                "indexFieldName",
                                "indexFieldType",
                                "dataSourceFieldName"
                            ]
                        }
                    ]
                }
            }
        },
        "required":
        [
            "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",  
            }  
          }  
        }  
      ]  
    }  
  }  
},  
"type": "array",  
"items":  
[  
  {  
    "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'"  
                        }  
                    }  
                }  
            ]  
        }  
    }  
}
```

```
        "type": "string",
        "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
    },
    "required": [
        "indexFieldName",
        "indexFieldType",
        "dataSourceFieldName"
    ]
},
],
"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"  
        ]  
    }  
},  
"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"  
]
```

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

```
        "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"
                ]
              },
              "dataSourceFieldName": {
                "type": "string"
              },
              "dateFieldFormat": {
                "type": "string",
                "pattern": "yyyy-MM-dd'T'HH:mm:ss'Z'"
              }
            },
            "required": [
              "indexFieldName",
              "indexFieldType",
              "dataSourceFieldName"
            ]
          }
        ]
      }
    }
  }
}

```

```
"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"  
            }  
        }  
    }  
}
```

```
        "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"
                },
                "isCrawlIndex": {
                    "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"
        }
    },
    "exclusionContractFileNamePatterns":{
        "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"
        }
    }
```

```
        "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"  
    }  
}
```

```
{  
    "type": "string"  
}  
},  
"exclusionChatterFileNamePatterns":{  
    "type": "array",  
    "items":  
    {  
        "type": "string"  
    }  
},  
"inclusionCustomEntityTypePatterns":{  
    "type": "array",  
    "items":  
    {  
        "type": "string"  
    }  
},  
"exclusionCustomEntityTypePatterns":{  
    "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"
    ]
  }
}

```

ServiceNowesquema de plantilla

Incluye un JSON que contiene el esquema de la fuente de datos como parte del [TemplateConfiguration](#) objeto. Usted proporciona el ServiceNow la URL del host, el tipo de autenticación y la versión de la instancia como parte de la configuración de la conexión o los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como SERVICENOWV2, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique TEMPLATE como elType cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [ServiceNowEsquema JSON \(p. 269\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
URL del servidor	El ServiceNow URL del servidor. Por ejemplo, <i>su-dominio.service-now.com</i> .
authType	El tipo de autenticación que utiliza, ya sea basicAuth o OAuth2.
servicenowInstanceVersion	El ServiceNow versión que está utilizando. Puedes elegir entre Tokio, Sandiego, Roma y otros.
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none"> Artículo de conocimiento attachment Catálogo de servicios incidente 	Una lista de objetos que mapean los atributos o nombres de campo de su ServiceNow artículos de conocimiento, archivos adjuntos, catálogo de servicios e incidentes para Amazon Kendra nombres de campos de índice. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos). El ServiceNow los nombres de los campos de la fuente de datos deben existir en su ServiceNow metadatos personalizados.

Configuración	Descripción
propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos.
<ul style="list-style-type: none"> • knowledgeArticleFilter • incidentQueryFilter • serviceCatalogQueryFiltrar • knowledgeArticleTitleRegExp • serviceCatalogTitleRegExp • incidentTitleRegExp. • inclusionFileTypePatrones • exclusionFileTypePatrones • inclusionFileNamePatrones • exclusionFileNamePatrones • incidentStateType 	Una lista de patrones de expresiones regulares para incluir o excluir ciertos archivos en suServiceNowfuente de datos. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
<ul style="list-style-type: none"> • isCrawlKnowledgeArtículo • isCrawlKnowledgeArticleAttachment • includePublicArticlesSolo • isCrawlServiceCatalogo • isCrawlServiceCatalogAttachment • isCrawlActiveServiceCatalog • isCrawlInactiveServiceCatalog • isCrawlIncident • isCrawlIncidentAdjuntivo • isCrawlActiveIncidente • isCrawlInactiveIncidente • Aplicar ACLForKnowledgeArticle • Aplicar ACLForServiceCatalog • Aplicar ACLForIncident 	trueindexarServiceNowartículos de conocimiento, catálogos de servicios, incidentes y archivos adjuntos.
type	El tipo de fuente de datos. EspecificarSERVICENOWV2como tipo de fuente de datos.
enableIdentityCrawler	truepara activar el rastreador de identidades. El rastreador de identidades está activado de forma predeterminada. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el PutPrincipalMappingAPI . Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte Filtrado en contexto de usuario .

Configuración	Descripción
Modo de sincronización	<p>Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:</p> <ul style="list-style-type: none"> • FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice • FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice • CHANGE_LOG para rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice
SecretarioN	<p>El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a su ServiceNow. El secreto debe contener una estructura JSON con las siguientes claves:</p> <pre>{ "username": "user name", "password": "password" }</pre> <p>Si usas la autenticación OAuth2, tu secreto debe contener una estructura JSON con las siguientes claves:</p> <pre>{ "username": "user name", "password": "password", "clientId": "client id", "clientSecret": "client secret" }</pre>
versión	La versión de la plantilla que se admite actualmente.

ServiceNow Esquema JSON

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
      "type": "object",
      "properties": {
        "repositoryEndpointMetadata": {
          "type": "object",
          "properties": {
            "hostUrl": {
              "type": "string",
              "pattern": "^(?!(^https?|ftp|file):\/\/)[a-z0-9-]+.service-now.com$",
              "minLength": 1,
              "maxLength": 2048
            },
            "authType": {
              "type": "string",
              "enum": [
                "BasicAuth"
              ]
            }
          }
        }
      }
    }
  }
}
```

```
        "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": {
    "isCrawlKnowledgeArticle": {
        "type": "boolean"
    },
    "isCrawlKnowledgeArticleAttachment": {
        "type": "boolean"
    },
    "includePublicArticlesOnly": {
        "type": "boolean"
    },
    "knowledgeArticleFilter": {
        "type": "string"
    },
    "incidentQueryFilter": {
        "type": "string"
    },
    "serviceCatalogQueryFilter": {
        "type": "string"
    },
    "isCrawlServiceCatalog": {
        "type": "boolean"
    },
    "isCrawlServiceCatalogAttachment": {
        "type": "boolean"
    },
    "isCrawlActiveServiceCatalog": {
        "type": "boolean"
    },
    "isCrawlInactiveServiceCatalog": {
        "type": "boolean"
    },
    "isCrawlIncident": {
        "type": "boolean"
    },
    "isCrawlIncidentAttachment": {
        "type": "boolean"
    },
    "isCrawlActiveIncident": {
        "type": "boolean"
    },
    "isCrawlInactiveIncident": {
        "type": "boolean"
    },
    "applyACLForKnowledgeArticle": {
        "type": "boolean"
    },
    "applyACLForServiceCatalog": {
        "type": "boolean"
    },
    "applyACLForIncident": {
        "type": "boolean"
    },
    "incidentStateType": {
        "type": "array",
        "items": {
            "type": "string",
            "enum": [
                "Open",
                "Open - Unassigned",
                "Resolved",
                "All"
            ]
        }
    },
    "knowledgeArticleTitleRegExp": {
```

```
        "type": "string"
    },
    "serviceCatalogTitleRegExp": {
        "type": "string"
    },
    "incidentTitleRegExp": {
        "type": "string"
    },
    "inclusionFileTypePatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionFileTypePatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "inclusionFileNamePatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    },
    "exclusionFileNamePatterns": {
        "type": "array",
        "items": {
            "type": "string"
        }
    }
},
"required": []
},
"type": {
    "type": "string",
    "pattern": "SERVICENOWV2"
},
"enableIdentityCrawler": {
    "type": "boolean"
},
"syncMode": {
    "type": "string",
    "enum": [
        "FORCED_FULL_CRAWL",
        "FULL_CRAWL"
    ]
},
"secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
},
"version": {
    "type": "string",
    "anyOf": [
        {
            "pattern": "1.0.0"
        }
    ]
},
"required": [
    "connectionConfiguration",
    "repositoryConfigurations",

```

```

    "syncMode",
    "additionalProperties",
    "secretArn",
    "type"
]
}

```

Esquema de plantillas de Zendesk

Incluye un JSON que contiene el esquema de la fuente de datos como parte de [TemplateConfiguration](#) objeto. La URL del host se proporciona como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Especifique también el tipo de fuente de datos como `ZENDESK`, un secreto para sus credenciales de autenticación y otras configuraciones necesarias. A continuación, especifique `TEMPLATE` como el `Tecleacuando` cuando llamas [CreateDataSource](#).

Puede utilizar la plantilla que se proporciona en esta guía para desarrolladores. Consulte [Esquema JSON de Zendesk \(p. 276\)](#).

A continuación se proporciona información sobre las claves JSON importantes que se deben configurar.

Configuración	Descripción
Configuración de conexión	Información de configuración del punto final de la fuente de datos.
repositoryEndpointMetadata	La información del punto final de la fuente de datos.
URL del servidor	La URL del host de Zendesk. Por ejemplo, <code>https://yoursubdomain.zendesk.com</code> .
Configuraciones de repositorios	Información de configuración del contenido de la fuente de datos. Por ejemplo, configurar tipos específicos de contenido y mapeos de campos.
<ul style="list-style-type: none"> • billete • Comentario sobre el billete • ticketCommentAttachment • article • Comentario del artículo • Artículo adjunto • Tema comunitario • communityPostComment 	Una lista de objetos que asignan atributos o nombres de campo de los tickets de Zendesk a los nombres de los campos de índice de Amazon Kendra. Para obtener más información, consulte Mapping data source fields (Asignación de campos de origen de datos).
SecretarioN	El nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto que contiene los pares clave-valor necesarios para conectarse a su cuenta de Zendesk. El secreto debe contener una estructura JSON con las siguientes claves: URL de host, ID de cliente, secreto de cliente, nombre de usuario y contraseña.
Propiedades adicionales	Opciones de configuración adicionales para el contenido de la fuente de datos
organizationNameFilter	Puede optar por indexar los tickets que existen dentro de un determinadoOrganización.

Configuración	Descripción
Desde la fecha	Puede optar por configurar <code>unsinceDate</code> parámetro para que el conector de Zendesk rastree el contenido en función de un parámetro específicos <code>sinceDate</code> .
Patrones de inclusión	Una lista de patrones de expresiones regulares para incluir ciertos archivos de su fuente de datos de Zendesk. Los archivos que coinciden con los patrones se incluyen en el índice. Los archivos que no coincidan con los patrones se excluyen del índice. Si un archivo coincide con un patrón de inclusión y otro de exclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
Patrones de exclusión	Una lista de patrones de expresiones regulares para excluir ciertos archivos de su fuente de datos de Zendesk. Los archivos que coinciden con los patrones se excluyen del índice. Los archivos que no coinciden con los patrones se incluyen en el índice. Si un archivo coincide con un patrón de exclusión y otro de inclusión, el patrón de exclusión tiene prioridad y el archivo no se incluye en el índice.
<ul style="list-style-type: none"> • <code>isCrawlTicket</code> • <code>isCrawlTicketComentario</code> • <code>isCrawlTicketCommentAttachment</code> • <code>isCrawlArticle</code> • <code>isCrawlArticleComentario</code> • <code>isCrawlArticleAdjuntivo</code> • <code>isCrawlCommunityTema</code> • <code>isCrawlCommunityPublicar</code> • <code>isCrawlCommunityPostComment</code> 	Entrada <code>true</code> para indexar el contenido.
<code>type</code>	Especificar <code>ZENDESK</code> como tipo de fuente de datos.
<code>useChangeLog</code>	Entrada <code>true</code> para usar el registro de cambios de Zendesk para determinar qué documentos deben actualizarse en el índice. Según el tamaño del registro de cambios, puede ser más rápido escanear los documentos en Zendesk. Si sincroniza su fuente de datos de Zendesk con su índice por primera vez, se escanean todos los documentos.

Esquema JSON de Zendesk

```
{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "type": "object",
  "properties": {
    "connectionConfiguration": {
```

```
"type": "object",
"properties": {
    "repositoryEndpointMetadata": {
        "type": "object",
        "properties": {
            "hostUrl": {
                "type": "string",
                "pattern": "https:.*"
            }
        },
        "required": [
            "hostUrl"
        ]
    }
},
"required": [
    "repositoryEndpointMetadata"
]
},
"repositoryConfigurations": {
    "type": "object",
    "properties": {
        "ticket": {
            "type": "object",
            "properties": {
                "fieldMappings": {
                    "type": "array",
                    "items": {
                        "anyOf": [
                            {
                                "type": "object",
                                "properties": {
                                    "indexFieldName": {
                                        "type": "string"
                                    },
                                    "indexFieldType": {
                                        "type": "string",
                                        "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                                    },
                                    "dataSourceFieldName": {
                                        "type": "string"
                                    },
                                    "dateFieldFormat": {
                                        "type": "string",
                                        "pattern": "dd-MM-yyyy HH:mm:ss"
                                    }
                                },
                                "required": [
                                    "indexFieldName",
                                    "indexFieldType",
                                    "dataSourceFieldName"
                                ]
                            }
                        ]
                    }
                }
            }
        },
        "required": [
            "fieldMappings"
        ]
    },
    "ticketComment": {
        "type": "object",
        "properties": {
            "fieldMappings": {

```

```
"type": "array",
"items": [
  "anyOf": [
    {
      "type": "object",
      "properties": {
        "indexFieldName": {
          "type": "string"
        },
        "indexFieldType": {
          "type": "string",
          "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
        },
        "dataSourceFieldName": {
          "type": "string"
        },
        "dateFieldFormat": {
          "type": "string",
          "pattern": "dd-MM-yyyy HH:mm:ss"
        }
      },
      "required": [
        "indexFieldName",
        "indexFieldType",
        "dataSourceFieldName"
      ]
    }
  ]
},
"required": [
  "fieldMappings"
]
},
"ticketCommentAttachment": {
  "type": "object",
  "properties": {
    "fieldMappings": {
      "type": "array",
      "items": [
        "anyOf": [
          {
            "type": "object",
            "properties": {
              "indexFieldName": {
                "type": "string"
              },
              "indexFieldType": {
                "type": "string",
                "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
              },
              "dataSourceFieldName": {
                "type": "string"
              },
              "dateFieldFormat": {
                "type": "string",
                "pattern": "dd-MM-yyyy HH:mm:ss"
              }
            },
            "required": [
              "indexFieldName",
              "indexFieldType",
              "dataSourceFieldName"
            ]
          }
        ]
      ]
    }
  }
}
```

```
        }
      ]
    }
  ],
  "required": [
    "fieldMappings"
  ]
},
"article": {
  "type": "object",
  "properties": {
    "fieldMappings": {
      "type": "array",
      "items": {
        "anyOf": [
          {
            "type": "object",
            "properties": {
              "indexFieldName": {
                "type": "string"
              },
              "indexFieldType": {
                "type": "string",
                "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
              },
              "dataSourceFieldName": {
                "type": "string"
              },
              "dateFieldFormat": {
                "type": "string",
                "pattern": "dd-MM-yyyy HH:mm:ss"
              }
            },
            "required": [
              "indexFieldName",
              "indexFieldType",
              "dataSourceFieldName"
            ]
          }
        ]
      }
    }
  ],
  "required": [
    "fieldMappings"
  ]
},
"communityPostComment": {
  "type": "object",
  "properties": {
    "fieldMappings": {
      "type": "array",
      "items": {
        "anyOf": [
          {
            "type": "object",
            "properties": {
              "indexFieldName": {
                "type": "string"
              },
              "indexFieldType": {
                "type": "string",
                "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
              },
              "dataSourceFieldName": {
                "type": "string"
              }
            }
          }
        ]
      }
    }
  ],
  "required": [
    "fieldMappings"
  ]
}
```

```
        "type": "string"
    },
    "dateFieldFormat": {
        "type": "string",
        "pattern": "dd-MM-yyyy HH:mm:ss"
    }
},
"required": [
    "indexFieldName",
    "indexFieldType",
    "dataSourceFieldName"
]
}
]
}
},
"required": [
    "fieldMappings"
]
},
"articleComment": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": {
                "anyOf": [
                    {
                        "type": "object",
                        "properties": {
                            "indexFieldName": {
                                "type": "string"
                            },
                            "indexFieldType": {
                                "type": "string",
                                "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
                            },
                            "dataSourceFieldName": {
                                "type": "string"
                            },
                            "dateFieldFormat": {
                                "type": "string",
                                "pattern": "dd-MM-yyyy HH:mm:ss"
                            }
                        },
                        "required": [
                            "indexFieldName",
                            "indexFieldType",
                            "dataSourceFieldName"
                        ]
                    }
                ]
            }
        },
        "required": [
            "fieldMappings"
        ]
    }
},
"articleAttachment": {
    "type": "object",
    "properties": {
        "fieldMappings": {
            "type": "array",
            "items": {
```

```
"anyOf": [
  {
    "type": "object",
    "properties": {
      "indexFieldName": {
        "type": "string"
      },
      "indexFieldType": {
        "type": "string",
        "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
      },
      "dataSourceFieldName": {
        "type": "string"
      },
      "dateFieldFormat": {
        "type": "string",
        "pattern": "dd-MM-yyyy HH:mm:ss"
      }
    },
    "required": [
      "indexFieldName",
      "indexFieldType",
      "dataSourceFieldName"
    ]
  }
],
"required": [
  "fieldMappings"
]
},
"communityTopic": {
  "type": "object",
  "properties": {
    "fieldMappings": {
      "type": "array",
      "items": {
        "anyOf": [
          {
            "type": "object",
            "properties": {
              "indexFieldName": {
                "type": "string"
              },
              "indexFieldType": {
                "type": "string",
                "enum": ["STRING", "STRING_LIST", "LONG", "DATE"]
              },
              "dataSourceFieldName": {
                "type": "string"
              },
              "dateFieldFormat": {
                "type": "string",
                "pattern": "dd-MM-yyyy HH:mm:ss"
              }
            },
            "required": [
              "indexFieldName",
              "indexFieldType",
              "dataSourceFieldName"
            ]
          }
        ]
      }
    }
  }
}
```

```
        }
      },
      "required": [
        "fieldMappings"
      ]
    }
  },
  "secretArn": {
    "type": "string",
    "minLength": 20,
    "maxLength": 2048
  },
  "additionalProperties": {
    "type": "object",
    "properties": {
      "organizationNameFilter": {
        "type": "array"
      },
      "sinceDate": {
        "type": "string",
        "pattern": "^[0-9]{4}-[0-9]{2}-[0-9]{2} [0-9]{2}:[0-9]{2}:[0-9]{2}$"
      },
      "inclusionPatterns": {
        "type": "array"
      },
      "exclusionPatterns": {
        "type": "array"
      },
      "isCrawTicket": {
        "type": "string"
      },
      "isCrawTicketComment": {
        "type": "string"
      },
      "isCrawTicketCommentAttachment": {
        "type": "string"
      },
      "isCrawlArticle": {
        "type": "string"
      },
      "isCrawlArticleAttachment": {
        "type": "string"
      },
      "isCrawlArticleComment": {
        "type": "string"
      },
      "isCrawlCommunityTopic": {
        "type": "string"
      },
      "isCrawlCommunityPost": {
        "type": "string"
      },
      "isCrawlCommunityPostComment": {
        "type": "string"
      }
    }
  },
  "type": {
    "type": "string",
    "pattern": "ZENDESK"
  },
  "useChangeLog": {
    "type": "string",
    "enum": ["true", "false"]
}
}
```

```
  },
  "version": {
    "type": "string",
    "anyOf": [
      {
        "pattern": "1.0.0"
      }
    ]
  },
  "additionalProperties": false,
  "required": [
    "connectionConfiguration",
    "repositoryConfigurations",
    "additionalProperties",
    "useChangeLog",
    "secretArn",
    "type"
  ]
}
```

Adobe Experience Manager

Adobe Experience Manager es un sistema de administración de contenido que se utiliza para crear contenido para sitios web o aplicaciones móviles. Puedes usar Amazon Kendra para conectarse a Adobe Experience Manager e indexar sus páginas y activos de contenido.

Amazon Kendra soporta Adobe Experience Manager (AEM) como instancia de autor de Cloud Service y Adobe Experience Manager Instancia de creación y publicación local.

Puedes conectarte a Amazon Kendra a tu Adobe Experience Manager fuente de datos mediante [Amazon Kendra console](#) o el [Template Configuration API](#).

Para solucionar problemas de Amazon Kendra Conector de fuentes de datos de Adobe Experience Manager, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 283\)](#)
- [Requisitos previos \(p. 283\)](#)
- [Instrucciones de conexión \(p. 285\)](#)

Características admitidas

Adobe Experience Manager el conector de fuente de datos admite las siguientes funciones:

- Mapeos de campo
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronización completa/ Sincronización de contenido nuevo y modificado/ Sincronización de contenido nuevo, modificado y eliminado
- Filtrado de contexto de usuario
- OAuth 2.0 y autenticación básica

Requisitos previos

Antes de poder usar Amazon Kendra para indexar sus Adobe Experience Manager fuente de datos, realice estos cambios en su Adobe Experience Manager AWS cuentas.

En Adobe Experience Manager, asegúrate de tener:

- Acceso a una cuenta con privilegios de administrador o a un usuario administrador.
- Copié tu Adobe Experience Manager URL del servidor.

Note

(En las instalaciones o en el servidor) Amazon Kendra comprueba si la información del punto final está incluida en AWS Secrets Manager es la misma que la información de punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a proteger contra la [problema de diputado confuso](#), que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, sino que usa Amazon Kendra como proxy para acceder al secreto configurado y realizar la acción. Si posteriormente cambias la información de tu terminal, debes crear un nuevo secreto para sincronizar esta información.

- Tomó nota de sus credenciales de autenticación básicas del nombre de usuario y la contraseña del administrador.
- Opcional: Credenciales de OAuth 2.0 generadas en Adobe Experience Manager (AEM) como servicio en la nube o AEM On-Premise. Si usa AEM On-Premise, las credenciales incluyen el ID del cliente, el secreto del cliente y la clave privada. Si usa AEM como servicio en la nube, las credenciales incluyen el ID de cliente, el secreto del cliente, la clave privada, el ID de la organización, el ID de cuenta técnica y Adobe Identity Management System host (IMS). Para obtener más información sobre cómo generar estas credenciales para AEM como servicio en la nube, consulte [Adobe Experience Manager documentación](#). Para AEM On-Premise, implementación del servidor Adobe Granite OAuth 2.0 (com.adobe.granite.oauth.server) proporciona soporte para las funcionalidades del servidor OAuth 2.0 en AEM.
- Se comprobó que cada documento es único en Adobe Experience Manager y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que desee utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu cuenta de AWS, asegúrate de tener:

- [Creó un Amazon Kendra índice](#), si usa la API, anota el ID del índice.
- [Creó un IAM papel](#) para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al correcto AWS Secrets Manager identificación secreta.

- Ha guardado sus credenciales de autenticación de Adobe Experience Manager en un AWS Secrets Manager secreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendable reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o secreto, puedes usar la consola para crear un nuevo IAM rol y Secrets Manager secreto al conectar la fuente de datos de Adobe Experience Manager a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existente IAM rol y Secrets Manager secreto y un identificador de índice.

Instrucciones de conexión

Para conectar Amazon Kendra a tu Adobe Experience Manager fuente de datos, debe proporcionar los detalles necesarios de su Adobe Experience Manager fuente de datos para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Adobe Experience Manager por Amazon Kendra, consulte [Requisitos previos \(p. 283\)](#).

Console

Para conectar Amazon Kendra a Adobe Experience Manager

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, selecciona Índices y, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar su Control de acceso de usuarios ajustes en Ajustes de índice.

3. En el Primeros pasospágina, elige Agregar fuente de datos.
4. En el Agregar fuente de datospágina, elige Conector de Adobe Experience Managery, a continuación, elija Agregar fuente de datos.
5. En el Especificar los detalles de la fuente de datospágina, introduzca la siguiente información:
 - a. En Nombre y descripción, para Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional) Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, para Idioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, para Añadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tus AWS costos.
 - e. Elija Siguiente.
6. En el Defina el acceso y la seguridadpágina, introduzca la siguiente información:
 - a. Fuente—Elija uno de los dos AEM local o AEM como servicio en la nube.

Introduce tu Adobe Experience Manager URL del servidor. Por ejemplo, si usa AEM On-Premise, incluye el nombre de host y el puerto: https://hostname:port. O bien, si usa AEM como servicio en la nube, puede usar la URL del autor: https://author-xxxxxx-xxxxxx.adobeaecloud.com.
 - b. Ubicación del certificado SSL—Introduzca la ruta al certificado SSL almacenado en un Amazon S3 balde. Se usa para conectarse a AEM On-Premise con una conexión SSL segura.
 - c. autenticación—Elige Autenticación básica o Autenticación OAuth 2.0. A continuación, elija una existente AWS Secrets Manager secreto o crea un nuevo secreto para guardar tu Adobe Experience Manager credenciales. Si eliges crear un secreto nuevo, un AWS Secrets Manager se abre una ventana secreta.

Si elegiste Autenticación básica, introduzca un nombre para el secreto, el Adobe Experience Manager nombre de usuario y contraseña del sitio. El usuario debe tener permiso de administrador o ser un usuario administrador.

Si elegiste Autenticación OAuth 2.0 y utilice AEM On-Premise, introduzca un nombre para el secreto, el ID de cliente, el secreto del cliente y la clave privada. Si usa AEM como servicio en la nube, introduzca un nombre para el secreto, el ID de cliente, el secreto del cliente, la clave

privada, el ID de la organización, el ID de la cuenta técnica yAdobe Identity Management Systemhost (IMS).

Seleccione Guardar.

- d. Rastreador de identidades—Elija rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos y guardarla enAmazon Kendra tienda principal/de identidad. Esto resulta útil para filtrar el contexto del usuario, donde los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos.
- e. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesGrupos de seguridad de VPC.
- f. IAMpapel—Elige uno existenteIAMrol o crea uno nuevoIAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- g. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. Alcance de sincronización—Establezca límites para rastrear ciertos tipos de contenido, componentes de página y rutas de raíces, y filtre el contenido mediante patrones de expresiones regulares.
 - i. Tipos de contenido—Elija si desea rastrear solo las páginas o los activos, o ambos.
 - ii. (Opcional)Configuración adicional—Configure los siguientes parámetros:
 - Componentes de página: los nombres específicos de los componentes de la página. El componente de página es un componente de página extensible diseñado para funcionar conAdobe Experience Managereditor de plantillas y permite ensamblar los componentes del encabezado, pie de página y estructura con el editor de plantillas.
 - Variaciones de fragmentos de contenido—Los nombres específicos de las variaciones de los fragmentos de contenido. Los fragmentos de contenido le permiten diseñar, crear, seleccionar y publicar contenido independiente de la página enAdobe Experience Manager. Permiten preparar contenido listo para su uso en múltiples ubicaciones/en múltiples canales.
 - Rutas raíz—Las rutas raíz a un contenido específico.
 - Patrones de expresiones regulares—Los patrones de expresiones regulares para incluir o excluir determinadas páginas y recursos.
 - b. Modo de sincronización—Elija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Cuando sincronizas tu fuente de datos conAmazon Kendrapor primera vez, todo el contenido se sincroniza de forma predeterminada.
 - Sincronización completa—Sincronice todo el contenido independientemente del estado de sincronización anterior.
 - Sincronización de documentos nuevos o modificados—Sincronice solo los documentos nuevos y modificados.
 - Sincronización de documentos nuevos, modificados o eliminados—Sincronice solo los documentos nuevos, modificados y eliminados.
 - c. ID de zona horaria—Si usa AEM On-Premise y la zona horaria de su servidor es diferente a la zona horaria delAmazon KendraConector o índice de AEM, puede especificar la zona horaria del servidor para alinearla con el conector o índice de AEM. La zona horaria predeterminada de AEM On-Premise es la zona horaria delAmazon KendraConector o índice

- AEM. La zona horaria predeterminada para AEM como servicio en la nube es la hora del meridiano de Greenwich.
- d. Cronograma de ejecución de sincronización—ParaFrecuencia, elige con qué frecuenciaAmazon Kendrase sincronizará con su fuente de datos.
 - e. Elija Siguiente.
8. En elDefinir mapeos de campospágina, introduzca la siguiente información:
 - a. Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice. Para agregar campos de fuente de datos personalizados, cree un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - b. Elija Siguiente.
 9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon KendraaAdobe Experience Manager

Debe especificar un JSON del[esquema de fuente de datos](#)utilizando el[TemplateConfigurationAPI](#).
Debe proporcionar la siguiente información:

- Fuente de datos—Especifique la fuente de datos comoAEM.
- URL del host de AEM—Especifique elAdobe Experience ManagerURL del servidor. Por ejemplo, si usa AEM On-Premise, incluye el nombre de host y el puerto:<https://hostname:port>. O bien, si usa AEM como servicio en la nube, puede usar la URL del autor:<https://author-xxxxxx-xxxxxx.adobeaecloud.com>.
- Tipo de autenticación—Especifique el tipo de autenticación que desea utilizar, ya seaBasicoOAuth2.
- Tipo de AEM—Especifica qué tipo deAdobe Experience Managerusas, ya seaCLOUDoON_PREMISE.
- Teclea—EspecificarTEMPLATEcomo el tipo cuando llamasCreateDataSource.
- Nombre secreto de recurso de Amazon (ARN)—Si desea utilizar la autenticación básica para AEM On-Premise o Cloud, proporcione un secreto que almacene las credenciales de autenticación de su nombre de usuario y contraseña. Usted proporciona el nombre de recurso de Amazon (ARN) de unAWS Secrets Managersecreto. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
  "aemUrl": "Adobe Experience Manager On-Premise host URL",  
  "username": "user name with admin permissions",  
  "password": "password with admin permissions"  
}
```

Si desea utilizar la autenticación OAuth 2.0 para AEM On-Premise, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
  "aemUrl": "Adobe Experience Manager host URL",  
  "clientId": "client ID",  
  "clientSecret": "client secret",  
  "privateKey": "private key"  
}
```

Si desea utilizar la autenticación OAuth 2.0 para AEM como servicio en la nube, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "clientId": "client ID",  
    "clientSecret": "client secret",  
    "privateKey": "private key",  
    "orgId": "organization ID",  
    "technicalAccountId": "technical account ID",  
    "imsHost": "Adobe Identity Management System (IMS) host"  
}
```

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSource para proporcionar un IAMrol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas necesarias para el conector de Adobe Experience Manager y Amazon Kendra. Para obtener más información, consulte [IAMfunciones para las fuentes de datos de Adobe Experience Manager](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfiguration cuando llamasCreateDataSource. Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).
- Modo de sincronización—Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puede elegir entre las siguientes opciones:
 - FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice.
 - FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice.
 - CHANGE_LOG para rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice.
- ID de zona horaria—Si usa AEM On-Premise y la zona horaria de su servidor es diferente a la zona horaria del Amazon Kendra Conector o índice de AEM, puede especificar la zona horaria del servidor para alinearla con el conector o índice de AEM.

La zona horaria predeterminada de AEM On-Premise es la zona horaria del Amazon Kendra Conector o índice AEM. La zona horaria predeterminada para AEM como servicio en la nube es la hora del meridiano de Greenwich.

Para obtener información sobre los identificadores de zonas horarias admitidas, consulte [Adobe Experience Manager Esquema JSON](#).

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinadas páginas y recursos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Rastreador de identidades—Especifica si se va a activar Amazon Kendra rastreador de identidad. Si el rastreador de identidades está desactivado, debe cargar la información de identidad o principal mediante el [PutPrincipalMapping](#) API. Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados

de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario](#).

- Mapeos de campo—Elija asignar los campos de la fuente de datos de Adobe Experience Manager a su Amazon Kendracampos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [Adobe Experience Manager esquema de plantilla](#).

Alfresco

Alfrescoes un servicio de administración de contenido que ayuda a los clientes a almacenar y administrar su contenido. Se puede utilizar Amazon Kendra para indexar la biblioteca de Alfresco documentos, la wiki y el blog.

Amazon Kendraes compatible con Alfresco entornos locales y Alfresco en la nube (plataforma como servicio).

Puede conectarse Amazon Kendra a la fuente Alfresco de datos mediante la [Amazon Kendraconsola](#) o la [TemplateConfigurationAPI](#).

Para solucionar problemas con el conector de fuente de datos de Amazon Kendra Alfresco, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 289\)](#)
- [Requisitos previos \(p. 289\)](#)
- [Instrucciones de conexión \(p. 290\)](#)
- [Más información \(p. 294\)](#)

Características admitidas

Amazon KendraAlfrescoel conector de fuente de datos admite las siguientes funciones:

- Mapeos de campo
- Filtros de inclusión/exclusión
- Sincronización completa/ Sincronización de contenido nuevo, modificado y eliminado
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Filtrado de contexto de usuario
- OAuth 2.0 y autenticación básica

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de datos de Alfresco, realice estos cambios en su Alfresco yCuentas de AWS.

EnAlfresco, asegúrese de tener:

- Copió la URL del Alfresco repositorio y la URL de la aplicación web. Si solo quieres indexar un Alfresco sitio específico, copia también el ID del sitio.

- Tomó nota Alfresco de sus credenciales de autenticación, que incluyen un nombre de usuario y una contraseña con al menos permisos de lectura. Si desea utilizar la autenticación OAuth 2.0, debe añadir el usuario al grupo de Alfresco administradores.
- Opcional: se generaron credenciales de OAuth 2.0 en Alfresco. Las credenciales incluyen el ID del cliente, el secreto del cliente y la URL del token. Para obtener más información sobre cómo configurar clientes para entornos Alfresco locales, consulte la documentación de [Alfresco](#). Si usa Alfresco Cloud (PaaS), debe ponerse en contacto con el [soporte de Hyland](#) para la autenticación Alfresco OAuth 2.0.
- Se ha comprobado que cada documento es único en Alfresco y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En su cuenta de AWS, asegúrese de tener:

- [Creó un Amazon Kendra índice](#) y, si utilizaba la API, anotó el ID del índice.
- [Creó una IAM función](#) para la fuente de datos y, si utilizaba la API, anotó el ARN de la IAM función.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al identificador AWS Secrets Manager secreto correcto.

- Almacenó sus credenciales de autenticación de Alfresco en AWS Secrets Manager secreto y, si utilizaba la API, anotó el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. No se recomienda reutilizar las credenciales y los secretos en las fuentes de datos ni en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o un secreto existentes, puede usar la consola para crear un IAM rol y un Secrets Manager secreto nuevos al conectar su fuente de datos de Alfresco a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes, así como un identificador de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Alfresco, debe proporcionar los detalles necesarios de su fuente de datos de Alfresco para Amazon Kendra poder acceder a sus datos. Si aún no ha configurado Alfresco para Amazon Kendra, consulte [Requisitos previos \(p. 289\)](#).

Console

Para conectarse Amazon Kendra a Alfresco

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, elija Índices y, a continuación, el índice que deseé utilizar de la lista de índices.

Note

Puede configurar o editar los ajustes del control de acceso de usuario en la configuración del índice.

3. En la página de introducción, selecciona Agregar fuente de datos.

4. En la página Agregar fuente de datos, seleccione el conector de Alfresco y, a continuación, elija Agregar fuente de datos.
5. En la página Especificar detalles de la fuente de datos, introduzca la siguiente información:
 - a. En Nombre y descripción, en Nombre de fuente de datos: introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. Descripción (opcional): introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, en Idioma predeterminado: un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, seleccione Agregar nueva etiqueta: etiquetas para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costos.
 - e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
 - a. Alfrescotipo: elija si utiliza Alfresco las instalaciones locales o Alfresco la nube (plataforma como servicio).
 - b. URL del repositorio de Alfresco: introduzca la URL del repositorio de Alfresco. Por ejemplo, si usas Alfresco Cloud (PaaS), la URL del repositorio podría ser `https://company.alfrescocloud.com`. O bien, si usa Alfresco On-Premises, la URL del repositorio podría ser `https://company-alfresco-instance.company-domain.suffix:port`
 - c. Aplicación de usuario de Alfresco. URL: introduzca la URL de la interfaz de Alfresco usuario. Puede obtener la URL del repositorio de su Alfresco administrador. Por ejemplo, la URL de la interfaz de usuario podría ser `https://example.com`.
 - d. Ubicación del certificado SSL: introduzca la ruta al certificado SSL almacenado en un Amazon S3 bucket. Se usa para conectarse a Alfresco On-Premises con una conexión SSL segura.
 - e. AWS Secrets Managersecreto: elija Autenticación básica o autenticación OAuth 2.0. A continuación, elija un Secrets Manager secreto existente o cree uno nuevo para almacenar sus Alfresco credenciales. Si decide crear un secreto nuevo, se abre una ventana AWS Secrets Manager secreta.

Si elige Autenticación básica, introduzca un nombre para el secreto, el nombre Alfresco de usuario y la contraseña.

Si eliges la autenticación OAuth 2.0, introduce un nombre para el secreto, el ID de cliente, el secreto del cliente y la URL del token.
 - f. Nube privada virtual (VPC): puedes elegir usar una VPC. Si es así, debe añadir subredes y grupos de seguridad de VPC.
 - g. Rastreador de identidades: elija rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos y guárdela en el Amazon Kendra almacén principal o de identidades. El almacén de identidades es útil para filtrar el contexto del usuario, donde los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos.
 - h. IAMrol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un rol nuevo para evitar errores.

- i. Elija Siguiente.
7. En la página Configurar los ajustes de sincronización, introduzca la siguiente información:

- a. Alcance de sincronización: establece límites para rastrear cierto contenido y filtra el contenido mediante patrones de expresiones regulares.
 - b.
 - i. Contenido: elige si deseas rastrear el contenido marcado con «Aspectos» en Alfresco, el contenido de un Alfresco sitio específico o el contenido de todos tus Alfresco sitios.
 - ii. Configuración adicional (opcional): defina los siguientes parámetros:
 - Incluir comentarios: elija incluir comentarios en la biblioteca de Alfresco documentos y el blog.
 - Patrones de expresiones regulares: patrones de expresiones regulares para incluir o excluir ciertos archivos.
 - c. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos con Amazon Kendra por primera vez, todo el contenido se sincroniza de forma predeterminada.
 - Sincronización completa: sincroniza todo el contenido independientemente del estado de sincronización anterior.
 - Sincronización de documentos nuevos, modificados o eliminados: sincronice solo los documentos nuevos, modificados y eliminados.
 - d. En Sincronizar programación de ejecución, en Frecuencia: elija la frecuencia con la Amazon Kendra que se sincronizará con la fuente de datos.
 - e. Elija Siguiente.
8. En la página Establecer mapeos de campos, introduzca la siguiente información:
 - a. Seleccione uno de los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.
 - b. Para agregar campos de fuente de datos personalizados, cree un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
 9. En la página Revisar y crear, compruebe que la información que ha introducido sea correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. La fuente de datos aparecerá en la página Fuentes de datos cuando se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Alfresco

Debe especificar un JSON del [esquema de la fuente de datos](#) mediante la [TemplateConfigurationAPI](#). Debe proporcionar la siguiente información:

- Fuente de datos: especifique la fuente de datos como ALFRESCO.
- AlfrescoID del sitio: especifique el ID del sitio de Alfresco.
- AlfrescoURL del repositorio: especifique la URL del Alfresco repositorio. Puede obtener la URL del repositorio de su Alfresco administrador. Por ejemplo, si usas Alfresco Cloud (PaaS), la URL del repositorio podría ser <https://company.alfrescocloud.com>. O bien, si usa Alfresco On-Premises, la URL del repositorio podría ser <https://company-alfresco-instance.company-domain.suffix:port>
- AlfrescoURL de la aplicación web: especifique la URL de la interfaz de Alfresco usuario. Puede obtener la URL del repositorio de su Alfresco administrador. Por ejemplo, la URL de la interfaz de usuario podría ser <https://example.com>.
- Tipo de autenticación: especifique qué tipo de autenticación desea utilizar, si OAuth2 o Basic.
- Alfrescotipo: especifique el tipo Alfresco que utiliza, ya sea PAAS (en la nube/plataforma como servicio) o ON_PREM (local).

- Tipo: especifique TEMPLATE como el tipo cuando llame `CreateDataSource`.
- Nombre secreto de recurso de Amazon (ARN): si desea utilizar la autenticación básica, proporcione un secreto que almacene las credenciales de autenticación de su nombre de usuario y contraseña. Usted proporciona el nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "user name",  
    "password": "password"  
}
```

Si quieras usar la autenticación OAuth 2.0, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "clientId": "client ID",  
    "clientSecret": "client secret",  
    "tokenUrl": "token URL"  
}
```

- IAMrol: especifique RoleArn cuándo llama `CreateDataSource` para proporcionar a un IAM rol los permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Alfresco y Amazon Kendra. Para obtener más información, consulte las [IAMfunciones de las fuentes de datos de Alfresco](#).

También puede añadir las siguientes funciones opcionales:

- Virtual Private Cloud (VPC): especifique VpcConfiguration cuándo llama `CreateDataSource`. Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).
- Modo de sincronización: especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:
 - FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice.
 - FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice.
- Tipo de contenido: el tipo de contenido que deseas rastrear, ya sea el contenido marcado con «Aspectos» en Alfresco, el contenido de un Alfresco sitio específico o el contenido de todos tus Alfresco sitios. También puedes enumerar contenido específico de «Aspectos».
- Filtros de inclusión y exclusión: especifique si se van a incluir o excluir determinados archivos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Rastreador de identidades: especifique si desea activar el rastreador de identidades. Si el rastreador de identidades está desactivado, debes cargar la información de identidad o principal mediante la API. [PutPrincipalMapping](#) Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la

búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado de contextos de usuario](#).

- Asignaciones de campos: elija asignar los campos de la fuente de datos de Alfresco a los campos de índice. Amazon Kendra Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).

Para obtener una lista de otras claves JSON importantes para configurar, consulte el [esquema Alfresco de la plantilla](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Alfresco, consulte:

- [Busque Alfresco contenido de forma inteligente utilizando Amazon Kendra](#)

Amazon RDS/Aurora

Puede indexar los documentos almacenados en una base de datos mediante una fuente de datos de base de datos. Después de proporcionar la información de conexión para la base de datos, Amazon Kendra conecta e indexa los documentos.

Amazon Kendraadmite las siguientes bases de datos:

- Amazon Aurora MySQL
- Amazon Aurora PostgreSQL
- Amazon RDS para MySQL
- Amazon RDS para PostgreSQL

Note

No se admiten las bases de datos Aurora sin servidor.

Puede conectarse Amazon Kendra a la fuente de datos de la base de datos mediante la [Amazon Kendraconsola](#) y la [DatabaseConfigurationAPI](#).

Para solucionar problemas con el conector de fuentes Amazon Kendra de datos de la base de datos, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 294\)](#)
- [Requisitos previos \(p. 295\)](#)
- [Instrucciones de conexión \(p. 296\)](#)

Características admitidas

Amazon Kendrael conector de fuente de datos de base de datos admite las siguientes funciones:

- Mapeos de campo
- Filtrado de contexto de usuario
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de utilizarla Amazon Kendra para indexar la fuente de datos de la base de datos, realice estos cambios en la base de datos y en AWS las cuentas.

En tu base de datos, asegúrate de tener:

- Apuntó su par de usuario y contraseña como credenciales de autenticación básicas para su fuente de datos.
- Copió el nombre del host, el número de puerto, la dirección del host y el nombre de la tabla de datos que contiene los datos de la base de datos. Para PostgreSQL, la tabla de datos debe ser pública.

Note

El host y el puerto indican Amazon Kendra dónde encontrar el servidor de base de datos en Internet. El nombre de la base de datos y el nombre de la tabla indican Amazon Kendra dónde se encuentran los datos del documento en el servidor de la base de datos.

- Copió los nombres de las columnas de la tabla de datos que contienen los datos del documento, el identificador del documento, una a cinco columnas para detectar si un documento ha cambiado y las columnas opcionales de la tabla de datos que se asignan a campos de índice personalizados. Puede asignar cualquiera de los nombres de campo Amazon Kendra reservados a una columna de la tabla.
- Copió la información del tipo de motor de base de datos, por ejemplo, si lo usa Amazon RDS para MySQL u otro tipo.
- Se ha comprobado que cada documento es único en la base de datos y en las demás fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En suCuenta de AWS, asegúrese de tener:

- [Creó un Amazon Kendra índice](#) y, si utilizaba la API, anotó el ID del índice.
- [Creó una IAM función](#) para la fuente de datos y, si utilizaba la API, anotó el ARN de la IAM función.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al identificador AWS Secrets Manager secreto correcto.

- Almacenó las credenciales de autenticación de la base de datos en AWS Secrets Manager secreto y, si utilizaba la API, anotó el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. No se recomienda reutilizar las credenciales y los secretos en las fuentes de datos ni en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o un secreto existentes, puede usar la consola para crear un IAM rol y un Secrets Manager secreto nuevos al conectar la fuente de datos de la base de datosAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes, así como un identificador de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a la fuente de datos de la base de datos, debe proporcionar los detalles necesarios de la fuente de datos de la base de datos para Amazon Kendra poder acceder a los datos. Si aún no ha configurado la base de datos para Amazon Kendra, consulte [Requisitos previos \(p. 295\)](#).

Console

Para conectar Amazon Kendra a una base de datos

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, elija Índices y, a continuación, el índice que deseé utilizar de la lista de índices.

Note

Puede configurar o editar los ajustes del control de acceso de usuario en la configuración del índice.

3. En la página de introducción, selecciona Agregar fuente de datos.
4. En la página Agregar fuente de datos, elija el conector de base de datos y, a continuación, elija Agregar fuente de datos.
5. En la página Especificar detalles de la fuente de datos, introduzca la siguiente información:
 - a. En Nombre y descripción, en Nombre de fuente de datos: introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. Descripción (opcional): introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, en Idioma predeterminado: un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, seleccione Agregar nueva etiqueta: etiquetas para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costos.
 - e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
 - a. Punto final: un nombre de host DNS, una dirección IPv4 o una dirección IPv6.
 - b. Puerto: número de puerto.
 - c. Base de datos: nombre de la base de datos.
 - d. Nombre de tabla: nombre de tabla.
 - e. En Tipo de autenticación, elija entre Existente y Nueva para almacenar las credenciales de autenticación de la base de datos. Si decide crear un nuevo secreto, se abre una ventana AWS Secrets Manager secreta.
 - Introduzca la siguiente información en la ventana Crear un AWS Secrets Manager secreto:
 - A. Nombre secreto: un nombre para tu secreto. El prefijo 'AmazonKendra-database-' se añade automáticamente a tu nombre secreto.
 - B. Para nombre de usuario y contraseña: introduzca los valores de las credenciales de autenticación de su cuenta de base de datos.
 - C. Selecciona Guardar autenticación.
 - f. Nube privada virtual (VPC): puedes elegir usar una VPC. Si es así, debe añadir subredes y grupos de seguridad de VPC.

Note

Debe utilizar una subred privada. Si la instancia de RDS se encuentra en una subred pública de la VPC, puede crear una subred privada que tenga acceso saliente a una puerta de enlace NAT de la subred pública. Las subredes proporcionadas en la configuración de VPC deben estar en Oeste de EE. UU. (Oregón), Este de EE. UU. (Norte de Virginia), Europa (Irlanda).

- g. IAMrol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales de su repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un rol nuevo para evitar errores.

- h. Elija Siguiente.

7. En la página Configurar los ajustes de sincronización, introduzca la siguiente información:

- a. Seleccione entre Aurora MySQL, MySQL, Aurora PostgreSQL y PostgreSQL según su caso de uso.
- b. Encerrar los identificadores SQL entre comillas dobles: seleccione esta opción para incluir los identificadores SQL entre comillas dobles. Por ejemplo, «ColumnName».
- c. Columna ACL y columnas de detección de cambios: se utilizan para configurar las columnas que se Amazon Kendra utilizan para las listas de detección de cambios y control de acceso.
- d. En Sincronizar programación de ejecución, en Frecuencia: elija la frecuencia con la Amazon Kendra que se sincronizará con la fuente de datos.
- e. Elija Siguiente.

8. En la página Establecer mapeos de campos, introduzca la siguiente información:

- a. Amazon Kendraasignaciones de campos predeterminadas: seleccione uno de los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice. Debe añadir los valores de la columna Base de datos para document_id y document_body
- b. Asignaciones de campos personalizadas: para agregar campos de fuentes de datos personalizados para crear un nombre de campo de índice al que asignar y el tipo de datos del campo.
- c. Elija Siguiente.

9. En la página Revisar y crear, compruebe que la información que ha introducido sea correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. La fuente de datos aparecerá en la página Fuentes de datos una vez que se haya agregado correctamente.

API

Para conectarse Amazon Kendra a una base de datos

Debe especificar lo siguiente en la [DatabaseConfiguration](#) API:

- ColumnConfiguration—Información sobre dónde debe obtener el índice la información del documento de la base de datos. Para obtener más información, consulte [ColumnConfiguration](#). Debe especificar los DocumentIdColumnName campos DocumentDataColumnName y. La columna asignada al DocumentIdColumnName campo debe ser una columna de números enteros. El siguiente ejemplo muestra una configuración de columna sencilla para una fuente de datos de base de datos:

```
"ColumnConfiguration": {  
    "ChangeDetectingColumns": [  
        "LastUpdateDate",  
        "LastUpdateTime"  
    ],  
    "DocumentDataColumnName": "TextColumn",  
    "DocumentIdColumnName": "IdentifierColumn",  
    "DocoumentTitleColumnName": "TitleColumn",  
    "FieldMappings": [  
        {  
            "DataSourceFieldName": "AbstractColumn",  
            "IndexFieldName": "Abstract"  
        }  
    ]  
}
```

- ConnectionConfiguration: el tipo de motor de base de datos que ejecuta la base de datos. El DatabaseHost campo debe ser el punto final de la instancia Amazon Relational Database Service (Amazon RDS) de la base de datos. No utilice el punto final del clúster.
- DatabaseEngineType—Información de configuración necesaria para conectarse a una base de datos. Para obtener más información, consulte [ConnectionConfiguration](#).
- Nombre de recurso secreto de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta de base de datos. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "user name",  
    "password": "password"  
}
```

El secreto puede contener más información. El siguiente ejemplo muestra una configuración de base de datos.

```
"DatabaseConfiguration": {  
    "ConnectionConfiguration": {  
        "DatabaseHost": "host.subdomain.domain.tld",  
        "DatabaseName": "DocumentDatabase",  
        "DatabasePort": 3306,  
        "SecretArn": "arn:aws:secretmanager:region:account ID:secret/secret name",  
        "TableName": "DocumentTable"  
    }  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. No se recomienda reutilizar las credenciales y los secretos en las fuentes de datos ni en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMrol: especifique RoleArn cuándo llama CreateDataSource para proporcionar a un IAM rol los permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de la base de datos y. Amazon Kendra Para obtener más información, consulte [IAMFunciones para fuentes de datos de bases de datos](#).

También puede añadir las siguientes funciones opcionales:

- Virtual Private Cloud (VPC): especifique `VpcConfiguration` como parte de la configuración de la fuente de datos. Consulte [Configurar Amazon Kendra para usar una VPC](#).

Note

Solo debe usar una subred privada. Si la instancia de RDS se encuentra en una subred pública de la VPC, puede crear una subred privada que tenga acceso saliente a una puerta de enlace NAT de la subred pública. Las subredes proporcionadas en la configuración de VPC deben estar en Oeste de EE. UU. (Oregón), Este de EE. UU. (Norte de Virginia), Europa (Irlanda).

- Asignaciones de campos: elija asignar los campos de la fuente de datos de la base de datos Amazon Kendra a los campos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario: Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de bases de datos](#).

Amazon FSx

Amazon FSx es un sistema de servidor de archivos basado en la nube totalmente gestionado que ofrece capacidades de almacenamiento compartido. Si es usuario de Amazon FSx, puede utilizar Amazon Kendra para indexar su fuente de datos de Amazon FSx.

Amazon Kendra actualmente solo admite Amazon FSx para Windows File Server.

Puedes conectarte a Amazon Kendra a su fuente de datos de Amazon FSx mediante el [Amazon Kendra console](#) y el [FsxConfiguration API](#).

Para solucionar problemas de su Amazon Kendra Conector de fuente de datos Amazon FSx, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 299\)](#)
- [Requisitos previos \(p. 299\)](#)
- [Instrucciones de conexión \(p. 300\)](#)
- [Más información \(p. 303\)](#)

Características admitidas

Amazon Kendra El conector de fuente de datos Amazon FSx admite las siguientes funciones:

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de poder usar Amazon Kendra para indexar su fuente de datos de Amazon FSx, realizar estos cambios en su Amazon FSx y AWS cuenta.

En Amazon FSx, asegúrese de tener:

- Una cuenta de Amazon FSx con permisos de lectura y montaje.
- Tomó nota de suAmazon FSx credenciales de autenticación para una cuenta de usuario de Active Directory. Esto incluye su nombre de usuario de Active Directory y su nombre de dominio del Sistema de nombres de dominio (DNS). Por ejemplo,*user@corp.example.com*.
- Copió tuAmazon FSx ID del sistema de archivos.
- Usó unAmazon VPC(AWSVPC) donde tuAmazon FSx reside.
- Se comprobó que cada documento es único en Amazon FSx y en otras fuentes de datos que tiene previsto utilizar para el mismo índice. Cada fuente de datos que desee utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tuIAMrol para acceder al correctoAWS Secrets Manager identificación secreta.

- Ha almacenado sus credenciales de autenticación de Amazon FSx en unAWS Secrets Manager secreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto al conectar su fuente de datos de Amazon FSx aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa su fuente de datos de Amazon FSx, debe proporcionar los detalles necesarios de su fuente de datos de Amazon FSx para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configurado Amazon FSx paraAmazon Kendra, consulte[Requisitos previos \(p. 299\)](#).

Console

Para conectarAmazon Kendraa Amazon FSx

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enConfiguración del índice.

3. En elCómo empezar página, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConector Amazon FSx, a continuación, elijaAregar fuente de datos.

5. En el Especificar los detalles de la fuente de datos página, introduzca la siguiente información:
 - a. En Nombre y descripción, para Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. (Opcional) Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, para idioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, para Añadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tus AWS costos.
 - e. Elija Siguiente.
6. En el Definir el acceso y la seguridad página, introduzca la siguiente información:
 - a. Fuente—Seleccione Amazon FSx para servidor de archivos de Windows.
 - b. Amazon FSx ID del sistema de archivos—Seleccione el ID del sistema de archivos o cree un directorio nuevo.
 - c. AWS Secrets Manager secreto—Elige un secreto existente o crea uno nuevo. Secrets Manager secreto para almacenar sus credenciales de autenticación de Amazon FSx. Si eliges crear un nuevo secreto, un AWS Secrets Manager abre una ventana secreta.
 - i. Introduzca la siguiente información en el Crea un AWS Secrets Manager ventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Amazon FSX' se añade automáticamente a su nombre secreto.
 - B. Para Nombre de usuario—Introduzca el nombre de usuario de Amazon FSx Cuenta de Active Directory.
 - C. Para Contraseña—Introduzca la contraseña de Amazon FSx Cuenta de Active Directory.
 - ii. Seleccione Guardar.
 - d. Nube privada virtual (VPC)—Debes usar una VPC con tu Amazon FSx fuente de datos. Debes añadir Subredes y Grupos de seguridad de VPC.
 - e. IAM papel—Elige uno existente IAM rol o crea uno nuevo IAM rol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAM roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

- f. Elija Siguiente.
7. En el Configurar los ajustes de sincronización página, introduzca la siguiente información:
 - a. Patrones Regex—Añade patrones de expresiones regulares para incluir o excluir cierto contenido. Puede añadir hasta 100 patrones.
 - b. En Cronograma de ejecución de sincronización, para Frecuencia—Elige con qué frecuencia Amazon Kendra debe sincronizarse con la fuente de datos.
 - c. Elija Siguiente.
8. En el Establecer mapeos de campos página, introduzca la siguiente información:
 - a. Mapeos de campos de Amazon FSx Windows—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.

- c. Elija Siguiente.
9. En el Revisar y crear página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en las Fuentes de datos página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Amazon FSx

Debe especificar lo siguiente mediante el [FsxConfiguration API](#):

- ID del sistema de archivos—El identificador del sistema de archivos Amazon FSx. Puede encontrar el identificador del sistema de archivos en el panel Sistemas de archivos de la consola de Amazon FSx.
- Tipo de sistema de archivos—Especifique el tipo de sistema de archivos como WINDOWS.
- Nube privada virtual (VPC)—Debe seleccionar un Amazon VPC donde reside su Amazon FSx. Esto incluye la subred de VPC y los grupos de seguridad. Consulte [Configuración de un Amazon VPC](#).
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contiene las credenciales de autenticación de su cuenta de Amazon FSx. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "user@corp.example.com",  
    "password": "password"  
}
```

- IAM papel—Especificando el rol cuando llamas a `CreateDataSource` para proporcionar un IAM role con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas necesarias para el conector Amazon FSx y Amazon Kendra. Para obtener más información, consulte [IAM funciones para las fuentes de datos de Amazon FSx](#).

También puede añadir las siguientes funciones opcionales:

- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir ciertos contenidos y tipos de contenido.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coinciden con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coinciden con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Amazon FSx](#).

Note

Para probar el filtrado del contexto de usuario en un usuario, debe incluir el nombre de dominio DNS como parte del nombre de usuario cuando emita la consulta. Debe tener permisos administrativos del dominio de Active Directory. También puede probar el filtrado de contexto de usuario en el nombre de un grupo.

- Mapeos de campo—Elija asignar los campos de su fuente de datos de Amazon FSx a su Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Amazon FSx, consulte:

- [Busque de forma segura datos no estructurados en sistemas de archivos de Windows con la Amazon Kendra conector para Amazon FSx para Windows File Server.](#)

Amazon S3

Amazon S3 es un servicio de almacenamiento de objetos que almacena datos como objetos dentro de cubos. Puede usarlo Amazon Kendra para indexar su repositorio de documentos de Amazon S3 bucket.

Warning

Amazon Kendra utiliza una política de bucket que otorgue permisos a un Amazon Kendra director para interactuar con un bucket de S3. En su lugar, usa IAM roles. Asegúrese de que Amazon Kendra no esté incluido como miembro de confianza en su política de bucket para evitar cualquier problema de seguridad de los datos al conceder permisos accidentalmente a directores arbitrarios. Sin embargo, puedes añadir una política de bucket para usar un Amazon S3 bucket en diferentes cuentas. Para obtener más información, consulte [Políticas para usar Amazon S3 en todas las cuentas](#) (en la pestaña IAM Funciones de S3, en IAM Funciones para fuentes de datos). Para obtener información sobre las IAM funciones de las fuentes de datos de S3, consulte las [IAMfunciones](#).

Para conectarse a Amazon S3, especifique la conexión y otra información en la consola, mediante [S3 DataSourceConfiguration](#) o mediante el [TemplateConfiguration](#). Si usa el `TemplateConfiguration` objeto, puede usar una VPC para conectarse a la fuente de datos. Tú `VpcConfiguration` específicas cuando llamas `CreateDataSource`. Para obtener más información, consulte [Configuring a VPC](#) (Configuración de una VPC).

Note

`S3DataSourceConfiguration` La API no admite la configuración de una VPC. Solo la `TemplateConfiguration` API y la consola admiten la configuración de VPC.

Para solucionar problemas con el conector de fuente de datos de Amazon Kendra S3, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 303\)](#)
- [Requisitos previos \(p. 304\)](#)
- [Instrucciones de conexión \(p. 304\)](#)
- [Creación de una fuente Amazon S3 de datos \(p. 307\)](#)
- [Amazon S3 metadatos del documento \(p. 309\)](#)
- [Control de acceso a fuentes Amazon S3 de datos \(p. 311\)](#)

Características admitidas

API de configuración S3

- Mapeos de campo
- Filtrado de contexto de usuario
- Patrones de inclusión/exclusión

Consola/API TemplateConfiguration

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados

Requisitos previos

Antes de utilizarla Amazon Kendra para indexar tu fuente de datos de S3, realiza estos cambios en tu S3 y en tus AWS cuentas.

En S3, asegúrate de tener:

- Copió el nombre de tu Amazon S3 bucket.

Note

El depósito debe estar en la misma región que el Amazon Kendra índice y el índice debe tener permiso para acceder al depósito que contiene los documentos.

- Se comprobó que cada documento es único en S3 y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu AWS cuenta, asegúrate de tener:

- [Creó un Amazon Kendra índice](#) y, si utilizaba la API, anotó el ID del índice.
- [Creó una IAM función](#) para la fuente de datos y, si utilizaba la API, anotó el ARN de la IAM función.

Si no tiene un IAM rol existente, puede usar la consola para crear un IAM rol nuevo al conectar su fuente de datos de S3 aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol existente y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de S3, debe proporcionar los detalles necesarios de su fuente de datos de S3 para Amazon Kendra poder acceder a sus datos. Si aún no ha configurado S3 paraAmazon Kendra, consulte[Requisitos previos \(p. 304\)](#).

Console

Para conectarse Amazon Kendra a Amazon S3

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, elija Índices y, a continuación, el índice que deseé utilizar de la lista de índices.

Note

Puede configurar o editar los ajustes del control de acceso de usuario en la configuración del índice.

3. En la página de introducción, selecciona Agregar fuente de datos.
4. En la página Agregar fuente de datos, elija el conector S3 y, a continuación, elija Agregar fuente de datos.
5. En la página Especificar detalles de la fuente de datos, introduzca la siguiente información:
 - a. En Nombre y descripción, en Nombre de fuente de datos: introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. Descripción (opcional): introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, en Idioma predeterminado: un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, seleccione Agregar nueva etiqueta: etiquetas para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costos.
 - e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información opcional:
 - a. IAMrol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un rol nuevo para evitar errores.

- b. Nube privada virtual (VPC): puedes elegir usar una VPC. Si es así, debe añadir subredes y grupos de seguridad de VPC.

Important

Asegúrese de tener:

- Configuró su VPC de acuerdo con los pasos de [Gateway endpoints](#) para Amazon S3
 - Escogió una subred privada en una zona de disponibilidad Amazon Kendra compatible. Consulte [Configurar Amazon Kendra para usar un Amazon VPC](#) para obtener más información.
 - Configuró su grupo de seguridad para Amazon Kendra permitir el acceso al Amazon S3 punto final.
- c. Elija Siguiente.
 7. En la página Configurar los ajustes de sincronización, introduzca la siguiente información:
 - a. En el ámbito de sincronización, para la ubicación de la fuente de datos: la ruta al Amazon S3 depósito donde se almacenan los datos. Selecciona Browse S3 para elegir tu bucket.
 - b. (Opcional) Ubicación de carpeta con prefijo de archivos de metadatos: la ruta a la carpeta en la que se almacenan los metadatos. Selecciona Browse S3 para localizar su carpeta de metadatos.
 - c. (Opcional) Ubicación del archivo de configuración de la lista de control de acceso: la ruta a la ubicación de un archivo que contiene una estructura JSON que especifica la configuración de acceso de los archivos almacenados en la fuente de datos de S3. Selecciona Browse S3 para localizar el archivo ACL.

- d. (Opcional) Seleccione la clave de descifrado: seleccione esta opción para utilizar una clave de descifrado. Puede optar por utilizar una AWS KMS clave existente o crear una nueva.
 - e. (Opcional) En Configuración adicional, para Patrones: añada patrones para incluir o excluir documentos del índice. Todas las rutas son relativas al bucket S3 de ubicación de la fuente de datos. Puede añadir hasta 100 patrones.
 - f. En el modo Sincronización, elija entre el modo de sincronización completa y la sincronización de contenido nuevo, modificado o eliminado para determinar cómo se actualiza el índice cuando cambia el contenido de la fuente de datos. Cuando sincroniza su fuente de datos con Amazon Kendra por primera vez, todo el contenido se sincroniza de forma predeterminada.
 - g. En Sincronizar programación de ejecución, en Frecuencia: elija la frecuencia con la Amazon Kendra que se sincronizará con la fuente de datos.
 - h. Elija Siguiente.
8. En la página Establecer mapeos de campos, introduzca la siguiente información opcional:
 - a. Mapeo de campos S3: seleccione uno de los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.
 - b. Agregar campo: elija agregar campos de fuente de datos personalizados para crear un nombre de campo de índice al que asignar y el tipo de datos del campo.
 - c. Elija Siguiente.
 9. En la página Revisar y crear, compruebe que la información que ha introducido sea correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. La fuente de datos aparecerá en la página Fuentes de datos cuando se haya agregado correctamente.

TemplateConfiguration API

Para conectarse Amazon Kendra a Amazon S3

Debe especificar un JSON del [esquema de la fuente de datos](#) mediante la [TemplateConfigurationAPI](#). Debe proporcionar la siguiente información:

- BucketName: el nombre del depósito que contiene los documentos.
- IAMrol: especifique RoleArn cuándo llama `CreateDataSource` para proporcionar a un IAM rol los permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector S3 y Amazon Kendra. Para obtener más información, consulte las [IAMfunciones de las fuentes de datos de S3](#).

También puede añadir las siguientes funciones opcionales:

- Filtros de inclusión y exclusión: especifican si se van a incluir o excluir ciertos nombres de archivos, tipos de archivos y rutas de archivos y utilizar patrones globales (patrones que pueden expandir un patrón comodín a una lista de nombres de rutas que coincidan con el patrón dado). Algunos ejemplos de patrones de globo son:
 - `/myapp/config/*`—Todos los archivos dentro del directorio de configuración
 - `/**/*.png`—Todos los archivos.png de todos los directorios
 - `/**/*.{png,ico,md}`—Todos los archivos.png, .ico o .md de todos los directorios
 - `/myapp/src/**/*.ts`—Todos los archivos.ts del directorio src (y todos sus subdirectorios)
 - `**!/(*.module).ts`—Todos los archivos.ts, pero no .module.ts
- Asignaciones de campos: elija asignar los campos de la fuente de datos de S3 a los campos de índice. Amazon Kendra Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).

- Virtual Private Cloud (VPC): especifique VpcConfiguration cuándo llama CreateDataSource. Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).
- Configuración de metadatos del documento: agregue archivos de metadatos del documento que contengan información como la información de control de acceso al documento, el URI de origen, el autor del documento y los atributos personalizados. Cada archivo de metadatos contiene metadatos sobre un solo documento.
- Filtrado de contexto de usuario: Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de S3](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [Amazon S3 esquema de plantilla \(p. 153\)](#).

S3DataSourceConfiguration API

Para conectarse Amazon Kendra a Amazon S3

Debe especificar lo siguiente mediante la DataSourceConfiguration API de [S3](#):

- BucketName: el nombre del depósito que contiene los documentos.
- IAMRole: especifique RoleArn cuándo llama CreateDataSource para proporcionar a un IAM rol los permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector S3 y Amazon Kendra. Para obtener más información, consulte las [IAM funciones de las fuentes de datos de S3](#).

También puede añadir las siguientes funciones opcionales:

- Filtros de inclusión y exclusión: especifican si se va a incluir o excluir cierto contenido mediante prefijos, nombre de archivo, tipo de archivo, ruta de archivo y patrones globales (patrones que pueden expandir un patrón comodín en una lista de nombres de rutas que coincidan con el patrón dado). Algunos ejemplos de patrones de globo son:
 - /myapp/config/*—Todos los archivos dentro del directorio de configuración
 - **/*.png—Todos los archivos.png de todos los directorios
 - **/*.{png,ico,md}—Todos los archivos.png, .ico o .md de todos los directorios
 - /myapp/src/**/*.ts—Todos los archivos.ts del directorio src (y todos sus subdirectorios)
 - **!/(*.module).ts—Todos los archivos.ts, pero no .module.ts
- Filtrado de contexto de usuario: Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de S3](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de S3, consulte:

- [Busque respuestas con precisión mediante el conector Amazon Kendra S3 compatible con VPC](#)

Creación de una fuente Amazon S3 de datos

Los siguientes ejemplos muestran la creación de una fuente de Amazon S3 datos. En los ejemplos se supone que ya ha creado un índice y un IAM rol con permiso para leer los datos del índice. Para obtener

más información sobre el IAM rol, consulte [roles de IAM acceso](#). Para obtener más información sobre la creación de un índice, consulte [Creación de un índice](#).

CLI

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

Python

El siguiente fragmento de código Python crea una fuente de Amazon S3 datos. Para ver el ejemplo completo, consulte [Introducción \(AWS SDK for Python \(Boto3\)\) \(p. 93\)](#).

```
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 source 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
)
```

Crear la fuente de datos puede llevar algún tiempo. Puede supervisar el progreso mediante la [DescribeDataSource API](#). Cuando el estado de la fuente de datos es, ACTIVE la fuente de datos está lista para usarse.

Los siguientes ejemplos muestran cómo obtener el estado de una fuente de datos.

CLI

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

Python

El siguiente fragmento de código Python obtiene información sobre una fuente de datos de S3. Para ver el ejemplo completo, consulte [Introducción \(AWS SDK for Python \(Boto3\)\) \(p. 93\)](#).

```
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
```

Esta fuente de datos no tiene una programación, por lo que no se ejecuta automáticamente. Para indexar la fuente de datos, llame [StartDataSourceSyncJob](#) para sincronizar el índice con la fuente de datos.

Los siguientes ejemplos muestran la sincronización de una fuente de datos.

CLI

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

Python

El siguiente fragmento de código Python sincroniza una Amazon S3 fuente de datos. Para ver el ejemplo completo, consulte[Introducción \(AWS SDK for Python \(Boto3\)\) \(p. 93\)](#).

```
print("Synchronize the data source.")

sync_response = kendra.start_data_source_sync_job(
    Id = "data-source-id",
    IndexId = "index-id"
)
```

Amazon S3 metadatos del documento

Puede añadir metadatos, información adicional sobre un documento, a los documentos de un Amazon S3 bucket mediante un archivo de metadatos. Cada archivo de metadatos está asociado a un documento indexado.

Los archivos de metadatos deben almacenarse en el mismo depósito que los archivos indexados. Puede especificar una ubicación dentro del bucket para los archivos de metadatos mediante la consola o el **S3Prefix** campo del **DocumentsMetadataConfiguration** parámetro al crear una fuente de Amazon S3 datos. Si no especificas ningún Amazon S3 prefijo, los archivos de metadatos deben almacenarse en la misma ubicación que los documentos indexados.

Si especifica un Amazon S3 prefijo para los archivos de metadatos, se encuentran en una estructura de directorios paralela a los documentos indexados. Amazon Kendra busca sus metadatos únicamente en el directorio especificado. Si no se leen los metadatos, compruebe que la ubicación del directorio coincida con la ubicación de los metadatos.

Los siguientes ejemplos muestran cómo la ubicación del documento indexado se corresponde con la ubicación del archivo de metadatos. Tenga en cuenta que la Amazon S3 clave del documento se añade al Amazon S3 prefijo de los metadatos y, a continuación, se le añade el sufijo `.metadata.json` para formar la ruta del archivo de metadatos. Amazon S3 La Amazon S3 clave combinada, con el Amazon S3

prefijo y el `.metadata.json` sufijo de los metadatos, no debe tener más de un total de 1024 caracteres. Se recomienda mantener la Amazon S3 clave por debajo de 1000 caracteres para tener en cuenta los caracteres adicionales al combinar la clave con el prefijo y el sufijo.

```
Bucket name:  
    s3://bucketName  
Document path:  
    documents  
Metadata path:  
    none  
File mapping  
    s3://bucketName/documents/file.txt ->  
        s3://bucketName/documents/file.txt.metadata.json
```

```
Bucket name:  
    s3://bucketName  
Document path:  
    documents/legal  
Metadata path:  
    metadata  
File mapping  
    s3://bucketName/documents/legal/file.txt ->  
        s3://bucketName/metadata/documents/legal/file.txt.metadata.json
```

Los metadatos del documento se definen en un archivo JSON. El archivo debe ser un archivo de texto UTF-8 sin un marcador BOM. El nombre del archivo JSON debe ser `<document>.<extension>.metadata.json`. En este ejemplo, «documento» es el nombre del documento al que se aplican los metadatos y «extensión» es la extensión del archivo del documento. El identificador del documento debe ser único en `<document>.<extension>.metadata.json`.

El contenido del archivo JSON sigue esta plantilla. Todos los atributos son opcionales. Si no especificas `el_source_uri`, los enlaces devueltos por Amazon Kendra en los resultados de búsqueda apuntan al Amazon S3 depósito que contiene el documento.

```
{  
    "DocumentId": "document ID",  
    "Attributes": {  
        "_category": "document category",  
        "_created_at": "ISO 8601 encoded string",  
        "_last_updated_at": "ISO 8601 encoded string",  
        "_source_uri": "document URI",  
        "_version": "file version",  
        "_view_count": "number of times document has been viewed",  
        "custom attribute key": "custom attribute value",  
        additional custom attributes  
    },  
    "AccessControlList": [  
        {  
            "Name": "user name",  
            "Type": "GROUP | USER",  
            "Access": "ALLOW | DENY"  
        }  
    ],  
    "Title": "document title",  
    "ContentType": "For example HTML | PDF. For supported content types, see Types of documents."  
}
```

Los campos `_created_at` y `_last_updated_at` metadatos son fechas codificadas en ISO 8601. Por ejemplo, 2012-03-25T12:30:10+01:00 es el formato de fecha y hora ISO 8601 para el 25 de marzo de 2012, a las 12.30 h (más 10 segundos) en la zona horaria de Europa Central.

Puede añadir información adicional al `Attributes` campo sobre un documento que utilice para filtrar consultas o agrupar las respuestas a las consultas. Para obtener más información, consulte [Creación de campos de documentos personalizados \(p. 116\)](#).

Puede utilizar el `AccessControlList` campo para filtrar la respuesta de una consulta. De esta forma, solo ciertos usuarios y grupos tienen acceso a los documentos. Para obtener más información, consulte [Filtrar según el contexto del usuario \(p. 525\)](#).

Control de acceso a fuentes Amazon S3 de datos

Puede controlar el acceso a los documentos de una fuente Amazon S3 de datos mediante un archivo de configuración. El archivo se especifica en la consola o como `AccessControlListConfiguration` parámetro cuando se llama a la [UpdateDataSourceAPI](#) [CreateDataSource](#).

El archivo de configuración contiene una estructura JSON que identifica un prefijo S3 y enumera la configuración de acceso del prefijo. El prefijo puede ser una ruta o un archivo individual. Si el prefijo es una ruta, la configuración de acceso se aplica a todos los archivos de esa ruta.

Puede especificar usuarios y grupos en la configuración de acceso. Cuando consulta el índice, especifica la información del usuario y del grupo. Para obtener más información, consulte [Filtrar por atributo de usuario \(p. 527\)](#).

La estructura JSON del archivo de configuración debe tener el siguiente formato:

```
[  
  {  
    "keyPrefix": "s3://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"  
          }  
        ]  
    ]  
  }]
```

Amazon KendraRastreador web

Puede utilizar Amazon Kendra Web Crawler para rastrear e indexar páginas web.

Solo puede rastrear sitios web públicos o sitios web internos de la empresa que utilicen el protocolo de comunicación seguro Hypertext Transfer Protocol Secure (HTTPS). Si recibe un error al rastrear un sitio web, es posible que el sitio web esté bloqueado para que no pueda rastrearse. Para rastrear sitios web internos, puedes configurar un proxy web. El proxy web debe estar orientado al público. También puede utilizar la autenticación para acceder a sitios web y rastrearlos.

Al seleccionar los sitios web que se van a indexar, se debe respetar la [Política de uso aceptable de Amazon](#) y todas las demás condiciones de Amazon. Recuerde que solo debe utilizar Amazon Kendra Web Crawler para indexar sus propias páginas web o las páginas web que tenga autorización para indexar. Para obtener información sobre cómo evitar que Amazon Kendra Web Crawler indexe sus sitios web, consulte. [Configuración del robots.txt archivo para Amazon Kendra Web Crawler \(p. 324\)](#)

Note

El uso Amazon Kendra indebido de Web Crawler para rastrear de forma agresiva sitios web o páginas web que no son de su propiedad no se considera un uso aceptable.

Amazon Kendra tiene dos versiones del web crawler conector. Las funciones compatibles de cada versión incluyen:

Amazon KendraConektor Web Crawler v1.0/API [WebCrawlerConfiguration](#)

- Proxy web
- Filtros de inclusión/exclusión

Amazon KendraConektor Web Crawler v2.0/API [TemplateConfiguration](#)

- Mapeos de campo
- Proxy web
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/sincronizar solo documentos nuevos, modificados o eliminados

Para solucionar problemas con el conector de fuente de datos del rastreador Amazon Kendra web, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Amazon KendraConektor Web Crawler v1.0 \(p. 312\)](#)
- [Amazon KendraConektor Web Crawler v2.0 \(p. 317\)](#)
- [Configuración del robots.txt archivo para Amazon Kendra Web Crawler \(p. 324\)](#)

Amazon KendraConektor Web Crawler v1.0

Puede utilizar Amazon Kendra Web Crawler para rastrear e indexar páginas web.

Solo puede rastrear sitios web orientados al público y sitios web que utilicen el protocolo de comunicación seguro Hypertext Transfer Protocol Secure (HTTPS). Si recibe un error al rastrear un sitio web, es posible que el sitio web esté bloqueado para que no pueda rastrearse. Para rastrear sitios web internos, puedes configurar un proxy web. El proxy web debe estar orientado al público.

Al seleccionar los sitios web que se van a indexar, se debe respetar la [Política de uso aceptable de Amazon](#) y todas las demás condiciones de Amazon. Recuerde que solo debe utilizar Amazon Kendra Web Crawler para indexar sus propias páginas web o las páginas web que tenga autorización para indexar. Para obtener información sobre cómo evitar que Amazon Kendra Web Crawler indexe sus sitios web, consulte [Configuración del robots.txt archivo para Amazon Kendra Web Crawler \(p. 324\)](#)

Note

El uso Amazon Kendra indebido de Web Crawler para rastrear de forma agresiva sitios web o páginas web que no son de su propiedad no se considera un uso aceptable.

Para solucionar problemas con el conector de fuente de datos del rastreador Amazon Kendra web, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 313\)](#)
- [Requisitos previos \(p. 313\)](#)
- [Instrucciones de conexión \(p. 314\)](#)
- [Más información \(p. 317\)](#)

Características admitidas

- Proxy web
- Filtros de inclusión/exclusión

Requisitos previos

Antes de utilizarlos Amazon Kendra para indexar sus sitios web, compruebe los detalles de sus sitios web y AWS cuentas.

Para tus sitios web, asegúrate de tener:

- Copió las URL iniciales o del mapa del sitio web de los sitios web que desea indexar.
- Para sitios web que requieren autenticación básica: anota el nombre de usuario y la contraseña y copia el nombre de host del sitio web y el número de puerto.
- Opcional: copia el nombre de host del sitio web y el número de puerto si quieras usar un proxy web para conectarte a los sitios web internos que deseas rastrear. El proxy web debe estar orientado al público. Amazon Kendra admite la conexión a servidores proxy web respaldados por una autenticación básica o puede conectarse sin autenticación.
- Marcó que cada documento de página web que desea indexar es único y está en todas las demás fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que desee utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu AWS cuenta, asegúrate de tener:

- [Creó un Amazon Kendra índice](#) y, si utilizaba la API, anotó el ID del índice.
- [Creó una IAM función](#) para la fuente de datos y, si utilizaba la API, anotó el ARN de la IAM función.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al identificador AWS Secrets Manager secreto correcto.

- En el caso de los sitios web que requieren autenticación, o si utilizan un proxy web con autenticación, guarde las credenciales de autenticación en AWS Secrets Manager secreto y, si utiliza la API, anote el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. No se recomienda reutilizar las credenciales y los secretos en las fuentes de datos ni en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o un secreto existentes, puede usar la consola para crear un IAM rol y un Secrets Manager secreto nuevos al conectar su fuente de web crawler datosAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes, así como un identificador de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de web crawler datos, debe proporcionar los detalles necesarios de su fuente de web crawler datos para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configurado web crawler, Amazon Kendra consulte [Requisitos previos \(p. 313\)](#).

Console

Para conectarse Amazon Kendra a web crawler

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, elija Índices y, a continuación, el índice que deseé utilizar de la lista de índices.

Note

Puede configurar o editar los ajustes del control de acceso de usuario en la configuración del índice.

3. En la página de introducción, selecciona Agregar fuente de datos.
4. En la página Agregar fuente de datos, elija el conector para rastreadores web y, a continuación, elija Agregar fuente de datos.
5. En la página Especificar detalles de la fuente de datos, introduzca la siguiente información:
 - a. En Nombre y descripción, en Nombre de fuente de datos: introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. Descripción (opcional): introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, en Idioma predeterminado: un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, seleccione Agregar nueva etiqueta: etiquetas para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costos.
 - e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
 - a. Para Fuente, elige entre las URL de origen y los mapas del sitio de origen según tu caso de uso e introduce los valores de cada una de ellas.

Puedes añadir hasta 10 URL de origen y tres mapas de sitio.

Note

Si quieres rastrear un mapa del sitio, comprueba que la URL base o raíz sea la misma que las URL que aparecen en la página del mapa del sitio. Por ejemplo, si la URL del mapa del sitio es <https://example.com/sitemap-page.html>, las URL que aparecen en esta página del mapa del sitio también deben usar la URL base "https://example.com/">.

- b. (Opcional) Para el proxy web, introduzca la siguiente información:
 - i. Nombre de host: nombre de host donde se requiere un proxy web.
 - ii. Número de puerto: el puerto utilizado por el protocolo de transporte URL del host. El número de puerto debe ser un valor numérico comprendido entre 0 y 65535.
 - iii. Para las credenciales de proxy web: si la conexión de proxy web requiere autenticación, elija un secreto existente o cree uno nuevo para almacenar las credenciales de autenticación. Si decide crear un nuevo secreto, se abre una ventana AWS Secrets Manager secreta.
 - iv. Introduzca la siguiente información en la ventana Crear un AWS Secrets Manager Secrets Manager secreto:
 - A. Nombre secreto: un nombre para tu secreto. El prefijo 'AmazonKendra-WebCrawler-' se añade automáticamente a tu nombre secreto.
 - B. Para nombre de usuario y contraseña: introduzca estas credenciales de autenticación básicas para sus sitios web.
 - C. Seleccione Guardar.
- c. (Opcional) Hosts con autenticación: seleccione esta opción para añadir hosts adicionales con autenticación.
- d. IAMrol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un rol nuevo para evitar errores.

- e. Elija Siguiente.
7. En la página Configurar los ajustes de sincronización, introduzca la siguiente información:
 - a. Intervalo de rastreo: elija el tipo de páginas web que desea rastrear.
 - b. Profundidad de rastreo: seleccione el número de niveles de la URL inicial que se Amazon Kendra deben rastrear.
 - c. En la configuración avanzada de rastreo y en la configuración adicional se introduce la siguiente información:
 - i. Tamaño máximo de archivo: el tamaño máximo de la página web o de los archivos adjuntos que se pueden rastrear. Mínimo 0,000001 MB (1 byte). Máximo 50 MB.
 - ii. Número máximo de enlaces por página: el número máximo de enlaces rastreados por página. Los enlaces se rastrean por orden de aparición. Mínimo 1 enlace/página. Máximo 1000 enlaces/página.
 - iii. Limitación máxima: la cantidad máxima de URL rastreadas por nombre de host por minuto. Mínimo 1 URL/nombre de host/minuto. Máximo de 300 URL/nombre de host/minuto.
 - iv. Patrones de expresiones regulares: agrega patrones de expresiones regulares para incluir o excluir determinadas URL. Puede añadir hasta 100 patrones.

- d. En Sincronizar programación de ejecución, en Frecuencia: elija la frecuencia con la Amazon Kendra que se sincronizará con la fuente de datos.
 - e. Elija Siguiente.
8. En la página Revisar y crear, compruebe que la información que ha introducido sea correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. La fuente de datos aparecerá en la página Fuentes de datos cuando se haya agregado correctamente.

API

Para conectarse Amazon Kendra a web crawler

Debe especificar lo siguiente mediante la [WebCrawlerConfigurationAPI](#):

- URL: especifique las URL iniciales o de punto de inicio de los sitios web o las URL del mapa del sitio web que desea rastrear utilizando y. [SeedUrlConfigurationSiteMapsConfiguration](#)

Note

Si quieres rastrear un mapa del sitio, comprueba que la URL base o raíz sea la misma que las URL que aparecen en la página del mapa del sitio. Por ejemplo, si la URL del mapa del sitio es <https://example.com/sitemap-page.html>, las URL que aparecen en esta página del mapa del sitio también deben usar la URL base "<https://example.com/>".

- Nombre secreto de recurso de Amazon (ARN): si un sitio web requiere una autenticación básica, debe proporcionar el nombre de host, el número de puerto y un secreto que almacena las credenciales de autenticación básicas de su nombre de usuario y contraseña. El ARN secreto se proporciona mediante la [AuthenticationConfigurationAPI](#). El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "user name",  
    "password": "password"  
}
```

También puede proporcionar credenciales de proxy web mediante un AWS Secrets Manager secreto. Utiliza la [ProxyConfigurationAPI](#) para proporcionar el nombre de host del sitio web y el número de puerto y, si lo deseas, el secreto que almacena sus credenciales de proxy web.

- IAMrol: especifique RoleArn cuándo llama `CreateDataSource` para proporcionar a un IAM rol los permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector del rastreador web y. Amazon Kendra Para obtener más información, consulte las [IAMfunciones de las fuentes de datos de los rastreadores web](#).

También puede añadir las siguientes funciones opcionales:

- Modo de rastreo: elija si desea rastrear solo los nombres de los sitios web o los nombres de los servidores con subdominios, o también rastrear otros dominios a los que se vinculan las páginas web.
- La «profundidad» o el número de niveles desde el nivel de semilla hasta el rastreo. Por ejemplo, la página URL inicial tiene la profundidad 1 y todos los hipervínculos de esta página que también se rastrean tienen la profundidad 2.
- El número máximo de URL que se pueden rastrear en una sola página web.
- El tamaño máximo en MB de una página web que se puede rastrear.
- El número de direcciones URL rastreadas por host de sitio web por minuto.

- El host del proxy web y el número de puerto para conectarse a los sitios web internos y rastrearlos. Por ejemplo, el nombre de host de `https://a.example.com/page1.htmles a.example.com »` y el número de puerto es 443, el puerto estándar para HTTPS. Si se requieren credenciales de proxy web para conectarse a un host de sitio web, puede crear una AWS Secrets Manager que almacene las credenciales.
- La información de autenticación para acceder y rastrear sitios web que requieren la autenticación del usuario.
- Puede extraer metaetiquetas HTML como campos mediante la herramienta de enriquecimiento de documentos personalizados. Para obtener más información, consulte [Customizing document metadata during the ingestion process](#) (Personalización de los metadatos del documento durante el proceso de ingestión). Para ver un ejemplo de cómo extraer metaetiquetas HTML, consulte los ejemplos [de CDE](#).
- Filtros de inclusión y exclusión: especifique si se van a incluir o excluir determinadas URL.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de web crawler datos, consulte:

- [Reimagine el descubrimiento de conocimientos con el Web Amazon Kendra Crawler](#)

Amazon KendraConektor Web Crawler v2.0

Puede utilizar Amazon Kendra Web Crawler para rastrear e indexar páginas web.

Solo puede rastrear sitios web públicos o sitios web internos de la empresa que utilicen el protocolo de comunicación seguro Hypertext Transfer Protocol Secure (HTTPS). Si recibe un error al rastrear un sitio web, es posible que el sitio web esté bloqueado para que no pueda rastrearse. Para rastrear sitios web internos, puedes configurar un proxy web. El proxy web debe estar orientado al público. También puede utilizar la autenticación para acceder a sitios web y rastrearlos.

Amazon KendraWeb Crawler v2.0 usa el paquete de rastreadores web Selenium y un controlador Chromium. Amazon Kendraactualiza automáticamente la versión de Selenium y el controlador Chromium mediante la integración continua (CI).

Al seleccionar los sitios web que se van a indexar, se debe respetar la [Política de uso aceptable de Amazon](#) y todas las demás condiciones de Amazon. Recuerde que solo debe utilizar Amazon Kendra Web Crawler para indexar sus propias páginas web o las páginas web que tenga autorización para indexar. Para obtener información sobre cómo evitar que Amazon Kendra Web Crawler indexe sus sitios web, consulte. [Configuración del robots.txt archivo para Amazon Kendra Web Crawler \(p. 324\)](#)

Note

El uso Amazon Kendra indebido de Web Crawler para rastrear de forma agresiva sitios web o páginas web que no son de su propiedad no se considera un uso aceptable.

Para solucionar problemas con el conector de fuente de datos del rastreador Amazon Kendra web, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 318\)](#)
- [Requisitos previos \(p. 318\)](#)
- [Instrucciones de conexión \(p. 319\)](#)

Características admitidas

- Mapeos de campo
- Proxy web
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/sincronizar solo documentos nuevos, modificados o eliminados
- Autenticación básica, NTLM/Kerberos, SAML y de formularios para sus sitios web

Requisitos previos

Antes de utilizarlos Amazon Kendra para indexar sus sitios web, compruebe los detalles de sus sitios web y AWS cuentas.

Para tus sitios web, asegúrate de tener:

- Copió las URL iniciales o del mapa del sitio web de los sitios web que desea indexar. Puedes almacenar las URL en un archivo de texto y subirlo a un Amazon S3 bucket. Cada URL del archivo de texto debe formatearse en una línea independiente. Si quieres almacenar tus mapas de sitio en un Amazon S3 depósito, asegúrate de haber copiado el XML del mapa del sitio y de guardarla en un archivo XML. También puedes agrupar varios archivos XML del mapa del sitio en un archivo ZIP.

Note

(En las instalaciones o en el servidor) Amazon Kendra comprueba si la información del punto final incluida en AWS Secrets Manager es la misma que la información del punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a evitar el [confuso problema del adjunto](#), que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, sino que Amazon Kendra lo utiliza como proxy para acceder al secreto configurado y realizar la acción. Si posteriormente cambias la información de tu terminal, debes crear un nuevo secreto para sincronizar esta información.

- Para sitios web que requieren autenticación básica, NTLM o Kerberos:
 - Tomó nota de las credenciales de autenticación de su sitio web, que incluyen un nombre de usuario y una contraseña.

Note

Amazon KendraWeb Crawler v2.0 admite el protocolo de autenticación NTLM, que incluye el cifrado de contraseñas, y el protocolo de autenticación Kerberos, que incluye el cifrado de contraseñas.

- Para sitios web que requieren autenticación mediante SAML o mediante un formulario de inicio de sesión:
 - Tomó nota de las credenciales de autenticación de su sitio web, que incluyen un nombre de usuario y una contraseña.
 - Copió los XPaths (lenguaje de rutas XML) del campo de nombre de usuario (y el botón del nombre de usuario si usa SAML), el campo de contraseña y el botón, y copió la URL de la página de inicio de sesión. Puedes encontrar los XPaths de los elementos utilizando las herramientas de desarrollo de tu navegador web. Los XPaths suelen seguir este formato://tagname[@Attribute='Value'].

Note

Amazon KendraWeb Crawler v2.0 usa un navegador Chrome sin encabezado y la información del formulario para autenticar y autorizar el acceso con una URL protegida con OAuth 2.0.

- Opcional: ha copiado el nombre del host y el número de puerto del servidor proxy web si desea utilizar un proxy web para conectarse a los sitios web internos que desea rastrear. El proxy web debe estar orientado al público. Amazon Kendraadmite la conexión a servidores proxy web respaldados por una autenticación básica o puede conectarse sin autenticación.
- Opcional: copia el ID de subred de la nube privada virtual (VPC) si quieres usar una VPC para conectarte a los sitios web internos que deseas rastrear. Para obtener más información, consulte [Configurar un Amazon VPC](#).
- Marcó que cada documento de página web que desea indexar es único y está en todas las demás fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que desee utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu AWS cuenta, asegúrate de tener:

- [Creó un Amazon Kendra índice](#) y, si utilizaba la API, anotó el ID del índice.
- [Creó una IAM función](#) para la fuente de datos y, si utilizaba la API, anotó el ARN de la IAM función.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al identificador AWS Secrets Manager secreto correcto.

- En el caso de los sitios web que requieren autenticación, o si utilizan un proxy web con autenticación, guarde las credenciales de autenticación en AWS Secrets Manager secreto y, si utiliza la API, anote el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. No se recomienda reutilizar las credenciales y los secretos en las fuentes de datos ni en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o un secreto existentes, puede usar la consola para crear un IAM rol y un Secrets Manager secreto nuevos al conectar su fuente de web crawler datosAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes, así como un identificador de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de web crawler datos, debe proporcionar los detalles necesarios de su fuente de web crawler datos para Amazon Kendra poder acceder a sus datos. Si aún no lo ha configurado web crawler, Amazon Kendra consulte [Requisitos previos \(p. 318\)](#).

Console

Para conectarse Amazon Kendra a web crawler

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, elija Índices y, a continuación, el índice que desee utilizar de la lista de índices.

Note

Puede configurar o editar los ajustes del control de acceso de usuario en la configuración del índice.

3. En la página de introducción, selecciona Agregar fuente de datos.
4. En la página Agregar fuente de datos, elija el conector para rastreadores web y, a continuación, elija Agregar fuente de datos.
5. En la página Especificar detalles de la fuente de datos, introduzca la siguiente información:
 - a. En Nombre y descripción, en Nombre de fuente de datos: introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. Descripción (opcional): introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, en Idioma predeterminado: un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, seleccione Agregar nueva etiqueta: etiquetas para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costos.
 - e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
 - a. Origen: elija las URL de origen, los mapas del sitio de origen, el archivo de URL de origen o el archivo de mapas de sitio de origen. Si decide utilizar un archivo de texto que incluya una lista de hasta 100 URL iniciales, especifique la ruta al Amazon S3 depósito en el que se almacena el archivo. Si decide utilizar un archivo XML de mapa del sitio, especifique la ruta al Amazon S3 depósito en el que se almacena el archivo. También puedes agrupar varios archivos XML del mapa del sitio en un archivo ZIP. De lo contrario, puedes introducir manualmente hasta 10 URL iniciales o de punto de inicio y hasta tres URL del mapa del sitio.

Note

Si quieras rastrear un mapa del sitio, comprueba que la URL base o raíz sea la misma que las URL que aparecen en la página del mapa del sitio. Por ejemplo, si la URL del mapa del sitio es <https://example.com/sitemap-page.html>, las URL que aparecen en esta página del mapa del sitio también deben usar la URL base "[https://example.com/»](https://example.com/).

Si sus sitios web requieren autenticación para acceder a ellos, puede elegir la autenticación básica, NTLM/Kerberos, SAML o mediante formularios. De lo contrario, elija la opción de no autenticación.

Note

Si quieras editar tu fuente de datos más adelante para cambiar las URL iniciales con la autenticación de los mapas de sitio, debes crear una nueva fuente de datos. Amazon Kendra configura la fuente de datos mediante la información de punto final de las URL de origen que figura en el Secrets Manager secreto para la autenticación y, por lo tanto, no puede volver a configurar la fuente de datos cuando se cambia a mapas de sitio.

- AWS Secrets Manager secreto: si sus sitios web requieren la misma autenticación para acceder a los sitios web, elija un secreto existente o cree uno nuevo Secrets Manager para almacenar las credenciales de su sitio web. Si decide crear un secreto nuevo, se abre una ventana AWS Secrets Manager secreta.

Si elige la autenticación básica o NTLM/Kerberos, introduzca un nombre para el secreto, además del nombre de usuario y la contraseña. El protocolo de autenticación NTLM

incluye el cifrado de contraseñas y el protocolo de autenticación Kerberos incluye el cifrado de contraseñas.

Si elige la autenticación mediante SAML o mediante formulario, introduzca un nombre para el secreto, además del nombre de usuario y la contraseña. Use XPath para el campo de nombre de usuario (y XPath para el botón de nombre de usuario si usa SAML). Utilice XPaths para el campo de contraseña y el botón y la URL de la página de inicio de sesión. Puedes encontrar los XPaths (XML Path Language) de los elementos utilizando las herramientas de desarrollo de tu navegador web. Los XPaths suelen seguir este formato://tagname[@Attribute='Value'].

- b. Proxy web (opcional): introduzca el nombre de host y el número de puerto del servidor proxy que desea utilizar para conectarse a sitios web internos. Por ejemplo, el nombre de host de `https://a.example.com/page1.html` es `a.example.com` » y el número de puerto es 443, el puerto estándar para HTTPS. Si se requieren credenciales de proxy web para conectarse a un host de sitio web, puede crear una AWS Secrets Manager que almacene las credenciales.
- c. Nube privada virtual (VPC): puedes elegir usar una VPC. Si es así, debe añadir subredes y grupos de seguridad de VPC.
- d. IAMrol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un rol nuevo para evitar errores.

- e. Elija Siguiente.
7. En la página Configurar los ajustes de sincronización, introduzca la siguiente información:
 - a. Alcance de sincronización: establece límites para rastrear páginas web, incluidos sus dominios, tamaños de archivo y enlaces; y filtra las URL mediante patrones de expresiones regulares.
 - i. (Opcional) Rastrear el rango de dominios: elija si desea rastrear los dominios de sitios web únicamente con subdominios o también rastrear otros dominios a los que se vinculan las páginas web. De forma predeterminada, Amazon Kendra solo rastrea los dominios de los sitios web que deseas rastrear.
 - ii. Configuración adicional (opcional): defina los siguientes parámetros:
 - Profundidad de rastreo: la «profundidad» o el número de niveles desde el nivel inicial hasta el rastreo. Por ejemplo, la página URL inicial tiene la profundidad 1 y todos los hipervínculos de esta página que también se rastrean tienen la profundidad 2.
 - Tamaño máximo de archivo: el tamaño máximo en MB de una página web o un archivo adjunto que se puede rastrear.
 - Número máximo de enlaces por página: el número máximo de URL que se pueden rastrear en una sola página web.
 - Limitación máxima de la velocidad de rastreo: la cantidad máxima de URL rastreadas por servidor de sitio web por minuto.
 - Archivos: elija rastrear los archivos a los que enlazan las páginas web.
 - Rastreo e indexación de URL: añada patrones de expresiones regulares para incluir o excluir el rastreo de determinadas URL y la indexación de cualquier hipervínculo en estas páginas web de URL.
 - b. Modo de sincronización: elija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Al sincronizar la fuente de datos con Amazon Kendra por primera vez, todo el contenido se sincroniza de forma predeterminada.

- Sincronización completa: sincroniza todo el contenido independientemente del estado de sincronización anterior.
 - Sincronización de documentos nuevos, modificados o eliminados: sincronice solo los documentos nuevos, modificados y eliminados.
- c. Programa de ejecución de sincronización: en Frecuencia, elija la frecuencia con la Amazon Kendra que se sincronizará con la fuente de datos.
- d. Elija Siguiente.
8. En la página Establecer mapeos de campos, introduzca la siguiente información:
- a. Seleccione uno de los campos predeterminados Amazon Kendra generados de las páginas web y los archivos que deseé asignar a su índice.
 - b. Elija Siguiente.
9. En la página Revisar y crear, compruebe que la información que ha introducido sea correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. La fuente de datos aparecerá en la página Fuentes de datos cuando se haya agregado correctamente.

API

Para conectarse Amazon Kendra a web crawler

Debe especificar un JSON del [esquema de la fuente de datos](#) mediante la [TemplateConfigurationAPI](#). Debe proporcionar la siguiente información:

- Fuente de datos: especifique la fuente de datos comoWEBCRAWLERV2.
- URL: especifique las URL iniciales o de punto de inicio de los sitios web o las URL del mapa del sitio web que desea rastrear. Puedes especificar la ruta a un Amazon S3 bucket que almacene tu lista de URL iniciales. Cada URL del archivo de texto de las URL iniciales debe formatearse en una línea independiente. También puedes especificar la ruta a un Amazon S3 depósito que almacena los archivos XML del mapa del sitio. Puedes agrupar varios archivos del mapa del sitio en un archivo ZIP y almacenar el archivo ZIP en tu Amazon S3 bucket.

Note

Si quieres rastrear un mapa del sitio, comprueba que la URL base o raíz sea la misma que las URL que aparecen en la página del mapa del sitio. Por ejemplo, si la URL del mapa del sitio es <https://example.com/sitemap-page.html>, las URL que aparecen en esta página del mapa del sitio también deben usar la URL base "<https://example.com/>".

- Tipo: especifique TEMPLATE como el tipo cuando llameCreateDataSource.
- Autenticación: si sus sitios web requieren la misma autenticación, especifique la autenticaciónBasicAuth, NTLM_KerberosSAML, o la Form autenticación. Si tus sitios web no requieren autenticación, especificaloNoAuthentication.
- Nombre secreto de recurso de Amazon (ARN): si sus sitios web requieren una autenticación básica, NTLM o Kerberos, debe proporcionar un secreto que almacene las credenciales de autenticación de su nombre de usuario y contraseña. Usted proporciona el nombre de recurso de Amazon (ARN) de un AWS Secrets Manager secreto. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "seedUrlsHash": "Hash representation of all seed URLs",  
    "userName": "user name",  
    "password": "password"  
}
```

Si sus sitios web requieren la autenticación SAML, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "seedUrlsHash": "Hash representation of all seed URLs",  
  
    "userName": "user name",  
    "password": "password",  
    "userNameFieldXpath": "XPath for user name field",  
    "userNameButtonXpath": "XPath for user name button",  
    "passwordFieldXpath": "XPath for password field",  
    "passwordButtonXpath": "XPath for password button",  
    "loginPageUrl": "Full URL for website login page"  
}
```

Si sus sitios web requieren autenticación mediante formularios, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "seedUrlsHash": "Hash representation of all seed URLs",  
    "userName": "user name",  
    "password": "password",  
    "userNameFieldXpath": "XPath for user name field",  
    "passwordFieldXpath": "XPath for password field",  
    "passwordButtonXpath": "XPath for password button",  
    "loginPageUrl": "Full URL for website login page"  
}
```

Puedes encontrar los XPaths (XML Path Language) de los elementos utilizando las herramientas de desarrollo de tu navegador web. Los XPaths suelen seguir este formato:// tagname[@Attribute='Value'].

También puede proporcionar credenciales de proxy web mediante AWS Secrets Manager un secreto.

- IAMrol: especifique RoleArn cuándo llama CreateDataSource para proporcionar a un IAM rol los permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector del rastreador web y Amazon Kendra Para obtener más información, consulte las [IAMfunciones de las fuentes de datos de los rastreadores web](#).

También puede añadir las siguientes funciones opcionales:

- Intervalo de dominios: elija si desea rastrear los dominios del sitio web únicamente con subdominios o también rastrear otros dominios a los que se vinculan las páginas web. De forma predeterminada, Amazon Kendra solo rastrea los dominios de los sitios web que deseas rastrear.
- La «profundidad» o el número de niveles desde el nivel de semilla hasta el rastreo. Por ejemplo, la página URL inicial tiene la profundidad 1 y todos los hipervínculos de esta página que también se rastrean tienen la profundidad 2.
- El número máximo de URL que se pueden rastrear en una sola página web.
- El tamaño máximo en MB de una página web o un archivo adjunto para rastrear.
- El número de direcciones URL rastreadas por host de sitio web por minuto.
- El host del proxy web y el número de puerto para conectarse a los sitios web internos y rastrearlos. Por ejemplo, el nombre de host de https://a.example.com/page1.htmls a.example.com » y el número de puerto es 443, el puerto estándar para HTTPS. Si se requieren credenciales de proxy web para conectarse a un host de sitio web, puede crear una AWS Secrets Manager que almacene las credenciales.

- Virtual Private Cloud (VPC): especifique VpcConfiguration cuando llama a `CreateDataSource`. Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).
- Asignaciones de campos: elija asignar los campos de las páginas web y los archivos de páginas web a los campos de índice. Amazon Kendra Para obtener más información, consulte [Mapping data source fields \(Asignación de campos de origen de datos\)](#).
- Modo de sincronización: especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:
 - FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice.
 - FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice.
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir el rastreo de determinadas URL y la indexación de cualquier hipervínculo en estas páginas web de URL.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

Para obtener una lista de otras claves JSON importantes para configurar, consulte el [esquema de plantillas de Amazon Kendra Web Crawler](#).

Configuración del robots.txt archivo para Amazon Kendra Web Crawler

Amazon Kendra es un servicio de búsqueda inteligente que AWS los clientes utilizan para indexar y buscar documentos de su elección. Para indexar documentos en la web, los clientes pueden utilizar Amazon Kendra Web Crawler, que indica qué URL deben indexarse y otros parámetros operativos. Amazon Kendra los clientes deben obtener una autorización antes de indexar cualquier sitio web en particular.

Amazon Kendra Web Crawler respeta las directivas estándar de robots.txt, como Allow y Disallow. Puede modificar el robots.txt archivo de su sitio web para controlar la forma en que Amazon Kendra Web Crawler rastrea su sitio web.

Configurar la forma en que Amazon Kendra Web Crawler accede a su sitio web

Puede controlar la forma en que el Amazon Kendra Web Crawler indexa su sitio web mediante directivas Allow y Disallow. También puede controlar qué páginas web se indexan y qué páginas web no se rastrean.

Para permitir que Amazon Kendra Web Crawler rastree todas las páginas web, excepto las páginas web no permitidas, utilice la siguiente directiva:

```
User-agent: amazon-kendra    # Amazon Kendra Web Crawler
Disallow: /credential-pages/ # disallow access to specific pages
```

Para permitir que Amazon Kendra Web Crawler rastree solo páginas web específicas, utilice la siguiente directiva:

```
User-agent: amazon-kendra # Amazon Kendra Web Crawler
Allow: /pages/ # allow access to specific pages
```

Para permitir que Amazon Kendra Web Crawler rastree todo el contenido del sitio web y no permitir el rastreo de ningún otro robot, utilice la siguiente directiva:

```
User-agent: amazon-kendra # Amazon Kendra Web Crawler
Allow: / # allow access to all pages
User-agent: * # any (other) robot
Disallow: / # disallow access to any pages
```

Impedir que Amazon Kendra Web Crawler rastree tu sitio web

Puede impedir que Amazon Kendra Web Crawler indexe su sitio web mediante la `Disallow` directiva. También puedes controlar qué páginas web se rastrean y cuáles no.

Para evitar que Amazon Kendra Web Crawler rastree el sitio web, utilice la siguiente directiva:

```
User-agent: amazon-kendra # Amazon Kendra Web Crawler
Disallow: / # disallow access to any pages
```

Amazon KendraWeb Crawler también admite los robots `noindex` y `nofollow` las directivas de las metaetiquetas de las páginas HTML. Estas directivas impiden que el rastreador web indexe una página web y deja de seguir los enlaces de la página web. Las metaetiquetas se colocan en la sección del documento para especificar las reglas de los robots.

Por ejemplo, la siguiente página web incluye las directivas robots `noindex` y `nofollow`:

```
<html>
<head>
    <meta name="robots" content="noindex, nofollow"/>
    ...
</head>
<body>...</body>
</html>
```

Si tiene alguna pregunta o duda sobre Amazon Kendra Web Crawler, puede ponerse en contacto con el [equipo de AWS soporte](#).

Amazon WorkDocs

Amazon WorkDocses un servicio seguro de colaboración de contenido para crear, editar, almacenar y compartir contenido. Se puede utilizar Amazon Kendra para indexar la fuente Amazon WorkDocs de datos.

Puede conectarse Amazon Kendra a la fuente Amazon WorkDocs de datos mediante la [Amazon Kendraconsola](#) y la [WorkDocsConfigurationAPI](#).

Amazon WorkDocsestá disponible en las regiones de Oregón, Virginia del Norte, Sídney, Singapur e Irlanda.

Para solucionar problemas con el conector de la fuente de Amazon Kendra WorkDocs datos, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 326\)](#)

- [Requisitos previos \(p. 326\)](#)
- [Instrucciones de conexión \(p. 326\)](#)
- [Más información \(p. 328\)](#)

Características admitidas

Amazon KendraWorkDocsel conector de fuente de datos admite las siguientes funciones:

- Change log
- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder utilizarla Amazon Kendra para indexar la fuente de WorkDocs datos, realice estos cambios en sus AWS cuentas WorkDocs de correo electrónico.

EnWorkDocs, asegúrese de tener:

- Apuntó el ID del Amazon WorkDocs directorio (ID de organización) de tu Amazon WorkDocs repositorio.
- Se ha comprobado que cada documento es único en WorkDocs las demás fuentes de datos que piensa utilizar para el mismo índice y entre ellas. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu AWS cuenta, asegúrate de tener:

- [Creó un Amazon Kendra índice](#) y, si utilizaba la API, anotó el ID del índice.
- [Creó una IAM función](#) para la fuente de datos y, si utilizaba la API, anotó el ARN de la IAM función.

Si no tiene un IAM rol existente, puede usar la consola para crear un IAM rol nuevo cuando conecte su fuente de WorkDocs datos aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de un IAM rol existente y un ID de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de WorkDocs datos, debe proporcionar los detalles necesarios de su fuente de WorkDocs datos para Amazon Kendra poder acceder a sus datos. Si aún no ha configurado WorkDocs paraAmazon Kendra, consulte[Requisitos previos \(p. 326\)](#).

Console

Para conectar Amazon Kendra a Amazon WorkDocs

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, elija Índices y, a continuación, el índice que deseé utilizar de la lista de índices.

Note

Puede configurar o editar los ajustes del control de acceso de usuario en la configuración del índice.

3. En la página de introducción, selecciona Agregar fuente de datos.
4. En la página Agregar fuente de datos, elija WorkDocsconector y, a continuación, elija Agregar fuente de datos.
5. En la página Especificar detalles de la fuente de datos, introduzca la siguiente información:
 - a. En Nombre y descripción, en Nombre de fuente de datos: introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. Descripción (opcional): introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, en Idioma predeterminado: un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, seleccione Agregar nueva etiqueta: etiquetas para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costos.
 - e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
 - a. ID de organización específico de tu Amazon WorkDocs sitio: selecciona el ID del Amazon WorkDocs sitio que deseas indexar. Ya debe haber creado un sitio.
 - b. IAMrol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un rol nuevo para evitar errores.

- c. Elija Siguiente.
7. En la página Configurar los ajustes de sincronización, introduzca la siguiente información:
 - a. Rastrear comentarios de documentos: Amazon WorkDocs las entidades o los tipos de contenido que desea rastrear.
 - b. Utilice registros de cambios: seleccione esta opción para actualizar el índice en lugar de sincronizar todos los archivos.
 - c. Patrones de expresiones regulares: patrones de expresiones regulares para incluir o excluir ciertos archivos. Puede añadir hasta 100 patrones.
 - d. En Sincronizar programación de ejecución por frecuencia: elija la frecuencia con la Amazon Kendra que se sincronizará con la fuente de datos.
 - e. Elija Siguiente.
8. En la página Establecer mapeos de campos, introduzca la siguiente información:
 - a. Campos de fuente de datos predeterminados: seleccione uno de los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar al índice.
 - b. Agregar campo: para agregar campos de fuente de datos personalizados para crear un nombre de campo de índice al que asignar y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En la página Revisar y crear, compruebe que la información que ha introducido sea correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. La fuente de datos aparecerá en la página Fuentes de datos cuando se haya agregado correctamente.

Debe especificar lo siguiente mediante la [WorkDocsConfigurationAPI](#):

- Amazon WorkDocsID de directorio: especifique el identificador de la organización de su Amazon WorkDocs directorio. Para encontrar el identificador de la organización en el servicio de directorio de AWS, vaya a Active Directory y, a continuación, a Directories.
- Función de IAM: especifique RoleArn cuándo llama `CreateDataSource` para proporcionar a una IAM función los permisos de acceso al WorkDocs directorio y para llamar a las API públicas necesarias para el conector y. WorkDocs Amazon Kendra Para obtener más información, consulte [Funciones de IAM para fuentes WorkDocs de datos](#).

También puede añadir las siguientes funciones opcionales:

- Registro de cambios: si se Amazon Kendra debe utilizar el mecanismo de registro de cambios de la fuente de WorkDocs datos para determinar si un documento debe agregarse, actualizarse o eliminarse del índice.

Note

Usa el registro de cambios si no Amazon Kendra quieres escanear todos los documentos. Si el registro de cambios es grande, puede llevar Amazon Kendra menos tiempo escanear los documentos de la fuente de WorkDocs datos que procesar el registro de cambios. Si sincroniza la fuente de WorkDocs datos con el índice por primera vez, se escanean todos los documentos.

- Filtros de inclusión y exclusión: especifique si se van a incluir o excluir ciertos documentos y comentarios de documentos. Cada comentario se indexa como un documento independiente.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Asignaciones de campos: elija asignar los campos de la fuente WorkDocs de datos Amazon Kendra a los campos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario: Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes WorkDocs de datos](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de WorkDocs datos, consulte:

- [Comience a utilizar el WorkDocs conector Amazon Kendra de Amazon](#)

Box

Box es un servicio de almacenamiento en la nube que ofrece capacidades de alojamiento de archivos. Puedes usar Amazon Kendra para indexar el contenido de tu Box, incluidos los comentarios, las tareas y los enlaces web.

Puedes conectarte a Amazon Kendra a su fuente de datos de Box mediante el [Amazon Kendra consola](#) y el [Box Configuration API](#).

Para solucionar problemas de Amazon Kendra Conector de fuente de datos de Box, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 329\)](#)
- [Requisitos previos \(p. 329\)](#)
- [Instrucciones de conexión \(p. 330\)](#)
- [Más información \(p. 333\)](#)

Características admitidas

Amazon Kendra El conector de fuente de datos de Box admite las siguientes funciones:

- Change log
- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de poder usar Amazon Kendra para indexar su fuente de datos de Box, realice estos cambios en su Box y AWS cuentas.

En Box, asegúrate de tener:

- Una cuenta de Box Enterprise o Box Enterprise Plus.
- Creé una aplicación personalizada de Box en la consola para desarrolladores de Box y la configuré para usar Autenticación de servidor (con JWT). Consulte [Documentación de Box sobre la creación de una aplicación personalizada](#) y [Documentación de Box sobre la configuración de JWT Auth](#) para obtener más información.
- Configura tu Nivel de acceso a la aplicación + Acceso empresarial y permitió que Realice llamadas a la API mediante el encabezado as-user.
- Usó el usuario administrador para agregar lo siguiente Ámbitos de aplicación en tu aplicación Box:
 - Escribe todos los archivos y carpetas almacenados en una caja
 - Administrar usuarios
 - Administrar grupos
 - Administrar propiedades empresariales
- Par de claves públicas/privadas generado y descargado que incluye un ID de cliente, un secreto de cliente, un ID de clave pública, un ID de clave privada, una contraseña y un ID de empresa para usarlos como credenciales de autenticación. Consulte [Par de claves público y privado](#) para obtener más información.
- Copió su ID empresarial de Box desde la configuración de la consola para desarrolladores de Box o desde la aplicación Box. Por ejemplo, [801234567](#).
- Marcó que cada documento es único en Box y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapelpara su fuente de datos y](#), si usa la API, anotó el ARN dellAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Has guardado tus credenciales de autenticación de Box en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Box aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa su fuente de datos de Box, debe proporcionar los detalles necesarios de su fuente de datos de Box para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configurado Box paraAmazon Kendra, consulte[Requisitos previos \(p. 329\)](#).

Console

Para conectarAmazon Kendraa Box

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConejor de cajay, a continuación, elijaAregar fuente de datos.
5. En elEspecificardetalles de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:

- a. ID empresarial de Box—Introduzca su ID empresarial de Box.
- b. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar sus credenciales de autenticación de Box. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - i. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Box-' se añade automáticamente a tu nombre secreto.
 - ii. ParaID de cliente,Secreto del cliente, ID de clave pública, ID de clave privada, yFrase de contraseña—Introduzca los valores de la clave pública/privada que generó en su cuenta de Box y que descargó de su cuenta de Box.
 - iii. Seleccione Guardar.
- c. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesyGrupos de seguridad de VPC.
- d. IAMpapel—Elige uno existenteIAMrol o crea uno nuevolAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- e. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. Seleccione entidades o tipos de contenido—Las entidades de la caja o los tipos de contenido que desea rastrear. Cada comentario se indexa como un documento independiente.
 - b. Registro de cambios—Seleccione esta opción para actualizar el índice en lugar de sincronizar todos los archivos.
 - c. Patrones de expresiones regulares—Patrones de expresiones regulares para incluir o excluir ciertos archivos. Puede añadir hasta 100 patrones.
 - d. EnCronograma de ejecución de sincronización, paraFrecuencia—Elige con qué frecuenciaAmazon Kendrase sincronizará con su fuente de datos.
 - e. Elija Siguiente.
8. En elDefinir mapeos de campospágina, introduzca la siguiente información:
 - a. ParaArchivos y carpetas,Comentarios,Tareas, yEnlaces web—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon Kendraa Box

Debe especificar lo siguiente mediante el[BoxConfigurationAPI](#):

ID empresarial de Box—Proporcione su ID empresarial de Box. Puedes encontrar el ID de empresa en la configuración de la consola para desarrolladores de Box o al crear una aplicación en Box.

- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de su cuenta de Box. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "clientID": "client-id",  
    "clientSecret": "client-secret",  
    "publicKeyID": "public-key-id",  
    "privateKey": "private-key",  
    "passphrase": "pass-phrase"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMPapel—EspecificarRoleArn cuando llamasCreateDataSourcepara proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para el conector Box yAmazon Kendra. Para obtener más información, consulte[AMfunciones para las fuentes de datos de Box](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfigurationcomo parte de la configuración de la fuente de datos. Consulte[ConfigurandoAmazon Kendrausar una VPC](#).
- Registro de cambios—SiAmazon Kendradebe utilizar el mecanismo de registro de cambios en la fuente de datos de Box para determinar si un documento debe agregarse, actualizarse o eliminarse del índice.

Note

Usa el registro de cambios si no quieresAmazon Kendrapara escanear todos los documentos. Si el registro de cambios es grande, podría tardarAmazon Kendramenos tiempo para escanear los documentos de la fuente de datos de Box que para procesar el registro de cambios. Si sincroniza su fuente de datos de Box con su índice por primera vez, se escanean todos los documentos.

- Filtros de inclusión y exclusión—Especifique si desea incluir o excluir ciertos archivos, carpetas, comentarios, tareas y enlaces web de Box.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elja asignar los campos de la fuente de datos de Box a suAmazon Kendracampos de índice. Para obtener más información, consulte[Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendrarastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte[Filtrado de contexto de usuario para fuentes de datos de Box](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Box, consulte:

- [Empezar con el Amazon Kendra Conector de caja](#)

Confluence

Confluence es una herramienta colaborativa de gestión del trabajo diseñada para compartir, almacenar y trabajar en la planificación de proyectos, el desarrollo de software y la gestión de productos. Puedes usar Amazon Kendra para indexar tus espacios, páginas (incluidas las páginas anidadas), blogs y comentarios y archivos adjuntos de páginas y blogs indexados.

Amazon Kendra es compatible con Confluence Server y Confluence Cloud.

Note

De forma predeterminada, Amazon Kendra no indexa los archivos ni los espacios personales de Confluence. Puede optar por indexarlos al crear la fuente de datos. Si no quieres que Amazon Kendra indexe un espacio, márcalo como privado en Confluence.

Puedes conectarte a Amazon Kendra a tu fuente de datos de Confluence mediante el [Amazon Kendra consola](#), el [Template Configuration API](#), o [Confluence Configuration API](#).

Amazon Kendra tiene dos versiones del conector Confluence. Las funciones compatibles de cada versión incluyen:

Conector Confluence V1.0/[Confluence Configuration API](#)

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- (Solo para Confluence Server) Nube privada virtual (VPC)

Conector Confluence V2.0/[Template Configuration API](#)

- Mapeos de campo
- Filtrado de contexto de usuario
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados
- Patrones de inclusión/exclusión

Note

Soporte para el conector Confluence V1.0/Confluence Configuration API. Está previsto que la API finalice en 2023. Recomendamos migrar a Confluence connector V2.0/Template Configuration API.

Para solucionar problemas de su Amazon Kendra Conector de fuente de datos de Confluence, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Conector Confluence V1.0 \(p. 334\)](#)
- [Conector Confluence V2.0 \(p. 339\)](#)

Conejero Confluence V1.0

Confluence es una herramienta colaborativa de gestión del trabajo diseñada para compartir, almacenar y trabajar en la planificación de proyectos, el desarrollo de software y la gestión de productos. Puedes usar Amazon Kendra para indexar tus espacios, páginas (incluidas las páginas anidadas), blogs y comentarios y archivos adjuntos de páginas y blogs indexados.

Note

Soporte para el conector Confluence V1.0/Confluence Configuration API. Está previsto que la API finalice en 2023. Recomendamos migrar a Confluence connector V2.0/Template Configuration API.

Para solucionar problemas de su Amazon Kendra Conejero de fuente de datos de Confluence, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 334\)](#)
- [Requisitos previos \(p. 334\)](#)
- [Instrucciones de conexión \(p. 335\)](#)
- [Más información \(p. 338\)](#)

Características admitidas

Amazon Kendra El conector de fuente de datos de Confluence admite las siguientes funciones:

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- (Solo para Confluence Server) Nube privada virtual (VPC)

Requisitos previos

Antes de poder usar Amazon Kendra para indexar tu fuente de datos de Confluence, realiza estos cambios en tu Confluence y AWS cuentas.

En Confluence, asegúrate de tener:

- Concedido Amazon Kendra permisos para ver todo el contenido de tu instancia de Confluence mediante:
 - Elaboración Amazon Kendra miembro de confluence-administrators grupo.
 - Otorgar permisos de administrador del sitio para todos los espacios, blogs y páginas existentes.
- Copiate la URL de tu instancia de Confluence.
- Para los usuarios de SSO (inicio de sesión único): Activó el Mostrar en la página de inicio de sesión para el nombre de usuario y la contraseña al configurar Confluence Métodos de autenticación en Confluence Data Center.
- Para Confluence Server
 - Anoté tus credenciales de autenticación básicas, que contienen el nombre de usuario y la contraseña de tu cuenta administrativa de Confluence para conectarte a Amazon Kendra.
 - Opcional: Has generado un token de acceso personal en tu cuenta de Confluence para conectarte a Amazon Kendra. Para obtener más información, consulte [Documentación de Confluence sobre la generación de tokens de acceso personales](#).
- Para Confluence Cloud
 - Anoté tus credenciales de autenticación básicas, que contienen el nombre de usuario y la contraseña de tu cuenta administrativa de Confluence para conectarte a Amazon Kendra.

- Has comprobado que cada documento es único en Confluence y en otras fuentes de datos que planeas usar para el mismo índice. Cada fuente de datos que deseas utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- Creó unAmazon Kendraíndicey, si usa la API, anota el ID del índice.
- Creó unIAMpapelpara su fuente de datos y, si usa la API, anotó el ARN dellAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Has guardado tus credenciales de autenticación de Confluence en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Confluence aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tu fuente de datos de Confluence, debes proporcionar detalles de tus credenciales de Confluence para queAmazon Kendrapuede acceder a sus datos. Si aún no has configurado Confluence paraAmazon KendraverRequisitos previos (p. 334).

Console

Para conectarAmazon Kendraa Confluence

- Inicia sesión en elAWSConsola de administración y abra elAmazon Kendraconsola.
- En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseas utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

- En elPrimeros pasospágina, eligeAregar fuente de datos.
- En elAregar fuente de datospágina, eligeConejor Confluence V1.0y, a continuación, elijaAregar fuente de datos.
- En elEspecificardetalles de la fuente de datospágina, introduzca la siguiente información:
 - EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

- d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:
 - a. Elige entreNube de confluenciayServidor Confluencesegún su caso de uso.
 - b. Si eligesNube de confluencia, introduzca la siguiente información:
 - i. URL de Confluence—Tu URL de Confluence.
 - ii. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar tus credenciales de autenticación de Confluence. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - Introduzca la siguiente información enCrea unAWS Secrets Managerventana secreta:
 - I. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Confluence-' se añade automáticamente a tu nombre secreto.
 - II. ParaNombre de usuarioyContraseña—Introduce tu nombre de usuario de Confluence y tu token de API de Confluence como contraseña.
 - III. EscojaGuardar autenticación.
 - c. Si eligesServidor Confluence, introduzca la siguiente información:
 - i. URL de Confluence—Tu nombre de usuario y contraseña de Confluence.
 - ii. (Opcional) ParaProxy webintroduzca la siguiente información:
 - A. Nombre del host—Nombre de host de tu cuenta de Confluence.
 - B. Número de puerto—Puerto utilizado por el protocolo de transporte de URL del host.
 - iii. Elige entreAutenticación básica yToken de acceso personal.
 - iv. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar tus credenciales de autenticación de Confluence. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - Introduzca la siguiente información enCrea unAWS Secrets Managerventana secreta:
 - I. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Confluence-' se añade automáticamente a tu nombre secreto.
 - II. ParaNombre de usuarioyContraseña—Introduce los valores de las credenciales de autenticación que generaste y descargaste de tu cuenta de Confluence. Si usas la autenticación básica, usa tu nombre de usuario y contraseña de Confluence como credencial de autenticación. Si utiliza un token de acceso personal, introduzca los detalles delToken de acceso personalque creaste en tu cuenta de Confluence.
 - III. EscojaGuardar autenticación.
 - d. IAMpapel—Elige uno existenteIAMrol o crea uno nuevolAMrol para acceder a las credenciales del repositorio e indexar el contenido.
- Note
- IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.
- e. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:

- a. Para incluir espacios personales y incluir espacios archivados: elija los tipos de espacio opcionales que se van a incluir en esta fuente de datos.
 - b. Para Configuración adicional—Especifique patrones de expresiones regulares para incluir o excluir cierto contenido. Puede añadir hasta 100 patrones.
 - c. También puedes elegir Rastrea los archivos adjuntos dentro de los espacios seleccionados.
 - d. En Cronograma de ejecución de sincronización, para Frecuencia—Elige con qué frecuencia Amazon Kendra se sincronizará con su fuente de datos.
 - e. Elija Siguiente.
8. En el Definir mapeos de campos de página, introduzca la siguiente información:
 - a. Para Espacio, página, Blog—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados o Mapeos de campos sugeridos adicionales para añadir campos de índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
 9. En el Revisar y crear página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en Fuentes de datos de página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Confluence

Debe especificar lo siguiente mediante [Confluence Configuration API](#):

- Versión Confluence—Especifica la versión de la instancia de Confluence como CLOUD SERVER.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contiene las credenciales de autenticación que creaste en tu cuenta de Confluence.

Si usas Confluence Server, puedes usar tu nombre de usuario y contraseña de Confluence o tu token de acceso personal como credenciales.

Cuando usas tu nombre de usuario y contraseña de Confluence como credenciales de autenticación, guardas las siguientes credenciales como una estructura JSON en tu Secrets Manager secreto:

```
{  
    "username": "user name",  
    "password": "password"  
}
```

Si utilizas un token de acceso personal para conectar Confluence Server a Amazon Kendra, almacena las siguientes credenciales como una estructura JSON en su Secrets Manager secreto:

```
{  
    "patToken": "personal access token"  
}
```

Si utilizas Confluence Cloud como Amazon Kendra fuente de datos, utilizas tu nombre de usuario de Confluence y un token de API generado en tu cuenta de Confluence como contraseña. Almacena las siguientes credenciales como una estructura JSON en su Secrets Manager secreto:

```
{  
    "username": "user name",  
    "password": "API token"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSourcepara proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas necesarias para el conector de Confluence yAmazon Kendra. Para obtener más información, consulte[IAMfunciones para las fuentes de datos de Confluence](#).

También puede añadir las siguientes funciones opcionales:

- Proxy web—Si debes conectarte a tu instancia URL de Confluence mediante un proxy web. Puedes usar esta opción para Confluence Server.
- (Solo para Confluence Server)Nube privada virtual (VPC)—EspecificarVpcConfigurationcomo parte de la configuración de la fuente de datos. Consulte[ConfiguraciónAmazon Kendrausar una VPC](#).
- Filtros de inclusión y exclusión—Especifique patrones de expresiones regulares para incluir o excluir ciertos espacios, publicaciones de blog, páginas, espacios y archivos adjuntos. Si elige indexar los archivos adjuntos, solo se indexarán los archivos adjuntos a las páginas y blogs indexados.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elige asignar los campos de tu fuente de datos de Confluence a tuAmazon Kendracampos de índice. Para obtener más información, consulte[Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendrarastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte[Filtrado de contexto de usuario para fuentes de datos de Confluence](#).

Más información

Para obtener más información sobre la integraciónAmazon Kendracon tu fuente de datos de Confluence, consulta:

- [Configuración de suAmazon KendraConejor de servidor Confluence](#)

Conejor Confluence V2.0

Confluence es una herramienta colaborativa de gestión del trabajo diseñada para compartir, almacenar y trabajar en la planificación de proyectos, el desarrollo de software y la gestión de productos. Puedes usar Amazon Kendrapara indexar tus espacios, páginas (incluidas las páginas anidadas), blogs y comentarios y archivos adjuntos de páginas y blogs indexados.

Para solucionar problemas de Amazon KendraConejor de fuente de datos de Confluence, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 339\)](#)
- [Requisitos previos \(p. 339\)](#)
- [Instrucciones de conexión \(p. 340\)](#)

Características admitidas

Amazon KendraEl conector de fuente de datos de Confluence admite las siguientes funciones:

- Mapeos de campo
- Filtrado de contexto de usuario
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados
- Patrones de inclusión/exclusión

Requisitos previos

Antes de poder usar Amazon Kendrapara indexar tu fuente de datos de Confluence, realiza estos cambios en tu Confluence y AWS cuentas.

En Confluence, asegúrate de tener:

- Copiate la URL de tu instancia de Confluence. Por ejemplo:<https://example.confluence.com>. Necesitas la URL de tu instancia de Confluence para conectarte a Amazon Kendra.

Note

(En las instalaciones o en el servidor)Amazon Kendrapara comprueba si la información del punto final está incluida en AWS Secrets Manageres la misma que la información de punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a proteger contra la [problema de diputado confuso](#), que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, sino que usa Amazon Kendrapara como proxy para acceder al secreto configurado y realizar la acción. Si posteriormente cambias la información de tu terminal, debes crear un nuevo secreto para sincronizar esta información.

- Credenciales de autenticación básicas configuradas que contienen un nombre de usuario (ID de correo electrónico que se utiliza para iniciar sesión en Confluence) y una contraseña (contraseña del servidor de Confluence) para permitir Amazon Kendrapara conectarte a tu instancia de Confluence. Para obtener información sobre cómo crear un token de API de Confluence, consulta [Gestiona los tokens de API para tu cuenta de Atlassian](#).
- Opcional:Credenciales de OAuth 2.0 configuradas que contienen una clave de aplicación de Confluence, un secreto de aplicación de Confluence, un token de acceso a Confluence y un token de actualización de Confluence para permitir Amazon Kendrapara conectarte a tu instancia de Confluence. Si su token de acceso caduca, puede usar el token de actualización para regenerar su token de acceso y actualizar el par de tokens. O bien, puede repetir el proceso de autorización. Para obtener más información sobre los tokens de acceso, consulte [Administra los tokens de acceso de OAuth](#).

- (Solo para Confluence Server)Opcional:Se configuró un token de acceso personal (PAT) que contiene un token de Confluence para permitirAmazon Kendrapara conectarte a tu instancia de Confluence. Para obtener información sobre cómo crear un token PAT, consulte[Uso de tokens de acceso personal](#).

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Has guardado tus credenciales de autenticación de Confluence en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Confluence aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tu fuente de datos de Confluence, debes proporcionar detalles de tus credenciales de Confluence para queAmazon Kendrapuede acceder a sus datos. Si aún no has configurado Confluence paraAmazon Kendraver[Requisitos previos \(p. 339\)](#).

Console

Para conectarAmazon Kendraa Confluence

1. Inicia sesión en elAWSConsola de administración y abra el[Amazon Kendraconsola](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConejor Confluence V2.0y, a continuación, elijaAñadir conector.
5. En elEspecificiar los detalles de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.

- d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:
- a. EnFuente, elige entreNube de ConfluenceyServidor Confluencesegún tu método de alojamiento de fuentes de datos de Confluence.
 - b. URL de Confluence—Introduzca la URL del host de Confluence. El formato de la URL de host que introduzca es`https://example.confluence.com`.
 - c. (Solo para Confluence Server)Ubicación del certificado SSL -opcional—Introduzca elAmazon S3ruta a tu archivo de certificado SSL para Confluence Server.
 - d. (Solo para Confluence Server)Proxy web -opcional—Introduzca el proxy webNombre del host(sin elhttp://ohttps://) yNúmero de puerto(puerto utilizado por el protocolo de transporte de URL del host). El número de puerto debe ser un valor numérico comprendido entre 0 y 65535.
 - e. (Solo para Confluence Server)Autorización—Elija habilitarLista de control de acceso (ACL). Luego, elige entreNombre de usuarioyCorreo electrónicopara seleccionar el campo que desea utilizar para el control de acceso.
 - f. Elige entreAutenticación básica,Autenticación OAuth 2.0y (solo para el servidor Confluence)Autenticación mediante token de acceso personalsegún su caso de uso.
 - g. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar tus credenciales de autenticación de Confluence. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta. Introduzca la siguiente información en la ventana:
 - i. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Confluence-' se añade automáticamente a tu nombre secreto.
 - ii. Si usaAutenticación básica—Introduzca elNombre secreto Nombre de usuario, yContraseña(contraseña del servidor de Confluence) que generaste y descargaste desde tu cuenta de Confluence.

Si usaAutenticación OAuth2.0—Introduzca elNombre secreto,Clave de aplicación,Secreto de la aplicación,Token de acceso, yToken de actualizaciónque creaste en tu cuenta de Confluence.

(Solo para el servidor Confluence) Si se utilizaAutenticación mediante token de acceso personal—Introduzca elNombre secretoyToken de Confluenceque creaste en tu cuenta de Confluence.
 - iii. EscojaGuardar y añadir secreto.
 - h. EnConfigurar VPC y grupo de seguridad (opcional), paraNube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesyGrupos de seguridad de VPC.
 - i. Rastreador de identidades—Elija activarAmazon Kendrarastreador de identidad para sincronizar la información de identidad. Si decide desactivar el rastreador de identidades, debe cargar la información principal mediante el[PutPrincipalMapping API](#).
 - j. IAMpapel—Elige uno existenteIAMrol o crea uno nuevoIAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- k. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:

- a. En Alcance de sincronización, para sincronizar contenido, elija sincronizar entre los siguientes tipos de entidades: Páginas, Comentarios de página, Adjuntos de página, Blogs, Comentarios del blog, Archivos adjuntos del blog, Espacios personales, y Espacios archivados.

Note

Comentarios de páginas y Adjuntos de páginas solo se puede seleccionar si eliges sincronizar Páginas. Comentarios del blog y archivos adjuntos del blog solo se puede seleccionar si eliges sincronizar Blogs.

- b. En Configuración adicional por Espacios: patrones de expresiones regulares, especifique si desea incluir o excluir espacios específicos en el índice mediante:
 - Tecla de espacio—Por ejemplo, *mi-espacio-123*.
 - URL—Por ejemplo, **/MySite/MyDocuments/*.
 - Tipo de archivo—Por ejemplo, *.*\ .pdf, .*\ .txt*.
 - c. Para Modo de sincronización elija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Cuando sincronizas tu fuente de datos con Amazon Kendra por primera vez, todo el contenido se sincroniza de forma predeterminada.
 - Sincronización completa—Sincronice todo el contenido independientemente del estado de sincronización anterior.
 - Sincronización de contenido nuevo, modificado o eliminado—Sincronice únicamente contenido nuevo, modificado y eliminado.
 - d. En Cronograma de ejecución de sincronización, para Frecuencia—Con qué frecuencia Amazon Kendra se sincronizará con su fuente de datos.
 - e. Elija Siguiente.
8. En el Definir mapeos de campos de página, introduzca la siguiente información:
 - a. Para Espacio, página, Blog, Comentario y Adjuntivo—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
 9. En el Revisar y crear página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en Fuentes de datos de página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Confluence

Debe especificar un JSON del [esquema de fuente de datos](#) utilizando el [TemplateConfiguration API](#).
Debe proporcionar la siguiente información:

- Fuente de datos—Especifique la fuente de datos como CONFLUENCEV2.
- URL del servidor—Especifica la versión de la instancia host de Confluence. Por ejemplo, <https://example.confluence.com>.

- (Opcional: solo para Confluence Server) Ubicación del certificado SSL—Especifique elS3bucketNamey s3certificateNameutilizaste para almacenar tu certificado SSL.
- Tipo de autenticación—Especifique el tipo de autenticación, siBasic, OAuth2, Personal-tokenpara tu instancia de Confluence.
- Teclea—EspecificarTEMPLATEcomo el tipo cuando llamasCreateDataSource.
- Modo de sincronización—Especifique siAmazon Kendradebe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:
 - FORCED_FULL_CRAWLpara volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice
 - FULL_CRAWLrastrea de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que tu fuente de datos se sincronice con tu índice
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación que creaste en tu cuenta de Confluence. Si utiliza la autenticación de cuenta básica, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "Confluence account user name",  
    "password": "Confluence API token"  
}
```

Si usas la autenticación OAuth 2.0, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "confluenceAppKey": "app key for your Confluence account",  
    "confluenceAppSecret": "app secret from your Confluence token",  
    "confluenceAccessToken": "access token created in Confluence",  
    "confluenceRefreshToken": "refresh token created in Confluence"  
}
```

(Solo para Confluence Server) Si usas la autenticación básica, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "hostUrl": "Confluence Server host URL",  
    "username": "Confluence Server user name",  
    "password": "Confluence Server password"  
}
```

(Solo para Confluence Server) Si usas la autenticación con token de acceso personal, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "hostUrl": "Confluence Server host URL",  
    "patToken": "Confluence token"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSource para proporcionar un IAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas necesarias para el conector de Confluence y Amazon Kendra. Para obtener más información, consulte [IAMfunciones para las fuentes de datos de Confluence](#).

También puede añadir las siguientes funciones opcionales:

- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir ciertos espacios, páginas, blogs y sus comentarios y archivos adjuntos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Habilitar el rastreador de identidades: especifique si desea activar el rastreador de identidades. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el [elPutPrincipalMapping](#) API. Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado en contexto de usuario](#).
- Nube privada virtual (VPC)—EspecificarVpcConfiguration cuando llamasCreateDataSource. Para obtener más información, consulte [ConfigurandoAmazon Kendrautilizar unAmazon VPC \(p. 465\)](#).
- Mapeos de campo—Elige asignar los campos de tu fuente de datos de Confluence a tuAmazon Kendracampos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Confluence](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [Esquema de plantillas de Confluence \(p. 165\)](#).

Notas

- El token de acceso personal (PAT) no está disponible para Confluence Cloud.

Conección de fuente de datos personalizado

Utilice una fuente de datos personalizada cuando tenga un repositorio que Amazon Kendra aún no proporciona un conector de fuente de datos para. Puede usarlo para ver las mismas métricas del historial de ejecución que Amazon Kendra las fuentes de datos proporcionan incluso cuando no puede utilizarlas Amazon Kendra de fuentes de datos para sincronizar tus repositorios. Utilice esto para crear una experiencia de monitorización de sincronización coherente entre Amazon Kendra fuentes de datos y fuentes personalizadas. Concretamente, utilice una fuente de datos personalizada para ver las métricas de sincronización de un conector de fuente de datos que haya creado mediante el [BatchPutDocument](#) y [BatchDeleteDocument](#) APIs.

Para solucionar problemas con su conector de fuente de datos personalizado de Amazon Kendra, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Al crear una fuente de datos personalizada, tiene un control total sobre cómo se seleccionan los documentos que se van a indexar. Amazon Kendra solo proporciona información métrica que puede utilizar para supervisar los trabajos de sincronización de fuentes de datos. Debe crear y ejecutar el rastreador que determina los documentos que indexa su fuente de datos.

Debe especificar el título principal de sus documentos mediante la [Documento](#) objeto, `y_source_uri` en [DocumentAttribute](#) con el fin de tener `DocumentTitle` y `DocumentURI` incluido en la respuesta de la `Query` resultado.

Crea un identificador para la fuente de datos personalizada mediante la consola o mediante la [CreateDataSource](#) API. Para usar la consola, asigne un nombre a la fuente de datos y, si lo desea, una descripción y etiquetas de recursos. Una vez creada la fuente de datos, se muestra un identificador de fuente de datos. Copie este identificador para usarlo cuando sincronice la fuente de datos con el índice.

Specify data source details

The screenshot shows the 'Specify data source details' step of a wizard. It has two main sections: 'Name data source' and 'Tags (0) - optional'.
Name data source:
- Data source name: A text input field containing 'my-data-source'. Below it is a note: 'Maximum of 1000 alphanumeric characters. Can include hyphens (-), but not spaces.'
- Description - optional: An empty text area.
Tags (0) - optional
- A note: 'A tag is an administrative label that you assign to AWS resources to make it easier to manage them. Each tag consists of a key and an optional value. Use tags to search and filter your resources or track your AWS costs.'
- A note below: 'This resource has no tags'
- A button: 'Add new tag'
- A note: 'You can add up to 50 more tags.'
At the bottom right are 'Cancel' and 'Next' buttons.

También puede crear una fuente de datos personalizada mediante `CreateDataSource` API. La API devuelve un ID para usarlo al sincronizar la fuente de datos. Cuando usa el `CreateDataSource` API para crear una fuente de datos personalizada, no puede configurar la `Configuration`, `RoleArn` o `Schedule` parámetros. Si configuras estos parámetros, Amazon Kendra devuelve un `ValidationException` excepción.

Para utilizar una fuente de datos personalizada, cree una aplicación que se encargue de actualizar la Amazon Kendra índice. La aplicación depende de un rastreador que usted cree. El rastreador lee los documentos de tu repositorio y determina cuáles se deben enviar a Amazon Kendra. La aplicación debe realizar los siguientes pasos:

1. Rastrea tu repositorio y haz una lista de los documentos del repositorio que se han añadido, actualizado o eliminado.
2. Llame al[StartDataSourceSyncJob](#)API para indicar que se está iniciando un trabajo de sincronización. Debe proporcionar un identificador de fuente de datos para identificar la fuente de datos que se está sincronizando. Amazon Kendra devuelve un identificador de ejecución para identificar un trabajo de sincronización concreto.
3. Llame al[BatchDeleteDocument](#)API para eliminar documentos del índice. Debe proporcionar el identificador de la fuente de datos y el identificador de ejecución para identificar la fuente de datos que se está sincronizando y el trabajo al que está asociada esta actualización.
4. Llame al[StopDataSourceSyncJob](#)API para indicar el final del trabajo de sincronización. Después de llamar al[StopDataSourceSyncJob](#)API, el ID de ejecución asociado ya no es válido.
5. Llame al[ListDataSourceSyncJobs](#)API con los identificadores de índice y fuente de datos para enumerar los trabajos de sincronización de la fuente de datos y ver las métricas de los trabajos de sincronización.

Después de finalizar un trabajo de sincronización, puede iniciar un nuevo trabajo de sincronización. Puede transcurrir un período de tiempo antes de que todos los documentos enviados se agreguen al índice. Utilice el[ListDataSourceSyncJobs](#)API para ver el estado del trabajo de sincronización. Si elStatusdevuelto para el trabajo de sincronización esSYNCING_INDEXING, algunos documentos aún se están indexando. Puede iniciar un nuevo trabajo de sincronización cuando el estado del trabajo anterior seaFAILED,SUCCEEDED, oSYNCING_INDEX.

Después de llamar al[StopDataSourceSyncJob](#)API, no puedes usar un identificador de trabajo de sincronización en una llamada al[BatchPutDocument](#)o[BatchDeleteDocument](#)APIs. Si lo hace, todos los documentos presentados se devolverán en elFailedDocumentsmensaje de respuesta de la API.

Atributos obligatorios.

Al enviar un documento aAmazon Kendrautilizando el[BatchPutDocument](#)API, cada documento requiere dos atributos para identificar la fuente de datos y la ejecución de sincronización a la que pertenece. Debe proporcionar los dos atributos siguientes:

- _data_source_id—El identificador de la fuente de datos. Se devuelve al crear la fuente de datos con la consola o el[CreateDataSource](#)API.
- _data_source_sync_job_execution_id: el identificador de la ejecución de sincronización. Se devuelve cuando se inicia la sincronización del índice con[StartDataSourceSyncJob](#)API.

El siguiente es el JSON necesario para indexar un documento mediante una fuente de datos personalizada.

```
{  
    "Documents": [  
        {  
            "Attributes": [  
                {  
                    "Key": "_data_source_id",  
                    "Value": {  
                        "StringValue": "data source identifier"  
                    }  
                },  
                {  
                    "Key": "_data_source_sync_job_execution_id",  
                    "Value": {  
                        "StringValue": "sync job identifier"  
                    }  
                }  
            ],  
        }  
    ]  
}
```

```
        "Blob": "document content",
        "ContentType": "content type",
        "Id": "document identifier",
        "Title": "document title"
    }
],
"IndexId": "index identifier",
"RoleArn": "IAM role ARN"
}
```

Al eliminar un documento del índice mediante elBatchDeleteDocumentAPI, debe especificar los dos campos siguientes enDataSourceSyncJobMetricTargetparámetro:

- **DataSourceId**—El identificador de la fuente de datos. Se devuelve al crear la fuente de datos con la consola o elCreateDataSourceAPI.
- **DataSourceSyncJobId**: el identificador de la ejecución de sincronización. Se devuelve cuando se inicia la sincronización del índice conStartDataSourceSyncJobAPI.

El siguiente es el JSON necesario para eliminar un documento del índice mediante elBatchDeleteDocumentAPI.

```
{
    "DataSourceSyncJobMetricTarget": {
        "DataSourceId": "data source identifier",
        "DataSourceSyncJobId": "sync job identifier"
    },
    "DocumentIdList": [
        "document identifier"
    ],
    "IndexId": "index identifier"
}
```

Visualización de métricas de

Una vez finalizado un trabajo de sincronización, puede utilizar el[DataSourceSyncJobMetrics](#)API para obtener las métricas asociadas al trabajo de sincronización. Úselo para supervisar las sincronizaciones de sus fuentes de datos personalizadas.

Si envía el mismo documento varias veces, ya sea como parte delBatchPutDocumentAPI, laBatchDeleteDocumentAPI, o si el documento se envía para agregarlo o eliminarlo, el documento solo se cuenta una vez en las métricas.

- **DocumentsAdded**—El número de documentos enviados mediante elBatchPutDocumentLa API asociada a este trabajo de sincronización se agregó al índice por primera vez. Si un documento se envía para su adición más de una vez en una sincronización, el documento solo se cuenta una vez en las métricas.
- **DocumentsDeleted**—El número de documentos enviados mediante elBatchDeleteDocumentLa API asociada a este trabajo de sincronización se ha eliminado del índice. Si un documento se envía para su eliminación más de una vez en una sincronización, el documento solo se cuenta una vez en las métricas.
- **DocumentsFailed**: el número de documentos asociados a este trabajo de sincronización que no se pudieron indexar. Estos son documentos que fueron aceptados porAmazon Kendrapara la indexación, pero no se pudo indexar ni eliminar. Si un documento no es aceptado porAmazon Kendra, el identificador del documento se devuelve enFailedDocumentspropiedad de respuesta delBatchPutDocumentyBatchDeleteDocumentAPIs.
- **DocumentsModified**—El número de documentos modificados enviados mediante elBatchPutDocumentAPI asociada a este trabajo de sincronización que se modificó enAmazon Kendraíndice.

Amazon Kendra también emite Amazon CloudWatch métricas al indexar documentos. Para obtener más información, consulte [Monitorización Amazon Kendra con Amazon CloudWatch](#).

Amazon Kendra devuelve el `DocumentsScanned` métrica para fuentes de datos personalizadas. También emite el `CloudWatch` métricas que figuran en el documento [Métricas para Amazon Kendra fuentes de datos](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos personalizada, consulte:

- [Agregar fuentes de datos personalizadas a Amazon Kendra](#)

Fuente de datos personalizada (Java)

El código siguiente proporciona un ejemplo de implementación de una fuente de datos personalizada mediante Java. El programa primero crea una fuente de datos personalizada y, a continuación, sincroniza los documentos recién agregados al índice con la fuente de datos personalizada.

El código siguiente muestra la creación y el uso de una fuente de datos personalizada. Al utilizar una fuente de datos personalizada en la aplicación, no es necesario crear una nueva fuente de datos (proceso único) cada vez que sincronice el índice con la fuente de datos. Utiliza el ID de índice y el ID de la fuente de datos para sincronizar los datos.

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import csoftware.amazon.awssdk.services.kendra.model.BatchPutDocumentRequest;
import csoftware.amazon.awssdk.services.kendra.model.BatchPutDocumentResponse;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.Document;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;
import software.amazon.awssdk.services.kendra.model.StopDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StopDataSourceSyncJobResponse;

public class SampleSyncForCustomDataSource {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String myIndexId = "yourIndexId";
        String dataSourceName = "custom data source";
        String dataSourceDescription = "Amazon Kendra custom data source connector"

        // Create custom data source
        CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
            .builder()
            .indexId(myIndexId)
            .name(dataSourceName)
            .description(dataSourceDescription)
            .type(DataSourceType.CUSTOM)
            .build();

        CreateDataSourceResponse createDataSourceResponse =
            kendra.createDataSource(createDataSourceRequest);
```

```
System.out.println(String.format("Response of creating data source: %s",
createDataSourceResponse);

// Get the data source ID from createDataSourceResponse
String dataSourceId = createDataSourceResponse.Id();

// Wait for the custom data source to become active
System.out.println(String.format("Waiting for Amazon Kendra to create the data source
%s", dataSourceId));
// You can use the DescribeDataSource API to check the status
DescribeDataSourceRequest describeDataSourceRequest = DescribeDataSourceRequest
    .builder()
    .indexId(myIndexId)
    .id(dataSourceId)
    .build();

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

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

    TimeUnit.SECONDS.sleep(60);
}

// Start syncing your 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();

// 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));

// Wait for the sync job status to succeed
// If the sync job status is SYNCING_INDEXING, documents are still being indexed
// If the sync job status is SYNCING, sync job has started
System.out.println(String.format("Waiting for the data source to sync with the index %s
for execution ID %s", indexId, startDataSourceSyncJobResponse.executionId()));
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("Syncing data source. Status: %s", job.status()));

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

    TimeUnit.SECONDS.sleep(60);
}

// Once custom data source synced, stop the sync job using the StopDataSourceSyncJob
API
StopDataSourceSyncJobResponse stopDataSourceSyncJobResponse =
kendra.stopDataSourceSyncJob(
    StopDataSourceSyncJobRequest()
        .indexId(myIndexId)
        .id(dataSourceId)
    );
}
```

Dropbox

Dropbox es un servicio de alojamiento de archivos que ofrece servicios de almacenamiento en la nube, organización de documentos y creación de plantillas de documentos. Si eres usuario de Dropbox, puedes usar Amazon Kendra para indexar tus archivos de Dropbox, Dropbox Paper, las plantillas de Dropbox Paper y los accesos directos a páginas web almacenados. También puede configurar Amazon Kendra para indexar archivos específicos de Dropbox, Dropbox Paper, plantillas de Dropbox Paper y accesos directos almacenados a páginas web.

Amazon Kendra es compatible con Dropbox y Dropbox Advanced para Dropbox Business.

Puedes conectarte a Amazon Kendra como tu fuente de datos de Dropbox mediante la [Amazon Kendra console](#) o la [Template Configuration API](#).

Para solucionar problemas de Amazon Kendra Conector de fuentes de datos de Dropbox, consulta [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 351\)](#)
- [Requisitos previos \(p. 351\)](#)
- [Instrucciones de conexión \(p. 352\)](#)
- [Más información \(p. 355\)](#)

Características admitidas

Amazon Kendra El conector de fuentes de datos de Dropbox admite las siguientes funciones:

- Change log
- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de poder usar Amazon Kendra para indexar tu fuente de datos de Dropbox, realiza estos cambios en tu Dropbox y AWS cuentas.

En Dropbox, asegúrate de tener:

- Creé una cuenta de Dropbox Advanced y configuré un usuario administrador.
- Creé una aplicación de Dropbox con un Nombre de la aplicación, activado Acceso limitado. Consulte [Documentación de Dropbox sobre la creación de una aplicación](#).
- Activado Dropbox completó permisos en la consola de Dropbox y se agregaron los siguientes permisos:
 - archivos.content.read
 - archivos.metadata.read
 - compartir.leer
 - file_requests.read
 - grupos.leer
 - team_info.read
 - team_data.content.read
- Apuntaste la clave de tu aplicación de Dropbox, el secreto de la aplicación de Dropbox y el token de acceso a Dropbox para las credenciales de autenticación básicas.
- Generaste y copiaste un token de acceso temporal de OAuth 2.0 para tu aplicación de Dropbox. Este token es temporal y caduca a las 4 horas. Consulte [Documentación de Dropbox sobre la autenticación con OAuth](#).

Note

Se recomienda crear un token de acceso de actualización de Dropbox que no caduque nunca, en lugar de utilizar un token de acceso único que caduque a las 4 horas. Un token de acceso a la actualización es permanente y nunca caduca, por lo que puedes seguir sincronizando tu fuente de datos en el futuro.

- Recomendado: Se configuró un token de actualización permanente de Dropbox que nunca caduca para permitir Amazon Kendra para seguir sincronizando la fuente de datos sin interrupciones. Consulte [Documentación de Dropbox sobre los tokens de actualización](#).
- Has comprobado que cada documento es único en Dropbox y en otras fuentes de datos que piensas usar para el mismo índice. Cada fuente de datos que deseas utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu cuenta de AWS, asegúrate de tener:

- [Creó un Amazon Kendra índice](#), si usa la API, anota el ID del índice.
- [Creó un IAM papel](#) para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al correcto AWS Secrets Manager identificación secreta.

- Has guardado tus credenciales de autenticación de Dropbox en un AWS Secrets Manager secreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es importante recomendar reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o secreto, puedes usar la consola para crear un nuevo IAM rol y Secrets Manager secreto cuando conectas tu fuente de datos de Dropbox a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existente IAM rol y Secrets Manager secreto y un identificador de índice.

Instrucciones de conexión

Para conectar Amazon Kendra a tu fuente de datos de Dropbox, debes proporcionar los detalles necesarios de tu fuente de datos de Dropbox para que Amazon Kendra pueda acceder a sus datos. Si aún no has configurado Dropbox para Amazon Kendra, consulta [Requisitos previos \(p. 351\)](#).

Console

Para conectar Amazon Kendra a Dropbox

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, selecciona Índices y, a continuación, elija el índice que deseas utilizar de la lista de índices.

Note

Puede elegir configurar o editar su control de acceso de usuarios ajustes en Ajustes de índice.

3. En el primer paso de la página, elige Agregar fuente de datos.
4. En el paso de agregar fuente de datos, elige Conector de Dropbox, a continuación, elija Agregar fuente de datos.
5. En el paso de especificar los detalles de la fuente de datos, introduzca la siguiente información:
 - a. En Nombre y descripción, para Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.

- b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paralidioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:
- a. Tipo de token de autenticación—Elige entreToken permanente (recomendado)yToken de acceso (uso temporal)según su caso de uso.
 - b. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar tus credenciales de autenticación de Dropbox. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - i. Introduzca la siguiente información enCrea unAWS Secrets Managerventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Dropbox-' se añade automáticamente a tu nombre secreto.
 - B. ParaClave de aplicación,Secreto de la aplicacióné información simbólica (permanente o temporal): introduce los valores de las credenciales de autenticación que generaste en tu cuenta de Dropbox.
 - ii. Seleccione Guardar.
 - c. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesyGrupos de seguridad de VPC.
 - d. IAMpapel—Elige uno existenteIAMrol o crea uno nuevolAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- e. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
- a. ParaSeleccione entidades o tipos de contenido—Elija las entidades o los tipos de contenido que deseé rastrear.
 - b. Cambiar modo de registro—Elija actualizar el índice en lugar de sincronizar todos los archivos.
 - c. EnConfiguración adicionalporPatrones Regex—Añada patrones de expresiones regulares para incluir o excluir determinados archivos.
 - d. EnCronograma de ejecución de sincronización, paraFrecuencia— Elige con qué frecuenciaAmazon Kendrase sincronizará con su fuente de datos.
 - e. Elija Siguiente.
8. En elDefinir mapeos de campospágina, introduzca la siguiente información:
- a. Archivos,Dropbox Paper, yPlantillas de Dropbox Paper—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.

9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon Kendraa Dropbox

Debe especificar un JSON delesquema de fuente de datosutilizando elTemplateConfigurationAPI.
Debe proporcionar la siguiente información:

- Fuente de datos—Debe especificar la fuente de datos comoDROPBOX.
- Teclea—EspecificartEMPLATEcomo el tipo cuando llamasCreateDataSource.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de tu cuenta de Dropbox. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "appKey": "Dropbox app key",  
    "appSecret": "Dropbox app secret",  
    "accesstoken": "temporary access token or refresh access token"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSourcepara proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para el conector de Dropbox yAmazon Kendra. Para obtener más información, consulteIAMfunciones para las fuentes de datos de Dropbox.

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfigurationcuando llamasCreateDataSource. Para obtener más información, consulte ConfigurandoAmazon Kendrautilizar unAmazon VPC (p. 465).
- Registro de cambios—Ya seaAmazon Kendradebe usar el mecanismo de registro de cambios en la fuente de datos de Dropbox para determinar si un documento debe agregarse, actualizarse o eliminarse del índice.

Note

Usa el registro de cambios si no quieresAmazon Kendrapara escanear todos los documentos. Si el registro de cambios es grande, podría tardarAmazon Kendramenos tiempo para escanear los documentos de la fuente de datos de Dropbox que para procesar el registro de cambios. Si sincronizas tu fuente de datos de Dropbox con tu índice por primera vez, se escanean todos los documentos.

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coinciden con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coinciden con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elige asignar los campos de tu fuente de datos de Dropbox a tu Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields \(Asignación de campos de origen de datos\)](#).
- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Dropbox](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [Esquema de plantillas de Dropbox \(p. 175\)](#).

Más información

Para obtener más información sobre la integración de Amazon Kendra con tu fuente de datos de Dropbox, consulta:

- [Indexa tu contenido de Dropbox mediante el conector de Dropbox para Amazon Kendra](#)

GitHub

GitHub es un servicio de alojamiento basado en la web para el desarrollo de software que proporciona servicios de almacenamiento y gestión de códigos con control de versiones. Puedes usar Amazon Kendra para indexar sus GitHub Nube empresarial (SaaS) y GitHub Archivos de repositorio de Enterprise Server (On Prem), solicitudes de emisión y extracción, comentarios de emisión y solicitud de extracción y archivos adjuntos de comentarios de emisión y solicitud de extracción. También puedes optar por incluir o excluir ciertos archivos.

Puedes conectarte a Amazon Kendra a través de GitHub como fuente de datos mediante la [Consola de Amazon Kendra](#) o la [API de configuración de GitHub](#).

Para solucionar problemas de Amazon Kendra GitHub conector de fuente de datos, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 355\)](#)
- [Requisitos previos \(p. 356\)](#)
- [Instrucciones de conexión \(p. 357\)](#)
- [Más información \(p. 360\)](#)

Características admitidas

Amazon Kendra GitHub conector de fuente de datos admite las siguientes funciones:

- Change log
- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder usar Amazon Kendra para indexar sus GitHub fuentes de datos, realice estos cambios en su GitHub AWS cuentas.

En GitHub, asegúrate de tener:

- Creó un GitHub usuario con permisos administrativos para el GitHub organización.
- Creó un token de acceso personal para las credenciales de autenticación.
Consulte [GitHub documentación sobre la creación de un token de acceso personal](#).
- Recomendado: Creó un token de OAuth para las credenciales de autenticación. Utilice el token OAuth para mejorar los límites de aceleración de la API y el rendimiento del conector.
Consulte [GitHub documentación sobre la autorización de OAuth](#).
- Opcional: Se instaló un certificado SSL.
- Tomó nota de la GitHub URL de host para el tipo de GitHub servicio que utilizas. Por ejemplo, la URL del servidor de GitHub la nube podría ser <https://api.github.com> la URL del servidor de GitHub el servidor podría ser <https://on-prem-host-url/api/v3/>.
- Tomó nota de la GitHub nombre de organización para sus repositorios desde su GitHub ajustes.
- Se agregaron los siguientes permisos:

Para GitHub Nube empresarial (SaaS)

- repositorio:estado
- repositorio público
- repositorio:invitar
- leer:org
- usuario:correo electrónico
- leer:usuario

Para GitHub Servidor empresarial (local)

- repositorio:estado
- repositorio público
- repositorio:invitar
- leer: org
- usuario:correo electrónico
- leer:usuario
- administrador del sitio
- Se ha marcado que cada documento es único en GitHub y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu Cuenta de AWS, asegúrate de tener:

- [Creó un Amazon Kendra índice](#), si usa la API, anota el ID del índice.

- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tuIAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Almacenó suGitHubcredenciales de autenticación en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tuGitHubfuente de datos paraAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tuGitHubfuente de datos, debe proporcionar los detalles necesarios de suGitHubfuente de datos para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configuradoGitHubporAmazon Kendra, consulte[Requisitos previos \(p. 356\)](#).

Console

Para conectarAmazon KendraaGitHub

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAgregar fuente de datos.
4. En elAgregar fuente de datospágina, eligeGitHubconectory, a continuación, elijaAgregar fuente de datos.
5. En elEspecificardetalless de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. Enidioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:
 - a. GitHubfuente—Elige entreGitHubNube empresarialyGitHubServidor empresarial.
 - b. GitHubURL del servidor—Introduzca suGitHubnombre de host.

- c. GitHub nombre de la organización—Introduzca su GitHub nombre de la organización. Puede encontrar la información de su organización en GitHub cuenta.
- d. AWS Secrets Manager secreto—Elige un secreto existente o crea uno nuevo Secrets Manager secreto para guardar tus GitHub credenciales de autenticación. Si eliges crear un nuevo secreto, un AWS Secrets Manager abre una ventana secreta.
 - i. Introduzca la siguiente información en Crea un AWS Secrets Manager ventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-GitHub-' se añade automáticamente a tu nombre secreto.
 - B. Para GitHub simbólico—Introduzca los valores de las credenciales de autenticación que creó en usted GitHub cuenta.
 - ii. Seleccione Guardar.
- e. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadir Subredes y Grupos de seguridad de VPC.
- f. IAM papel—Elige uno existente IAM rol o crea uno nuevo IAM rol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

- g. Elija Siguiente.
7. En el Configurar los ajustes de sincronización página, introduzca la siguiente información:
 - a. Seleccione los repositorios que desee rastrear—La GitHub entidades o tipos de contenido que desee rastrear.
 - b. Registro de cambios—Seleccione actualizar el índice en lugar de sincronizar todos los archivos.
 - c. Tipos de contenido—Seleccione los tipos de archivos que desee incluir.
 - d. Patrones Regex—Patrones de expresiones regulares para incluir o excluir ciertos archivos. Puede añadir hasta 100 patrones.
 - e. En Cronograma de ejecución de sincronización, para Frecuencia—Con qué frecuencia Amazon Kendra se sincronizará con su fuente de datos.
 - f. Elija Siguiente.
8. En el Definir mapeos de campos página, introduzca la siguiente información:
 - a. Para Repositorio, Confirmación de repositorio, Documento de emisión, Emitir comentario, Emitir archivo adjunto, Comentario de solicitud de extracción, Documento de solicitud de extracción, Anexo de solicitud de extracción—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En el Revisar y crear página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en Fuentes de datos página después de que se haya agregado correctamente.

Debe especificar lo siguiente mediante el [GitHubConfiguration](#) Objeto de API:

- Tipo de fuente de datos: especifique el tipo de fuente de datos como:SAASoON_PREMISE.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de suGitHubcuenta. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "personalToken": "token"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSourcepara proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para elGitHubconector yAmazon Kendra. Para obtener más información, consulte[IAMfunciones paraGitHubfuentes de datos](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfigurationcomo parte de la configuración de la fuente de datos. Consulte[ConfiguraciónAmazon Kendrausar una VPC](#).

Note

Si usaGitHubservidor, debe utilizar unAmazon VPCpara conectarse a suGitHubservidor.

- Registro de cambios—SiAmazon Kendradebe utilizar elGitHubmecanismo de registro de cambios en la fuente de datos para determinar si un documento debe agregarse, actualizarse o eliminarse en el índice.

Note

Usa el registro de cambios si no quieresAmazon Kendrapara escanear todos los documentos. Si el registro de cambios es grande, podría tardarAmazon Kendramenos tiempo para escanear los documentos delGitHubfuente de datos que procesar el registro de cambios. Si estás sincronizando tuGitHubfuente de datos con su índice por primera vez, se escanean todos los documentos.

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elja mapear suGitHubcampos de fuente de datos para suAmazon Kendracampos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendrarastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los

resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para GitHub fuentes de datos](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con tu GitHub fuente de datos, consulte:

- [Reimagine la búsqueda en GitHub repositorios con el poder de Amazon Kendra GitHub conector](#)

Gmail

Gmail es un cliente de correo desarrollado por Google a través del cual puedes enviar mensajes de correo electrónico con archivos adjuntos. Los mensajes de Gmail se pueden ordenar y almacenar en la bandeja de entrada del correo electrónico mediante carpetas y etiquetas. Puedes usar Amazon Kendra para indexar sus mensajes de correo electrónico y sus archivos adjuntos. También puede configurar Amazon Kendra para incluir o excluir mensajes de correo electrónico, archivos adjuntos de mensajes y etiquetas específicos para su indexación.

Puedes conectarte Amazon Kendra a tu fuente de datos de Gmail mediante el [Amazon Kendra console](#) y el [Template Configuration API](#).

Para solucionar problemas de su Amazon Kendra Conector de fuentes de datos de Gmail, consulta [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 360\)](#)
- [Requisitos previos \(p. 360\)](#)
- [Instrucciones de conexión \(p. 361\)](#)
- [Más información \(p. 365\)](#)
- [Notas \(p. 365\)](#)

Características admitidas

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados

Requisitos previos

Antes de poder usar Amazon Kendra para indexar tu fuente de datos de Gmail, realiza estos cambios en tu Gmail y AWS cuentas.

En Gmail, asegúrate de tener:

- Creé una cuenta de administrador de Google Cloud Platform y creé un proyecto de Google Cloud.
- Has activado la API de Gmail y la API del SDK de administración en tu cuenta de administrador.
- Creaste una cuenta de servicio y descargaste una clave privada JSON para tu Gmail. Para obtener información sobre cómo crear tu clave privada y acceder a ella, consulta la documentación de Google Cloud sobre cómo [Crear una clave de cuenta de servicio](#) y [Credenciales de cuenta de servicio](#).

- Copié el correo electrónico de tu cuenta de administrador, el correo electrónico de tu cuenta de servicio y tu clave privada para usarlos en la autenticación.
- Se agregaron los siguientes ámbitos de OAuth (con una función de administrador) para tu usuario y los directorios compartidos que deseas indexar:
 - https://www.googleapis.com/auth/admin.directory.user.readonly
 - https://www.googleapis.com/auth/gmail.readonly
- Has comprobado que cada documento es único en Gmail y en las demás fuentes de datos que piensas usar para el mismo índice. Cada fuente de datos que deseas utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellIAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulIAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Has guardado tus credenciales de autenticación de Gmail en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Gmail aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tu fuente de datos de Gmail debes proporcionar detalles de tus credenciales de Gmail para queAmazon Kendrapuede acceder a sus datos. Si aún no has configurado Gmail paraAmazon Kendra, consulte[Requisitos previos \(p. 360\)](#).

Console

Para conectarAmazon Kendraa Gmail

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseas utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConektor de Gmaily, a continuación, elijaAregar fuente de datos.

5. En el Especificar los detalles de la fuente de datos página, introduzca la siguiente información:
 - a. En Nombre y descripción, para Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional) Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, para idioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, para Añadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tus AWS costos.
 - e. Elija Siguiente.
6. En el Definir el acceso y la seguridad página, introduzca la siguiente información:
 - a. En Autenticación por AWS Secrets Manager secreto—Elige un secreto existente o crea uno nuevo Secrets Manager secreto para almacenar tus credenciales de autenticación de Gmail. Si eliges crear un nuevo secreto, un AWS Secrets Manager abre una ventana secreta.
 - Introduzca la siguiente información en el Crea un AWS Secrets Manager ventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto.
 - B. Correo electrónico del cliente—El correo electrónico del cliente que copiaste de tu cuenta de servicio de Google.
 - C. Correo electrónico de la cuenta de administrador—El correo electrónico de la cuenta de administrador que desea utilizar.
 - D. Clave privada—La clave privada que copiaste de tu cuenta de servicio de Google.
 - E. Seleccione Guardar.
 - b. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadir Subredes y Grupos de seguridad de VPC.
 - c. IAM papel—Elige uno existente IAM rol o crea uno nuevo IAM rol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAM roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

- d. Elija Siguiente.
7. En el Configurar los ajustes de sincronización página, introduzca la siguiente información:
 - a. En Alcance de sincronización, para Tipos de entidades—Seleccione Adjuntos de mensajes para sincronizar los archivos adjuntos de los mensajes. Los mensajes se sincronizarán de forma predeterminada.
 - b. (Opcional) Para Configuración adicional, introduzca la siguiente información:
 - i. Intervalo de fechas—Introduzca un intervalo de fechas para especificar la fecha de inicio y finalización de los correos electrónicos que se van a rastrear.
 - ii. Dominios de correo electrónico—Incluya o excluya correos electrónicos basados en dominios.
 - iii. Palabras clave en las asignaturas—Incluya o excluya correos electrónicos basados en palabras clave en sus asuntos.

Note

También puede optar por incluir cualquier documento que coincida con todas las palabras clave del asunto que ha introducido.

- iv. Etiquetas—Añada patrones de expresiones regulares para incluir o excluir etiquetas específicas. Puede añadir hasta 100 patrones.
 - v. Adjuntos—Añada patrones de expresiones regulares para incluir o excluir adjuntos específicos. Puede añadir hasta 100 patrones.
- c. ParaModo de sincronizaciónelija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Cuando sincronizas tu fuente de datos conAmazon Kendrapor primera vez, todo el contenido se sincroniza de forma predeterminada.
 - Sincronización completa—Sincronice todo el contenido independientemente del estado de sincronización anterior.
 - Sincronización de contenido nuevo, modificado o eliminado—Sincronice únicamente contenido nuevo, modificado y eliminado.

Important

Como no existe una API para actualizar los mensajes de Gmail eliminados permanentemente, unSincronización de contenido nuevo, modificado o eliminado:

- No eliminará de tu cuenta los mensajes que se eliminaron permanentemente de GmailAmazon Kendraíndice
- No sincronizará los cambios en las etiquetas de correo electrónico de Gmail Para sincronizar tu fuente de datos de Gmail, los cambios en las etiquetas y los mensajes de correo electrónico eliminados permanentemente con tuAmazon Kendraíndice, debe ejecutar rastreos completos periódicamente.

- d. EnCronograma de ejecución de sincronización, paraFrecuencia— Con qué frecuenciaAmazon Kendrase sincronizará con su fuente de datos.
 - e. Elija Siguiente.
8. En elDefinir mapeos de campospágina, introduzca la siguiente información:
 - a. ParaMensajesyAdjuntos de mensajes—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.

Note

Amazon KendraEl conector de fuente de datos de Gmail no admite la creación de campos de índice personalizados debido a las limitaciones de la API.

- b. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAgregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en elFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon Kendraa Gmail

Debe especificar lo siguiente mediante el[TemplateConfigurationAPI](#):

- Fuente de datos—Debe especificar la fuente de datos comoGMAIL.
- Esquema de fuente de datos—Incluya un JSON que contenga el esquema de la fuente de datos. Para ver el esquema de la plantilla, consulte[Esquemas de fuentes de datos](#).
- Teclea—EspecificartEMPLATEcomo el tipo cuando llamasCreateDataSource.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de tu cuenta de Gmail. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "adminAccountEmailId": "service account email",  
    "clientEmailId": "user account email",  
    "privateKey": "private key"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSource para proporcionar un IAMrol con permisos para acceder a su Secrets Managersecreto y para llamar a las API públicas requeridas para el conector de Gmail y Amazon Kendra. Para obtener más información, consulte [IAMfunciones para fuentes de datos de Gmail](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfiguration cuando llamasCreateDataSource. Para obtener más información, consulte [Configurando Amazon Kendra utilizar un Amazon VPC \(p. 465\)](#).
- Modo de sincronización—Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:
 - FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice
 - FULL_CRAWL rastrea de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que tu fuente de datos se sincronice con tu índice

Important

Como no existe una API para actualizar los mensajes de Gmail eliminados permanentemente, un FULL_CRAWL/Sincronización de contenido nuevo, modificado o eliminado:

- No eliminará de tu cuenta los mensajes que se eliminaron permanentemente de Gmail Amazon Kendra índice
- No sincronizará los cambios en las etiquetas de correo electrónico de Gmail Para sincronizar tu fuente de datos de Gmail, los cambios en las etiquetas y los mensajes de correo electrónico eliminados permanentemente con tu Amazon Kendra índice, debe ejecutar rastreos completos periódicamente.
- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir mensajes y archivos adjuntos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elige asignar los campos de la fuente de datos de Gmail a tu Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields \(Asignación de campos de origen de datos\)](#).

Note

Amazon KendraEl conector de fuente de datos de Gmail no admite la creación de campos de índice personalizados debido a las limitaciones de la API.

- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Gmail](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con tu fuente de datos de Gmail, consulta:

- [Realice una búsqueda inteligente en los correos electrónicos de su espacio de trabajo de Google mediante el conector de Gmail para Amazon Kendra.](#)

Notas

- Como no existe una API para actualizar los mensajes de Gmail eliminados permanentemente, un FULL_CRAWL/Sincronización de contenido nuevo, modificado o eliminado:
 - No eliminará de tu cuenta los mensajes que se eliminaron permanentemente de Gmail Amazon Kendra índice
 - No sincronizará los cambios en las etiquetas de correo electrónico de Gmail

Para sincronizar tu fuente de datos de Gmail, los cambios en las etiquetas y los mensajes de correo electrónico eliminados permanentemente con tu Amazon Kendra índice, debe ejecutar rastreos completos periódicamente.

- Amazon Kendra El conector de fuente de datos de Gmail no admite la creación de campos de índice personalizados debido a las limitaciones de la API.

Google Drive

Google Drive es un servicio de almacenamiento de archivos basado en la nube. Puedes usar Amazon Kendra para indexar los documentos almacenados en las carpetas de unidades compartidas, Mis unidades y Compartido conmigo en tu fuente de datos de Google Drive. Puede indexar tanto los documentos de Google Workspace como los documentos que figuran en [Tipos de documentación](#). También puede utilizar filtros de inclusión y exclusión para indexar el contenido por nombre de archivo, tipo de archivo y ruta de archivo.

Puedes conectarte Amazon Kendra a tu fuente de datos de Google Drive mediante la [Amazon Kendra consola](#), la [TemplateConfiguration API](#), o la [GoogleDriveConfiguration API](#).

Amazon Kendra tiene dos versiones del conector de Google Drive. Las funciones compatibles de cada versión incluyen:

Conector Google Drive V1.0/[GoogleDriveConfiguration API](#)

- Mapeos de campo
- Filtros de inclusión/exclusión

Conector Google Drive V2.0/[TemplateConfiguration API](#)

- Filtrado de contexto de usuario
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Filtros de inclusión/exclusión
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados

Note

Soporte para el conector Google Drive V1.0/GoogleDriveConfigurationEstá previsto que la API finalice en 2023. Recomendamos migrar al conector V2.0 de Google Drive o utilizarlo/ TemplateConfigurationAPI.

Para solucionar problemas de suAmazon KendraConektor de fuentes de datos de Google Drive, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Conector Google Drive V1.0 \(p. 366\)](#)
- [Conector Google Drive V2.0 \(p. 370\)](#)

Conektor Google Drive V1.0

Google Drive es un servicio de almacenamiento de archivos basado en la nube. Puedes usarAmazon Kendrapara indexar los documentos y comentarios almacenados en las carpetas de unidades compartidas, Mis unidades y Compartido conmigo en tu fuente de datos de Google Drive. Puede indexar los documentos de Google Workspace, así como los documentos que figuran en[Tipos de documentación](#). También puede utilizar filtros de inclusión y exclusión para indexar el contenido por nombre de archivo, tipo de archivo y ruta de archivo.

Note

Soporte para el conector Google Drive V1.0/GoogleDriveConfigurationEstá previsto que la API finalice en 2023. Recomendamos migrar al conector V2.0 de Google Drive o utilizarlo/ TemplateConfigurationAPI.

Para solucionar problemas de suAmazon KendraConektor de fuentes de datos de Google Drive, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 366\)](#)
- [Requisitos previos \(p. 366\)](#)
- [Instrucciones de conexión \(p. 367\)](#)
- [Más información \(p. 370\)](#)

Características admitidas

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder usarAmazon Kendrapara indexar tu fuente de datos de Google Drive, realiza estos cambios en tu Google Drive yAWS cuentas.

En Google Drive, asegúrate de tener:

- Ya se te ha otorgado acceso mediante un rol de superadministrador a un usuario con privilegios de administrador. No necesitas un rol de superadministrador para ti si un rol de superadministrador te ha concedido el acceso.
- Creó una cuenta de servicio conHabilitar la delegación de todo el dominio de G Suiteactivada y una clave JSON como clave privada utilizando la cuenta.
- Copió el correo electrónico de su cuenta de usuario y el correo electrónico de su cuenta de servicio. Cuando te conectas aAmazon Kendraintroduce el correo electrónico de su cuenta de usuario como correo electrónico de la cuenta de administrador y el correo de su cuenta de servicio como correo electrónico de cliente en suSecrets Managersecreto.
- Se agregaron la API del SDK de administración y la API de Google Drive a tu cuenta.
- Se agregaron (o se le pidió a un usuario con un rol de superadministrador que los agregara) los siguientes permisos a tu cuenta de servicio mediante un rol de superadministrador:
 - https://www.googleapis.com/auth/drive.readonly
 - https://www.googleapis.com/auth/drive.metadata.readonly
 - https://www.googleapis.com/auth/admin.directory.user.readonly
 - https://www.googleapis.com/auth/admin.directory.group.readonly
- Has comprobado que cada documento es único en Google Drive y en otras fuentes de datos que piensas utilizar para el mismo índice. Cada fuente de datos que deseas utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- Creó unAmazon Kendraíndicey, si usa la API, anota el ID del índice.
- Creó unIAMpapelpara su fuente de datos y, si usa la API, anotó el ARN delIAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tuIAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Has guardado tus credenciales de autenticación de Google Drive en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendable reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Google Drive aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tu fuente de datos de Google Drive, debes proporcionar los detalles necesarios de tu fuente de datos de Google Drive para queAmazon Kendrapuede acceder a sus datos. Si aún no has configurado Google Drive paraAmazon KendraverRequisitos previos (p. 366).

Console

Para conectarAmazon Kendraa Google Drive

1. Inicia sesión enAWSConsola de administración y abra elAmazon Kendraconsola.

2. En el panel de navegación de la izquierda, selecciona Índices y, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar su Control de acceso de usuarios ajustes en Ajustes de índice.

3. En el Primeros pasospágina, elige Agregar fuente de datos.
4. En el Agregar fuente de datospágina, elige Conector Google Drive V1.0y, a continuación, elija Añadir conector.
5. En el Especificar los detalles de la fuente de datospágina, introduzca la siguiente información:
 - a. En Nombre y descripción, para Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. (Opcional) Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, para idioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, para Añadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tus AWS costos.
 - e. Elija Siguiente.
6. En el Defina el acceso y la seguridadpágina, introduzca la siguiente información:
 - a. Para Tipo de autenticación—Elige entre Existente y Nuevo. Si eliges usar un secreto existente, usa Selecciona el secreto para elegir tu secreto.
 - b. Si eliges crear un nuevo secreto, un AWS Secrets Manager abre la opción secreta:
 - Introduzca la siguiente información en el Crea un AWS Secrets Manager ventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Google Drive-' se añade automáticamente a tu nombre secreto.
 - B. Para Correo electrónico de la cuenta de administrador, Correo electrónico del cliente, y Clave privada—Introduzca los valores de las credenciales de autenticación que generó y descargó de su cuenta de Google Drive.
 - C. Escoja Guardar autenticación.
 - c. IAM papel—Elige uno existente IAM rol o crea uno nuevo IAM rol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

- d. Elija Siguiente.
7. En el Configurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. Excluir cuentas de usuario—Los usuarios de Google Drive que quieras excluir del índice. Puedes añadir hasta 100 cuentas de usuario.
 - b. Excluir unidades compartidas—Las unidades compartidas de Google Drive que quieras excluir de tu índice. Puedes añadir hasta 100 unidades compartidas.
 - c. Excluir tipos de archivos (unidades)—Los tipos de archivos de Google Drive que quieras excluir del índice. También puedes optar por editar las selecciones de tipos MIME.
 - d. Configuraciones adicionales—Patrones de expresiones regulares para incluir o excluir cierto contenido. Puedes añadir hasta 100 patrones.

- e. Frecuencia— Con qué frecuencia Amazon Kendra se sincronizará con su fuente de datos.
 - f. Elija Siguiente.
8. En el **Establecer mapeos de campo** página, introduzca la siguiente información:
 - a. Para **Nombre de campo** y **Mapeos de campos sugeridos adicionales**—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. **Añadir campo**—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
 9. En el **Revisar y crear** página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione **Agregar fuente de datos**. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en **Fuentes de datos** página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Google Drive

Debe especificar lo siguiente mediante el [GoogleDriveConfiguration API](#):

- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contiene las credenciales de autenticación de tu cuenta de Google Drive. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "clientAccount": "service account email",  
    "adminAccount": "user account email",  
    "privateKey": "private key"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendado reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAM papel—Especificar `RoleArn` cuando llamas `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Google Drive y Amazon Kendra. Para obtener más información, consulte [IAM funciones para las fuentes de datos de Google Drive](#).

También puede añadir las siguientes funciones opcionales:

- Filtros de inclusión y exclusión—De forma predeterminada Amazon Kendra indexa todos los documentos de Google Drive. Puede especificar si desea incluir o excluir cierto contenido en las unidades compartidas, las cuentas de usuario, los tipos MIME de documentos y los archivos. Si eliges excluir las cuentas de usuario, no se indexa ninguno de los archivos de Mi Drive propiedad de la cuenta. Los archivos compartidos con el usuario se indexan a menos que también se excluya al propietario del archivo.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión,

solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coinciden con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coinciden con el filtro de exclusión no se indexarán, aunque coinciden con el filtro de inclusión.

- Mapeos de campo—Elija asignar los campos de la fuente de datos de Google Drive a su Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Google Drive](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con tu fuente de datos de Google Drive, consulta:

- [Empezar con el Amazon Kendra Conector Google Drive](#)

Conector Google Drive V2.0

Google Drive es un servicio de almacenamiento de archivos basado en la nube. Puedes usar Amazon Kendra para indexar los documentos y comentarios almacenados en las carpetas de unidades compartidas, Mis unidades y Compartido conmigo en tu fuente de datos de Google Drive. Puede indexar los documentos de Google Workspace, así como los documentos que figuran en [Tipos de documentación](#). También puede utilizar filtros de inclusión y exclusión para indexar el contenido por nombre de archivo, tipo de archivo y ruta de archivo.

Note

Soporte para el conector Google Drive V1.0/GoogleDriveConfigurationEstá previsto que la API finalice en 2023. Recomendamos migrar al conector V2.0 de Google Drive o utilizarlo/TemplateConfigurationAPI.

Para solucionar problemas de su Amazon Kendra Conector de fuentes de datos de Google Drive, consulta [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 370\)](#)
- [Requisitos previos \(p. 371\)](#)
- [Instrucciones de conexión \(p. 372\)](#)
- [Notas \(p. 376\)](#)

Características admitidas

- Mapeos de campo
- Filtrado de contexto de usuario
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder usar Amazon Kendra para indexar tu fuente de datos de Google Drive, realiza estos cambios en tu Google Drive y AWS cuentas.

En Google Drive, asegúrate de tener:

- Ya se te ha otorgado acceso mediante un rol de superadministrador o un usuario con privilegios de administrador. No necesitas un rol de superadministrador para ti si un rol de superadministrador te ha concedido el acceso.
- Creó una cuenta de servicio con [Habilitar la delegación de todo el dominio de G Suite](#) activó y generó una clave privada JSON con la cuenta.
- Se agregaron la API del SDK de administración y la API de Google Drive a tu cuenta de usuario.
- Credenciales de conexión configuradas a la cuenta de servicio de Google Drive que contienen el correo electrónico de la cuenta de administrador, el correo electrónico del cliente (correo electrónico de la cuenta de servicio) y la clave privada. Consulte [Documentación de Google Cloud sobre la creación y eliminación de claves de cuentas de servicio](#).
- Opcional: Se configuró un token de credenciales de OAuth 2.0 que puede identificar Amazon Kendra y genere un identificador de cliente de OAuth, un secreto de cliente y un token de actualización como credenciales de conexión. Consulte [Documentación de Google sobre el uso de OAuth 2.0 para acceder a las API](#).
- Se agregaron (o se le pidió a un usuario con un rol de superadministrador que añadiera) los siguientes ámbitos de OAuth a tu cuenta de servicio mediante un rol de superadministrador:
 - <https://www.googleapis.com/auth/drive>
 - <https://www.googleapis.com/auth/drive.file>
 - <https://www.googleapis.com/auth/drive.readonly>
 - <https://www.googleapis.com/auth/cloud-platform>
 - <https://www.googleapis.com/auth/forms.body.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>
 - <https://www.googleapis.com/auth/admin.directory.user.alias.readonly>
 - <https://www.googleapis.com/auth/admin.directory.userschema.readonly>
 - <https://www.googleapis.com/auth/admin.directory.group.member.readonly>
- Has comprobado que cada documento es único en Google Drive y en otras fuentes de datos que piensas utilizar para el mismo índice. Cada fuente de datos que deseas utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu cuenta de AWS, asegúrate de tener:

- [Creó un Amazon Kendra índice](#), si usa la API, anota el ID del índice.
- [Creó un IAM papel](#) para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al correcto AWS Secrets Manager identificación secreta.

- Has guardado tus credenciales de autenticación de Google Drive en un AWS Secrets Manager secreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendable reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAMrol o secreto, puedes usar la consola para crear un nuevo IAMrol y Secrets Manager secreto cuando conectas tu fuente de datos de Google Drive a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existente IAMrol y Secrets Manager secreto y un identificador de índice.

Instrucciones de conexión

Para conectar Amazon Kendra a tu fuente de datos de Google Drive, debes proporcionar los detalles necesarios de tu fuente de datos de Google Drive para que Amazon Kendra pueda acceder a sus datos. Si aún no has configurado Google Drive para Amazon Kendra vea [Requisitos previos \(p. 371\)](#).

Console

Para conectar Amazon Kendra a Google Drive

1. Inicia sesión en AWS Consola de administración y abra el [Amazon Kendra consola](#).
2. En el panel de navegación de la izquierda, selecciona Índices y, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar su Control de acceso de usuarios ajustes en Ajustes de índice.

3. En el Primeros pasos página, elige Agregar fuente de datos.
4. En el Agregar fuente de datos página, elige Conector Google Drive V2.0 y, a continuación, elija Añadir conector.
5. En el Especificar los detalles de la fuente de datos página, introduzca la siguiente información:
 - a. En Nombre y descripción, para Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. (Opcional) Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, para idioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, para Añadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tus AWS costos.
 - e. Elija Siguiente.
6. En el Defina el acceso y la seguridad página, introduzca la siguiente información:
 - a. Para Autorización—Elija habilitar Lista de control de acceso (ACL) según su caso de uso. La ACL está habilitada de forma predeterminada.
 - b. Para autenticación—Elige entre Cuenta de servicio de Google y Autenticación OAuth 2.0 según su caso de uso.
 - c. AWS Secrets Manager secreto—Elige un secreto existente o crea uno nuevo Secrets Manager secreto para almacenar tus credenciales de autenticación de Google Drive. Si eliges crear un nuevo secreto, un AWS Secrets Manager abre una ventana secreta.

- i. Si elegisteCuenta de servicio de Google, introduzca elNombre secreto,Correo electrónico de la cuenta de administrador,Correo electrónico del cliente, yClave privadaque creaste en tu cuenta de servicio y seleccionasteGuardar y añadir secreto.
- ii. Si elegisteAutenticación OAuth 2.0, introduzca los detalles deNombre secreto, ID de cliente, Secreto de cliente y Token de actualizaciónque creaste en tu cuenta de servicio y seleccionasteGuardar y añadir secreto.
- d. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesyGrupos de seguridad de VPC.
- e. (Solo para usuarios de autenticación de cuentas de servicio de Google)Rastreador de identidades—Elija activarAmazon Kendrarastreador de identidad para sincronizar la información de identidad. Si optas por desactivar el rastreador de identidades, debes cargar la información principal mediante el[PutPrincipalMapping API](#).
- f. IAMpapel—Elige uno existenteIAMrol o crea uno nuevoIAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- g. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
- a. Sincronizar contenido—Seleccione una de las siguientes opciones para indexarMi impulso y lo compartí conmigo,Unidades compartidas, yComentarios. Los archivos se rastrean de forma predeterminada.
 - b. EnConfiguración adicional: opcionalintroduzca la siguiente información opcional:
 - i. Tamaño máximo de archivo—Elija el límite máximo de tamaño en MB de los archivos que se van a rastrearAmazon Kendra.
 - ii. Correo electrónico del usuario—Añada los correos electrónicos de los usuarios que deseé incluir o excluir.
 - iii. Unidades compartidas—Añada los nombres de las unidades compartidas que deseé incluir o excluir.
 - iv. Tipos de mímica—Añada los tipos MIME que deseé incluir o excluir.
 - v. Patrones de expresiones regulares de adjunto—Añada patrones de expresiones regulares para incluir o excluir ciertos archivos adjuntos de todas las entidades compatibles. Puede añadir hasta 100 patrones.
 - c. ParaModo de sincronización, elija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Cuando sincronizas tu fuente de datos conAmazon Kendrapor primera vez, todo el contenido se sincroniza de forma predeterminada.
 - Sincronización completa—Sincronice todo el contenido independientemente del estado de sincronización anterior.
 - Sincronización de documentos nuevos o modificados—Sincronice solo los documentos nuevos y modificados.
 - Sincronización de documentos nuevos, modificados o eliminados—Sincronice solo los documentos nuevos, modificados y eliminados.

Important

La API de Google Drive no admite la recuperación de comentarios de un archivo eliminado permanentemente. Los comentarios de los archivos desechados

- se pueden recuperar. Cuando se elimina un archivo, el conector eliminará los comentarios del Amazon Kendraíndice.
- d. En Cronograma de ejecución de sincronización, para Frecuencia— Con qué frecuencia Amazon Kendra sincronizará con su fuente de datos.
 - e. Elija Siguiente.
8. En el Establecer mapeos de campos página, introduzca la siguiente información:
 - a. Para Archivos—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Para Comentarios—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.
- Note**
- La API de Google Drive no admite la creación de campos personalizados. El mapeo de campos personalizado no está disponible para el conector de Google Drive.
- c. Elija Siguiente.
 9. En el Revisar y crear página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en el Fuentes de datos página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Google Drive

Debe especificar un JSON de [esquema de fuente de datos](#) utilizando el [TemplateConfiguration API](#). Debe proporcionar la siguiente información:

- Fuente de datos—Especifique la fuente de datos como GOOGLEDRIVEV2.
- Téclea—Especificando TEMPLATE como el tipo cuando llamas CreateDataSource.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contiene las credenciales de autenticación que creaste en tu cuenta de Google Drive. Si utilizas la autenticación de cuentas de servicio de Google, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "clientEmail": "user account email",  
    "adminAccountEmail": "service account email",  
    "privateKey": "private key"  
}
```

Si usas la autenticación OAuth 2.0, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "clientID": "OAuth client ID",  
    "clientSecret": "client secret",  
    "refreshToken": "refresh token"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendado reutilizar

las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSource para proporcionar un IAMrol con permisos para acceder a su Secrets Managersecreto y para llamar a las API públicas requeridas para el conector de Google Drive y Amazon Kendra. Para obtener más información, consulte [IAMfunciones para las fuentes de datos de Google Drive](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfiguration cuando llamasCreateDataSource. Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).
- Modo de sincronización—Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:
 - FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice
 - FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice
 - CHANGE_LOG para rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice

Important

La API de Google Drive no admite la recuperación de comentarios de un archivo eliminado permanentemente. Los comentarios de los archivos desecharados se pueden recuperar. Cuando se elimina un archivo, el conector eliminará los comentarios del Amazon Kendraíndice.

- Habilitar el rastreador de identidades: especifique si desea activar el rastreador de identidades. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el [PutPrincipalMapping](#) API. Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado en contexto de usuario](#).
- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir determinadas cuentas de usuario, unidades compartidas, tipos de MIME y archivos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Solo puede incluir campos de indexación para Amazon KendraConektor Google Drive. La asignación de campos personalizada no está disponible para el conector de Google Drive debido a las limitaciones de la API. Para obtener más información, consulte [Mapping data source fields \(Asignación de campos de origen de datos\)](#).
- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Google Drive](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [Esquema de plantillas de Google Drive](#).

Notas

- El mapeo de campos personalizados no está disponible para el conector de Google Drive, ya que la interfaz de usuario de Google Drive no admite la creación de campos personalizados.
- La API de Google Drive no admite la recuperación de comentarios de un archivo eliminado permanentemente. Sin embargo, los comentarios se pueden recuperar en el caso de los archivos desecharados. Cuando se elimina un archivo, el Amazon Kendra el conector eliminará los comentarios del Amazon Kendra índice.
- La API de Google Drive no devuelve los comentarios presentes en un archivo.docx.

Jira

Jira es una herramienta de gestión de proyectos para el desarrollo de software, la gestión de productos y el seguimiento de errores. Puedes usar Amazon Kendra para indexar tus proyectos, problemas, comentarios, archivos adjuntos, registros de trabajo y estados de Jira.

Amazon Kendra actualmente solo es compatible con Jira Cloud.

Puedes conectarte a Amazon Kendra a tu fuente de datos de Jira mediante el [Amazon Kendra console](#) o la [Jira Configuration API](#). Para obtener una lista de las funciones compatibles con cada una, consulte [Características admitidas \(p. 376\)](#).

Para solucionar problemas de su Amazon Kendra Conector de fuente de datos de Jira, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 376\)](#)
- [Requisitos previos \(p. 376\)](#)
- [Instrucciones de conexión \(p. 377\)](#)
- [Más información \(p. 380\)](#)

Características admitidas

Amazon Kendra El conector de fuente de datos de Jira admite las siguientes funciones:

- Change log
- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de poder usar Amazon Kendra para indexar tu fuente de datos de Jira, realiza estos cambios en tu Jira y AWS cuentas.

En Jira, asegúrate de tener:

- Cree credenciales de autenticación con token de la API de Jira que incluyen un ID de Jira (nombre de usuario o correo electrónico) y una credencial de Jira (token de API de Jira). Consulte [Documentación de Atlassian sobre la gestión de los tokens de API](#).
- Apunté la URL de la cuenta de Jira en la configuración de tu cuenta de Jira. Por ejemplo, `company.atlassian.net`.
- Se comprobó que cada documento es único en Jira y en otras fuentes de datos que planeas usar para el mismo índice. Cada fuente de datos que deseas utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tuIAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Has guardado tus credenciales de autenticación de Jira en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Jira aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tu fuente de datos de Jira, debes proporcionar los detalles necesarios de tu fuente de datos de Jira para queAmazon Kendrapuede acceder a sus datos. Si aún no has configurado Jira paraAmazon Kendra, consulte [Requisitos previos \(p. 376\)](#).

Console

Para conectarAmazon Kendraa Jira

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseas utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConejor Jiray, a continuación, elijaAregar fuente de datos.
5. En elEspecificardetalles de la fuente de datospágina, introduzca la siguiente información:

- a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, parIdioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:
- a. URL de la cuenta de Jira—Introduce la URL de tu cuenta de Jira.
 - b. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar tus credenciales de autenticación de Jira. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - i. Introduzca la siguiente información en elCrea unAWS Secrets Managerventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Jira-' se añade automáticamente a tu nombre secreto.
 - B. ParaID de Jira—Introduzca el nombre de usuario o el correo electrónico de Jira.
 - C. ParaContraseña/Token—Introduce el token de la API de Jira que creaste desde tu cuenta de Jira.
 - ii. Seleccione Guardar.
 - c. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesyGrupos de seguridad de VPC.
 - d. IAMpapel—Elige uno existenteIAMrol o crea uno nuevolAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- e. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
- a. Selecciona los proyectos de Jira que quieras indexar—Las entidades o tipos de contenido de Jira que quieras rastrear.
 - b. Estados,Elementos adicionales, yTipos de problemas—Seleccione el contenido para refinar el alcance del índice.
 - c. Registro de cambios—Seleccione esta opción para actualizar el índice en lugar de sincronizar todos los archivos.
 - d. Patrones Regex—Patrones de expresiones regulares para incluir o excluir ciertos archivos. Puede añadir hasta 100 patrones.
 - e. EnCronograma de ejecución de sincronización, paraFrecuencia—Elija la frecuencia con la que Amazon Kendra se sincronizará con su fuente de datos.
 - f. Elija Siguiente.
8. En elDefinir mapeos de campospágina, introduzca la siguiente información:
- a. ParaProyecto,Asunto,Comentario,Adjuntivo,Registro de trabajo—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.

- b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en elFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Jira

Debe especificar lo siguiente mediante el[JiraConfiguration API](#):

- URL de la fuente de datos—Especifica la URL de tu cuenta de Jira. Por ejemplo,[company.atlassian.net](#).
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de tu cuenta de Jira. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "jiraId": "Jira user name or email",  
    "jiraCredential": "Jira API token"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendable reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAM papel—EspecificareRoleArn cuando llamas `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para el conector de Jira y Amazon Kendra. Para obtener más información, consulte [IAM funciones para las fuentes de datos de Jira](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—Especificar `VpcConfiguration` como parte de la configuración de la fuente de datos. Consulte [Configurando Amazon Kendra para usar una VPC](#).
- Registro de cambios—Si Amazon Kendra debe utilizar el mecanismo de registro de cambios de la fuente de datos de Jira para determinar si un documento debe agregarse, actualizarse o eliminarse del índice.

Note

Usa el registro de cambios si no quieres Amazon Kendra para escanear todos los documentos. Si el registro de cambios es grande, podría tardar Amazon Kendra menos tiempo para escanear los documentos de la fuente de datos de Jira que para procesar el registro de cambios. Si sincronizas tu fuente de datos de Jira con tu índice por primera vez, se escanean todos los documentos.

- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir ciertos proyectos, problemas, comentarios, archivos adjuntos, registros de trabajo y estados.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elige asignar los campos de tu fuente de datos de Jira a tu Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Jira](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con tu fuente de datos de Jira, consulta:

- [Busca tus proyectos de Jira de forma inteligente con Amazon Kendra Conector Jira Cloud](#)

Microsoft Exchange

Microsoft Exchange es una herramienta de colaboración empresarial para mensajería, reuniones y uso compartido de archivos. Si es usuario de Microsoft Exchange, puede utilizar Amazon Kendra para indexar la fuente de datos de Microsoft Exchange.

Puedes conectarte Amazon Kendra a su fuente de datos de Microsoft Exchange mediante [Amazon Kendra console](#) y el [Template Configuration API](#).

Para solucionar problemas de Amazon Kendra Conector de fuentes de datos de Microsoft Exchange, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Características admitidas

- Mapeos de campo
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados

Requisitos previos

Antes de poder usar Amazon Kendra para indexar su fuente de datos de Microsoft Exchange, realice estos cambios en su Microsoft Exchange y AWS cuentas.

En Microsoft Exchange, asegúrese de tener:

- Creó una cuenta de Microsoft Exchange en Office 365.
- Tomó nota de su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
- Creó una aplicación OAuth en el portal de Azure y anotó el ID de inquilino, el ID del cliente y el secreto del cliente o las credenciales del cliente. Consulte [Tutorial de Microsoft](#) para obtener más información.

- Se agregaron los siguientes permisos para la aplicación de conectores:

Microsoft Graph	Office 365 Exchange en línea
<ul style="list-style-type: none">• Mail.Read (Aplicación)• Correo.ReadBasic(Solicitud)• Correo.ReadBasic.All (Aplicación)• Calendars.Read (Aplicación)• User.Read.All (Aplicación)• Contacts.Read (Aplicación)• Notes.Read.All (Aplicación)• Directory.Read.All (Aplicación)• NOTICIAS.AccessAsUser.Todos (delegados) <ul style="list-style-type: none">• Se comprobó que cada documento es único en Microsoft Exchange y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.	full_access_as_app (Aplicación)

En tuCuenta de AWS, asegúrate de tener:

- Creó unAmazon Kendraíndicey, si usa la API, anota el ID del índice.
- Creó unIAMpapelpara su fuente de datos y, si usa la API, anotó el ARN dellIAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulIAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Almacenó sus credenciales de autenticación de Microsoft Exchange en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto al conectar la fuente de datos de Microsoft Exchange aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa su fuente de datos de Microsoft Exchange, debe proporcionar los detalles necesarios de su fuente de datos de Microsoft Exchange para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configurado Microsoft Exchange paraAmazon Kendra, consulte[Requisitos previos \(p. 380\)](#).

Console

Para conectarAmazon Kendraa Microsoft Exchange

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).

2. En el panel de navegación de la izquierda, seleccionalíndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAgregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConector Microsoft Exchangey, a continuación, elijaAregar fuente de datos.
5. En elEspecificardetalle de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefinardel acceso y la seguridadpágina, introduzca la siguiente información:
 - a. Fuente—Introduzca su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
 - b. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar sus credenciales de autenticación de Microsoft Exchange. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - i. Introduzca la siguiente información enCrea unAWS Secrets Managerventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Microsoft Exchange'
 - B. ParalD de clienteySecreto del cliente—Introduzca los valores de las credenciales de autenticación que creó en su cuenta de Microsoft Exchange en el portal de Azure.
 - ii. Seleccione Guardar.
 - c. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesyGrupos de seguridad de VPC.
 - d. IAMpapel—Elige uno existenteIAMrol o crea uno nuevolAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- e. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. Sincronizar contenido—Seleccione el contenido que desea sincronizar.
 - b. Configuración adicional—Si lo desea, puede indexar el siguiente contenido en lugar de sincronizar todos los documentos.
 - Tipos de entidades— Elige las entidades que quieras sincronizar. Puedes elegir entreCalendario,OneNotesyContactos.

- Rastreo de calendarios— Introduzca la fecha de inicio y finalización de la sincronización de su calendario.
 - Incluir correo electrónico— Introduzca el correo electrónico deyEnvíe un correo electrónico adominios y cualquierAsunto líneas que desea incluir o excluir en su índice.
 - Regex para dominios— Añada patrones para incluir y excluir ciertos dominios de correo electrónico de su índice.
 - Patrones Regex— Añada patrones de expresiones regulares para incluir o excluir ciertos archivos. Puede añadir hasta 100 patrones.
- c. Modo de sincronización—Puede elegir cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos.
- i. Si eliges la sincronización completa,Amazon Kendrasincronizará todo el contenido de todas las entidades, independientemente del estado de sincronización anterior.
 - ii. Si eliges la sincronización de contenido nuevo o modificado,Amazon Kendrasolo sincronizará contenido nuevo o modificado.
 - iii. Si eliges la sincronización de contenido nuevo, modificado o eliminado,Amazon Kendrasolo sincronizará contenido nuevo, modificado o eliminado.
8. En elEstablecer mapeos de campospágina, introduzca la siguiente información:
- a. Campos de fuente de datos predeterminados—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Microsoft Exchange

Debe especificar un JSON de [el esquema de fuente de datos](#) utilizando el [TemplateConfiguration API](#). Debe proporcionar la siguiente información:

- Fuente de datos—Debe especificar la fuente de datos como MSEXCHANGE.
- ID de inquilino—Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
- Escriba—EspecificareTEMPLATEcomo el tipo cuando llamasCreateDataSource.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de su cuenta de Microsoft Exchange. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "clientId": "client ID",  
    "clientSecret": "client secret"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar

las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSource para proporcionar un IAMrol con permisos para acceder a su Secrets Managersecreto y para llamar a las API públicas requeridas para el conector de Microsoft Exchange y Amazon Kendra. Para obtener más información, consulte [IAMfunciones para fuentes de datos de Microsoft Exchange](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfiguration cuando llamasCreateDataSource. Para obtener más información, consulte [Configurando Amazon Kendra utilizar un Amazon VPC \(p. 465\)](#).
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinadas páginas y recursos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elija asignar los campos de la fuente de datos de Microsoft Exchange a su Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields \(Asignación de campos de origen de datos\)](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Microsoft Exchange, consulte:

- [Indexe su contenido de Microsoft Exchange mediante el conector de Exchange para Amazon Kendra](#)

MicrosoftOneDrive

MicrosoftOneDrive es un servicio de almacenamiento basado en la nube que puedes usar para almacenar, compartir y alojar tu contenido. Puedes usar Amazon Kendra para indexar sus OneDrive fuente de datos.

Puedes conectarte Amazon Kendra a tu OneDrive fuente de datos mediante [Amazon Kendra console](#) y [el OneDrive Configuration API](#).

Amazon Kendra tiene dos versiones del OneDrive conector. Las funciones compatibles de cada versión incluyen:

MicrosoftOneDrive conector V1.0/[OneDriveConfiguration API](#)

- Mapeos de campo
- Filtros de inclusión/exclusión

MicrosoftOneDrive conector V2.0/[TemplateConfiguration API](#)

- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión

- Funcionalidad de rastreador de identidades
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados

Note

Soporte paraOneDriveconector V1.0/OneDriveConfigurationEstá previsto que la API finalice en junio de 2023. Recomendamos usarOneDriveconector V2.0/TemplateConfigurationAPI.

Para solucionar problemas deAmazon Kendra OneDriveconector de fuente de datos, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [MicrosoftOneDriveconector V1.0 \(p. 385\)](#)
- [MicrosoftOneDriveconector V2.0 \(p. 389\)](#)
- [Más información \(p. 393\)](#)

MicrosoftOneDriveconector V1.0

MicrosoftOneDrivees un servicio de almacenamiento basado en la nube que puedes usar para almacenar, compartir y alojar tu contenido. Puedes usarAmazon Kendrapara indexar tu MicrosoftOneDrivefuente de datos.

Note

Soporte paraOneDriveconector V1.0/MicrosoftOneDriveEstá previsto que la API finalice en junio de 2023. Recomendamos usarOneDriveconector V2.0/TemplateConfigurationAPI.

Para solucionar problemas deAmazon Kendra OneDriveconector de fuente de datos, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 385\)](#)
- [Requisitos previos \(p. 385\)](#)
- [Instrucciones de conexión \(p. 386\)](#)

Características admitidas

- Mapeos de campo
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder usarAmazon Kendrapara indexar susOneDrivefuente de datos, realice estos cambios en suOneDriveyAWS cuentas.

En su Azure Active Directory (AD), asegúrese de tener:

- Creó una aplicación de Azure Active Directory (AD).
- Utilizó el ID de la aplicación de AD para registrar una clave secreta para la aplicación en el sitio de AD. La clave secreta debe contener el ID de la aplicación y una clave secreta.
- Copió el dominio de AD de la organización.
- Se agregaron los siguientes permisos a la aplicación de AD en la opción Microsoft Graph:
 - Lea los archivos de todas las colecciones de sitios (File.Read.All)

- Lea el perfil completo de todos los usuarios (User.Read.All)
- Lea los datos del directorio (Directory.Read.All)
- Leer todos los grupos (Group.Read.All)
- Lea los elementos de todas las colecciones del sitio (Site.Read.All)
- Copió la lista de usuarios cuyos documentos se deben indexar. Puede optar por proporcionar una lista de nombres de usuario o puede proporcionar los nombres de usuario de un archivo almacenado en unAmazon S3. Después de crear la fuente de datos, puede:
 - Modifique la lista de usuarios.
 - Cambiar de una lista de usuarios a una lista almacenada en unAmazon S3balde.
 - Cambia elAmazon S3ubicación del bucket de una lista de usuarios. Si cambias la ubicación del depósito, también debes actualizar elIAMrol de la fuente de datos para que tenga acceso al bucket.

Note

Si almacena la lista de nombres de usuario en unAmazon S3balde, elIAMla política de la fuente de datos debe proporcionar acceso al bucket y el acceso a la clave con la que se cifró el bucket, si la hubiera.

- Se ha marcado que cada documento es único enOneDrivey en otras fuentes de datos que planee utilizar para el mismo índice. Cada fuente de datos que desee utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Almacenó suOneDrivecredenciales de autenticación en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tuOneDrivefuente de datos paraAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tuOneDrivefuente de datos: debe proporcionar detalles de suOneDrivecredenciales para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configuradoOneDriveporAmazon Kendraver[Requisitos previos \(p. 385\)](#).

Console

Para conectarAmazon KendraaOneDrive

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).

2. En el panel de navegación de la izquierda, selecciona Índices y, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar su Control de acceso de usuarios en Ajustes de índice.

3. En el Primeros pasospágina, elige Agregar fuente de datos.
4. En el Agregar fuente de datospágina, elige OneDrive conectory, a continuación, elija Agregar fuente de datos.
5. En el Especificar los detalles de la fuente de datospágina, introduzca la siguiente información:
 - a. En Nombre y descripción, para Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. (Opcional) Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, para idioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, para Añadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tus AWS costos.
 - e. Elija Siguiente.
6. En el Defina el acceso y la seguridadpágina, introduzca la siguiente información:
 - a. OneDriveID de inquilino—Introduzca el OneDriveID de inquilino sin el protocolo.
 - b. Tipo de autenticación—Elige entre Nuevo y Existente.
 - i. Si eliges Existente, selecciona un secreto existente para Selecciona el secreto.
 - ii. Si eliges Nuevo, introduzca la siguiente información en el Nuevo AWS Secrets Manager secretosección:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-OneDrive-' se añade automáticamente a tu nombre secreto.
 - B. Para ID de aplicación y Contraseña de aplicación—Introduzca los valores de las credenciales de autenticación de su OneDrive cuenta y, a continuación, elige Guardar autenticación.
 - d. IAM papel—Elige uno existente IAM rol o crea uno nuevo IAM rol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

- e. Elija Siguiente.
7. En el Configurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. Elige entre Archivo de listay Lista de nombres según su caso de uso.
 - i. Si eliges Archivo de lista, introduzca la siguiente información:
 - Selecciona ubicación—Introduzca la ruta a su Amazon S3 balde.Agregar archivo de lista de usuarios a Amazon S3—Seleccione esta opción para añadir los archivos de su lista de usuarios a su Amazon S3 balde.
 - Mapeos de grupos locales de usuarios—Seleccione utilizar el mapeo de grupos locales para filtrar el contenido.

- ii. Si eligesLista de nombres, introduzca la siguiente información:
 - Nombre de usuario—Introduzca hasta 10 unidades de usuario para indexarlas. Para añadir más de 10 usuarios, cree un archivo que contenga los nombres.

Añadir otro—Elija añadir más usuarios.

Mapeos de grupos locales de usuarios—Seleccione utilizar el mapeo de grupos locales para filtrar el contenido.

 - b. ParaConfiguraciones adicionales—Añada patrones de expresiones regulares para incluir o excluir determinados archivos. Puede añadir hasta 100 patrones.
 - c. EnCronograma de ejecución de sincronización, paraFrecuencia—Elige con qué frecuenciaAmazon Kendrase sincronizará con su fuente de datos.
 - d. Elija Siguiente.
8. En elEstablecer mapeos de campospágina, introduzca la siguiente información:
 - a. ParaCampos de fuente de datos predeterminadosyMapeos de campos sugeridos adicionales—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAgregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en elFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon KendraaOneDrive

Debe especificar lo siguiente mediante el[OneDriveConfiguration](#)API:

- ID de inquilino—Especifique el dominio de Azure Active Directory de la organización.
- OneDriveUsuarios—Especifique la lista de cuentas de usuario cuyos documentos se deben indexar.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de suOneDrivecuenta. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "OAuth client ID",  
    "password": "client secret"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificareRoleArn cuando llamasCreateDataSourcepara proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para elOneDriveconector yAmazon Kendra. Para obtener más información, consulte[IAMfunciones paraOneDrivefuentes de datos](#).

También puede añadir las siguientes funciones opcionales:

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir ciertos documentos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elija mapear suOneDrivecampos de fuente de datos para suAmazon Kendracampos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendrarastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario paraOneDrivefuentes de datos](#).

MicrosoftOneDriveconector V2.0

MicrosoftOneDrivees un servicio de almacenamiento basado en la nube que puedes usar para almacenar, compartir y alojar tu contenido. Puedes usarAmazon Kendrapara indexar susOneDrivefuente de datos.

Puedes conectarteAmazon Kendraa tuOneDrivefuente de datos mediante [Amazon Kendraconsolay elOneDriveConfigurationAPI](#).

Note

Soporte paraOneDriveConektor V1.0/OneDriveConfigurationEstá previsto que la API finalice en junio de 2023. Recomendamos usarOneDriveConektor V2.0/TemplateConfigurationAPI. La versión 2.0 proporciona ACL adicionales y funciones de rastreador de identidades.

Para solucionar problemas deAmazon Kendra OneDriveconector de fuente de datos, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 389\)](#)
- [Requisitos previos \(p. 389\)](#)
- [Instrucciones de conexión \(p. 390\)](#)

Características admitidas

Amazon Kendra OneDriveel conector de fuente de datos admite las siguientes funciones:

- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Funcionalidad de rastreador de identidades
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados

Requisitos previos

Antes de poder usarAmazon Kendrapara indexar susOneDrivefuente de datos, realice estos cambios en suOneDriveyAWS cuentas.

En su Azure Active Directory (AD), asegúrese de tener:

- Creó unOneDrivecuenta en Office 365.
- Tomó nota de su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
- Creó una aplicación de Azure Active Directory (AD).
- Utilizó el ID de la aplicación de AD para registrar una clave secreta para la aplicación en el sitio de AD. La clave secreta debe contener el ID de la aplicación y una clave secreta.
- Copió el dominio de AD de la organización.
- Se agregaron los siguientes permisos a la aplicación de AD en la opción Microsoft Graph:
 - Lea los archivos de todas las colecciones de sitios (File.Read.All)
 - Lea los perfiles completos de todos los usuarios (User.Read.All)
 - Leer todos los grupos (Group.Read.All)
 - Lea todas las notas (Notes.Read.All)
- Copió la lista de usuarios cuyos documentos se deben indexar. Puede optar por proporcionar una lista de nombres de usuario o puede proporcionar los nombres de usuario de un archivo almacenado en unAmazon S3. Después de crear la fuente de datos, puede:
 - Modifique la lista de usuarios.
 - Cambiar de una lista de usuarios a una lista almacenada en unAmazon S3balde.
 - Cambia elAmazon S3ubicación del bucket de una lista de usuarios. Si cambias la ubicación del depósito, también debes actualizar elIAMrol de la fuente de datos para que tenga acceso al bucket.

Note

Si almacena la lista de nombres de usuario en unAmazon S3balde, elIAMla política de la fuente de datos debe proporcionar acceso al bucket y el acceso a la clave con la que se cifró el bucket, si la hubiera.
ElOneDriveusos del conectorCorreo electrónico de la información de contactopresente en elPropiedades de usuario de Onedrive. Asegúrese de que el usuario cuyos datos desea rastrear tenga el campo de correo electrónico configurado en elInformación de contactoen cuanto a los nuevos usuarios, esta página podría estar en blanco.

En tuAWScuenta, asegúrate de tener:

- Creó unAmazon Kendraíndice y, si usa la API, anota el identificador del índice.
- Creó unIAMrol para su fuente de datos y, si usa la API, anotó el ARN dellAMrol.
- Almacenó suOneDrivecredenciales de autenticación en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tuOneDrivefuente de datos paraAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tuOneDrivefuente de datos: debe proporcionar detalles de suOneDrivecredenciales para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configuradoOneDriveporAmazon Kendra, consulte[Requisitos previos \(p. 389\)](#).

Console

Para conectarAmazon KendraaOneDrive

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).

2. En el panel de navegación de la izquierda, selecciona Índices y, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar su Control de acceso de usuarios en Ajustes de índice.

3. En el Primeros pasospágina, elige Agregar fuente de datos.
4. En el Agregar fuente de datospágina, elige OneDriveconectory, a continuación, elija Agregar fuente de datos.
5. En el Especificar los detalles de la fuente de datospágina, introduzca la siguiente información:
 - a. En Nombre y descripción, para Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. (Opcional) Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, para idioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, para Añadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tus AWS costos.
 - e. Elija Siguiente.
6. En el Defina el acceso y la seguridadpágina, introduzca la siguiente información:
 - a. OneDriveID de inquilino—Introduzca el OneDriveID de inquilino sin el protocolo.
 - b. En autenticación—Elige entre Nuevo y Existente.
 - i. Si eliges Existente, selecciona un secreto existente para Selecciona el secreto.
 - ii. Si eliges Nuevo, introduzca la siguiente información en el Nuevo AWS Secrets Manager secreto sección:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-OneDrive-' se añade automáticamente a tu nombre secreto.
 - B. Para ID de cliente y Secreto del cliente—Introduzca el ID del cliente y el secreto del cliente y, a continuación, seleccione Guardar autenticación.
 - c. En Configurar VPC y grupo de seguridad (opcional), para Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadir Subredes y Grupos de seguridad de VPC.
 - d. Rastreador de identidades—Elija activar Amazon Kendra rastreador de identidad para sincronizar la información de identidad. Si decide desactivar el rastreador de identidades, debe cargar la información principal mediante el [PutPrincipalMapping API](#).
 - e. IAM papel—Elige uno existente IAM rol o crea uno nuevo IAM rol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

- g. Elija Siguiente.
7. En el Configurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
8. a. Para Alcance de sincronización—Elige qué usuarios OneDrive datos para indexar. Puede añadir un máximo de 10 usuarios manualmente.
- b. Para Configuraciones adicionales—Añada patrones de expresiones regulares para incluir o excluir cierto contenido. Puede añadir hasta 100 patrones.
- c. En Modo de sincronización, elija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Sincronización completa indexa todo el contenido, independientemente

- del estado de sincronización anterior. Sincronización de documentos nuevos, modificados o eliminados solo sincroniza los documentos nuevos, modificados o eliminados.
- d. En Cronograma de ejecución de sincronización, para Frecuencia—Elige con qué frecuencia Amazon Kendra sincronizará con su fuente de datos.
 - e. Elija Siguiente.
9. En el Establecer mapeos de campospágina, introduzca la siguiente información:
 - a. Para Campos de fuente de datos predeterminadosy Mapeos de campos sugeridos adicionales—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Elija Siguiente.
 10. En el Revisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en el Fuentes de datos página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a OneDrive

Debe especificar un JSON del [esquema de fuente de datos](#) utilizando el [TemplateConfiguration API](#). Debe proporcionar la siguiente información:

- Fuente de datos—Especifique la fuente de datos como ONEDRIVEV2.
- ID de inquilino—Especifique el ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
- Técnica—Especificare TEMPLATE como el tipo cuando llamas `CreateDataSource`.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contiene las credenciales de autenticación que creó en su OneDrive cuenta.

Si usas la autenticación OAuth 2.0, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "clientID": "OAuth client ID",  
    "clientSecret": "client secret"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendado reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAM papel—Especificar RoleArn cuando llamas `CreateDataSource` para proporcionar un IAM role con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el OneDrive conector y Amazon Kendra. Para obtener más información, consulte [IAM funciones para OneDrive fuentes de datos](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—Especificar `VpcConfiguration` cuando llamas `CreateDataSource`. Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).

- Modo de sincronización—Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:
 - FORCED_FULL_CRAWL para rastrear y sincronizar todo el contenido con tu índice
 - FULL_CRAWL para rastrear todo el contenido y sincronizar solo el contenido nuevo, modificado o eliminado
 - CHANGE_LOG para rastrear y sincronizar solo contenido nuevo, modificado y eliminado.
- Rastreador de identidades—Especifica si se va a activar Amazon Kendra rastreador de identidad. Si el rastreador de identidades está desactivado, debe cargar la información de identidad o principal mediante el [PutPrincipalMapping API](#). Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario](#).
- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir ciertos archivos, OneNote secciones, y OneNote páginas.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Solo puede incluir campos de indexación para Amazon Kendra OneDrive conector. El mapeo de campos personalizado no está disponible para OneDrive conector debido a las limitaciones de la API. Para obtener más información, consulte [Mapping data source fields \(Asignación de campos de origen de datos\)](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [MicrosoftOneDrive esquema de plantilla \(p. 201\)](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con tu OneDrive fuente de datos, consulte:

- [Anunciamos la actualización de MicrosoftOneDrive conector \(V2\) para Amazon Kendra.](#)

MicrosoftSharePoint

SharePoint es un servicio colaborativo de creación de sitios web que puede utilizar para personalizar el contenido web y crear páginas, sitios, bibliotecas de documentos y listas. Puedes usar Amazon Kendra para indexar sus SharePoint fuente de datos.

Amazon Kendra actualmente es compatible con SharePoint En línea y SharePoint Servidor (versiones 2013, 2016, 2019 y edición de suscripción).

Puedes conectarte a Amazon Kendra a tu SharePoint fuente de datos mediante el [Amazon Kendra consola](#), el [TemplateConfiguration API](#), o [SharePoint Configuration API](#).

Amazon Kendra tiene dos versiones del SharePoint conector. Las funciones compatibles de cada versión incluyen:

SharePoint Conector V1.0/[SharePoint Configuration API](#)

- Change log
- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

SharePointConejor V2.0/[TemplateConfigurationAPI](#)

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos o modificados/Sincronizar solo documentos nuevos, modificados o eliminados

Note

Soporte para SharePointConejor V1.0/SharePointConfigurationEstá previsto que la API finalice en 2023. Recomendamos migrar a o usar SharePointConejor V2.0/TemplateConfigurationAPI.

Para solucionar problemas de Amazon Kendra SharePointConejor de fuente de datos, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [SharePointConejor V1.0 \(p. 394\)](#)
- [SharePointConejor V2.0 \(p. 401\)](#)

SharePointConejor V1.0

SharePointes un servicio colaborativo de creación de sitios web que puede utilizar para personalizar el contenido web y crear páginas, sitios, bibliotecas de documentos y listas. Si eres un SharePointusuario, puedes usar Amazon Kendrapara indexar sus SharePointfuente de datos.

Note

Soporte para SharePointConejor V1.0/SharePointConfigurationEstá previsto que la API finalice en 2023. Recomendamos migrar a o usar SharePointConejor V2.0/TemplateConfigurationAPI.

Para solucionar problemas de Amazon Kendra SharePointConejor de fuente de datos, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 394\)](#)
- [Requisitos previos \(p. 395\)](#)
- [Instrucciones de conexión \(p. 396\)](#)
- [Más información \(p. 400\)](#)

Características admitidas

- Change log
- Mapeos de campo
- Filtrado de contexto de usuario

- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de poder usar Amazon Kendra para indexar su SharePoint fuente de datos, realice estos cambios en su SharePoint AWS cuentas.

En SharePoint, asegúrate de tener:

- Tomó nota de la URL del SharePoint sitios que desea indexar.
- Para SharePoint En línea:
 - Tomó nota de sus credenciales de autenticación básicas, que contienen un nombre de usuario y una contraseña con permisos de administrador del sitio.
 - Opcional: Credenciales de OAuth 2.0 generadas que contienen un nombre de usuario, una contraseña, un identificador de cliente y un secreto de cliente.
 - Desactivado Valores predeterminados de seguridad en su portal de Azure mediante un usuario administrativo. Para obtener más información sobre la administración de la configuración predeterminada de seguridad en el portal de Azure, consulte [Documentación de Microsoft sobre cómo habilitar/deshabilitar los valores predeterminados de seguridad](#).
- Para SharePoint Servidor:
 - Tomó nota de su SharePoint Nombre de dominio del servidor (el nombre de NetBIOS en Active Directory). Usas esto, junto con tu SharePoint nombre de usuario y contraseña de autenticación básica, para conectarse SharePoint Servidor para Amazon Kendra.

Note

Si usa SharePoint Servidor y necesita convertir su lista de control de acceso (ACL) a formato de correo electrónico para filtrar según el contexto del usuario, proporcionar la URL del servidor LDAP y la base de búsqueda de LDAP. O puede utilizar la anulación del dominio del directorio. La URL del servidor LDAP es el nombre de dominio completo y el número de puerto (por ejemplo, ldap://example.com:389). La base de búsqueda de LDAP son los controladores de dominio «example» y «com». Con la anulación del dominio del directorio, puede utilizar el dominio de correo electrónico en lugar de utilizar la URL del servidor LDAP y la base de búsqueda LDAP. Por ejemplo, el dominio de correo electrónico de username@example.com es «example.com». Puedes usar esta anulación si no te preocupa validar tu dominio y simplemente quieres usar tu dominio de correo electrónico.

- Se agregaron los siguientes permisos a su SharePoint cuenta:

Para SharePoint listas

- Abrir elementos: permite ver el origen de los documentos con los controladores de archivos del lado del servidor.
- Ver páginas de aplicaciones: permite ver formularios, vistas y páginas de aplicaciones. Enumere las listas.
- Ver elementos: permite ver los elementos de las listas y los documentos de las bibliotecas de documentos.
- Ver versiones: permite ver versiones anteriores de un elemento o documento de la lista.

Para SharePoint sitios web

- Navegar por directorios: enumere los archivos y carpetas de un sitio web mediante SharePoint Diseñador e interfaz DAV web.
- Buscar información de usuario: permite ver información sobre los usuarios del sitio web.
- Enumerar permisos: enumere los permisos del sitio web, la lista, la carpeta, el documento o el elemento de la lista.

- Abrir: abre un sitio web, una lista o una carpeta para acceder a los elementos del contenedor.
- Utilice las funciones de integración de clientes: utilice SOAP, WebDAV, el modelo de objetos del cliente o SharePointInterfaces de diseñador para acceder al sitio web.
- Utilice interfaces remotas: utilice funciones que inicien aplicaciones cliente.
- Ver páginas: permite ver las páginas de un sitio web.
- Se ha marcado que cada documento es único en SharePoint en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que desee utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tuIAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Almacenó suSharePointcredenciales de autenticación en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tuSharePointfuente de datos paraAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tuSharePointfuente de datos: debe proporcionar detalles de suSharePointcredenciales para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configuradoSharePointporAmazon Kendraver[Requisitos previos \(p. 395\)](#).

Console

Para conectarAmazon Kendraa SharePoint

1. Inicia sesión en elAWSConsola de administración y abra el[Amazon Kendraconsola](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeSharePointconector v1.0y, a continuación, elijaAregar fuente de datos.
5. En elEspecificar los detalles de la fuente de datospágina, introduzca la siguiente información:

- a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, parIdioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:
- a. ParaMétodo de alojamiento—Elige entreSharePointEn línea y SharePointServidor.
 - i. ParaSharePointEn línea—Introduzca elURL del sitio específicas para suSharePointrepositorio.
 - ii. ParaSharePointServidor—Elige tuSharePointversión, introduceURL del sitio específicas para suSharePointrepositorio, e introduzca elAmazon S3camino hacia tuUbicación del certificado SSL.
 - b. (SharePointSolo servidor) ParaProxy web—Introduzca elNombre del hostyNúmero de puertode tu interiorSharePointinstancia. El número de puerto debe ser un valor numérico comprendido entre 0 y 65535.
 - c. Paraautenticación—Elija entre las siguientes opciones en función de su caso de uso:
 - i. ParaSharePointEn línea: elija entreAutenticación básica y Autenticación OAuth 2.0.
 - ii. ParaSharePointServidor: elija entreNinguna, LDAP, yManual.
 - d. ParaAWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para guardar tusSharePointcredenciales de autenticación. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta. Debe introducir unNombre secreto. El prefijo 'AmazonKendra-SharePoint-' se añade automáticamente a tu nombre secreto.
 - e. Introduzca la siguiente información adicional en elCrea unAWS Secrets Managerventana secreta:
 - i. Elige una de las siguientes opcionesSharePointOpciones de autenticación en la nube, según su caso de uso:
 - A. Autenticación básica—Introduzca suSharePointnombre de usuario de la cuenta comoNombre de usuarioySharePointcontraseña de cuenta comoContraseña.
 - B. Autenticación OAuth 2.0—Introduzca suSharePointnombre de usuario de la cuenta comoNombre de usuario, SharePointcontraseña de cuenta comoContraseña, tu único generado automáticamenteSharePointID comolD de cliente, y la cadena secreta compartida utilizada por ambosSharePointyAmazon KendratanSecreto de cliente.
 - ii. Elige una de las siguientes opcionesSharePointOpciones de autenticación del servidor, según su caso de uso:
 - A. Ninguna—Introduzca suSharePointnombre de usuario de la cuenta comoNombre de usuario, tuSharePointcontraseña de cuenta comoContraseña, y tuNombre de dominio del servidor.
 - B. LDAP—Introduzca suSharePointnombre de usuario de la cuenta comoNombre de usuario, SharePointcontraseña de cuenta comoContraseña, tuTerminal de servidor LDAP(incluidos el protocolo y el número de puerto, por ejemplo`ldap://example.com:389`), y tuBase de búsqueda LDAP(por ejemplo,`dc=ejemplo, dc=com`).

- C. Manual—Introduzca su SharePoint nombre de usuario de la cuenta como Nombre de usuario, tu SharePoint contraseña de cuenta como Contraseña, y tu Anulación de dominio de correo electrónico (dominio de correo electrónico del usuario o grupo del directorio).
 - iii. Seleccione Guardar.
 - f. Nube privada virtual (VPC)— También debes añadir Subredes y Grupos de seguridad de VPC.

Note

Debe usar una VPC si usa SharePoint Servidor. Amazon VPC es opcional para otros SharePoint versiones.
 - g. IAM papel—Elige uno existente IAM rol o crea uno nuevo IAM rol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.
- h. Elija Siguiente.
7. En el Configurar los ajustes de sincronización página, introduzca la siguiente información:
 - a. Utilice el registro de cambios—Seleccione esta opción para actualizar el índice en lugar de sincronizar todos los archivos.
 - b. Rastrea archivos adjuntos—Seleccione esta opción para rastrear los archivos adjuntos.
 - c. Utilice mapeos de grupos locales—Seleccione esta opción para asegurarse de que los documentos estén filtrados correctamente.
 - d. Configuración adicional—Añada patrones de expresiones regulares para incluir o excluir determinados archivos. Puede añadir hasta 100 patrones.
 - e. En Cronograma de ejecución de sincronización por Frecuencia— Con qué frecuencia Amazon Kendra se sincronizará con su fuente de datos.
 - f. Elija Siguiente.
 8. En el Definir mapeos de campos página, introduzca la siguiente información:
 - a. Amazon Kendra mapeos de campos predeterminados—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Para Mapeos de campos personalizados—Añada campos de fuente de datos personalizados para crear un nombre de campo de índice al que asignar y el tipo de datos del campo.
 - c. Elija Siguiente.
 9. En el Revisar y crear página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en Fuentes de datos página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a SharePoint

Debe especificar lo siguiente mediante [SharePoint Configuration API](#):

- SharePoint Versión—Especifique el SharePoint versión que usa al configurar SharePoint. Este es el caso sin importar si usa SharePoint Servidor 2013, SharePoint Servidor 2016, SharePoint Server 2019, o SharePoint En línea.

- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación que creó en suSharePointCuenta. El secreto se almacena en una estructura JSON.

Para SharePoint Autenticación básica en línea, la siguiente es la estructura JSON mínima que debe estar en su secreto:

```
{  
    "username": "user name",  
    "password": "password"  
}
```

Para SharePoint Autenticación OAuth 2.0 en línea, la siguiente es la estructura JSON mínima que debe estar en su secreto:

```
{  
    "username": "SharePoint account user name",  
    "password": "SharePoint account password",  
    "clientId": "SharePoint auto-generated unique client id",  
    "clientSecret": "secret string shared by Amazon Kendra and SharePoint to  
    authorize communications"  
}
```

Para SharePoint Autenticación básica del servidor, la siguiente es la estructura JSON mínima que debe estar en su secreto:

```
{  
    "username": "user name",  
    "password": "password",  
    "domain": "server domain name"  
}
```

Para SharePoint Autenticación LDAP del servidor(si necesita convertir su lista de control de acceso (ACL) a formato de correo electrónico para filtrar según el contexto del usuario, puede incluir la URL del servidor LDAP y la base de búsqueda de LDAP en su secreto), la siguiente es la estructura JSON mínima que debe contener su secreto:

```
{  
    "username": "user name",  
    "password": "password",  
    "domain": "server domain name",  
    "ldapServerUrl": "ldap://example.com:389",  
    "ldapSearchBase": "dc=example,dc=com"  
}
```

Para SharePoint Autenticación manual del servidor, la siguiente es la estructura JSON mínima que debe estar en su secreto:

```
{  
    "username": "user name",  
    "password": "password",  
    "domain": "server domain name",  
    "emailDomainOverride": "example.com"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSourcepara proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para elSharePointconector yAmazon Kendra. Para obtener más información, consulte[IAMfunciones paraSharePointfuentes de datos](#).
- Amazon VPC—Si usaSharePointServidor, especifiqueVpcConfigurationcomo parte de la configuración de la fuente de datos. Consulte[ConfigurandoAmazon Kendrausar una VPC](#).

También puede añadir las siguientes funciones opcionales:

- Proxy web—Si desea conectarse a suSharePointURL del sitio a través de un proxy web. Puede utilizar esta opción solo paraSharePointServidor.
- Listas de indexación—SiAmazon Kendradebe indexar el contenido de los archivos adjuntos aSharePointelementos de la lista.
- Registro de cambios—SiAmazon Kendradebe utilizar elSharePointmecanismo de registro de cambios en la fuente de datos para determinar si un documento debe agregarse, actualizarse o eliminarse en el índice.

Note

Usa el registro de cambios si no quieresAmazon Kendrapara escanear todos los documentos. Si el registro de cambios es grande, podría tardarAmazon Kendramenos tiempo para escanear los documentos en elSharePointfuente de datos que procesar el registro de cambios. Si estás sincronizando tuSharePointfuente de datos con su índice por primera vez, se escanean todos los documentos.

- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir cierto contenido.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elija mapear suSharePointcampos de fuente de datos para suAmazon Kendracampos de índice. Para obtener más información, consulte[Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendrarastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte[Filtrado de contexto de usuario paraSharePointfuentes de datos](#).

Más información

Para obtener más información sobre la integraciónAmazon Kendracon tuSharePointfuente de datos, consulte:

- [Empezar con elAmazon Kendra SharePointConector en línea](#)

SharePointconector V2.0

SharePointes un servicio colaborativo de creación de sitios web que puede utilizar para personalizar el contenido web y crear páginas, sitios, bibliotecas de documentos y listas. Puedes usarAmazon Kendrapara indexar susSharePointfuente de datos.

Amazon Kendraactualmente es compatibleSharePointNube ySharePointServidor (2013, 2016, 2019 y edición de suscripción).

Note

Soporte paraSharePointconector V1.0/SharePointConfigurationEstá previsto que la API finalice en 2023. Recomendamos migrar a o usarSharePointconector V2.0/TemplateConfigurationAPI.

Para solucionar problemas deAmazon Kendra SharePointconector de fuente de datos, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 401\)](#)
- [Requisitos previos \(p. 401\)](#)
- [Instrucciones de conexión \(p. 404\)](#)
- [Notas \(p. 414\)](#)

Características admitidas

Amazon Kendra SharePointel conector de fuente de datos admite las siguientes funciones:

- Mapeos de campo
- Rastreo de identidad
- Filtrado de contexto de usuario
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Patrones de inclusión/exclusión
- Sincronización completa/Sincronización de contenido nuevo o modificado/Sincronización de contenido nuevo, modificado o eliminado

Requisitos previos

Antes de poder usarAmazon Kendrapara indexar susSharePointfuente de datos, realice estos cambios en suSharePointyAWS cuentas.

EnSharePointEn línea, asegúrate de tener:

- Copié tuSharePointURL de instancia. El formato de la URL de host que introduzca es`https://yourdomain.sharepoint.com/sites/mysite`. La URL debe empezar porhttpsy contienensharepoint.com.
- Copió el nombre de dominio de suSharePointURL de instancia.
- Tomó nota de sus credenciales de autenticación básicas, que contienen el nombre de usuario y la contraseña con permisos de administrador del sitio para conectarse y conectarseSharePointEn línea.
- Si utiliza un tipo de autenticación distinto de la autenticación básica:Copió el ID de inquilino de suSharePointinstancia. Para obtener más información sobre cómo encontrar su ID de inquilino, consulte[Encuentra tu ID de inquilino de Microsoft 365](#).
- Para la autenticación OAuth 2.0:Tomó nota de suAutenticación básica credenciales que contienen el nombre de usuario y la contraseña que utiliza para conectarseSharePointEn línea y el ID de cliente y el secreto del cliente generados después del registroSharePointcon Azure AD.

- Si no está utilizando ACL, agregó los siguientes permisos:

Microsoft Graph	SharePoint
<ul style="list-style-type: none">• Notes.Read.All (Aplicación): lee todoOneNotecuadernos• Sites.Read.All (Aplicación): lee los elementos de todas las colecciones de sitios	<ul style="list-style-type: none">• AllSites.Leer (delegado): lee los elementos de todas las colecciones de sitios
Note	

Note.Read.All y Sites.Read.All solo son necesarios si desea rastrearOneNoteDocumentos.

- Si está utilizando ACL, agregó los siguientes permisos:

Microsoft Graph	SharePoint
<ul style="list-style-type: none">• Group.Member.Read.All (Aplicación): lee todas las membresías de grupos• Notes.Read.All (Aplicación): lee todoOneNotecuadernos• Sitios.FullControl.Todos (delegados): tienen el control total de todas las colecciones de sitios• Sites.Read.All (Aplicación): lee los elementos de todas las colecciones de sitios• User.Read.All (Aplicación): lee todos los perfiles de los usuarios	<ul style="list-style-type: none">• AllSites.Leer (delegado): lee los elementos de todas las colecciones de sitios

Note

GroupMember.Read.All y User.Read.All son obligatorios solo si Rastreador de identidades está activado.

- Para la autenticación solo con la aplicación Azure AD: Tomó nota del certificado X.509, la clave privada y el ID de cliente que generó después de registrarse SharePoint con Azure AD.
- Si no está utilizando ACL, agregó los siguientes permisos:

SharePoint
<ul style="list-style-type: none">• AllSites.Leer (delegado): lee los elementos de todas las colecciones de sitios• Sites.Manage.All (Aplicación): lee y escribe elementos y listas en todas las colecciones de sitios

- Si está utilizando ACL, agregó los siguientes permisos:

SharePoint
<ul style="list-style-type: none">• AllSites.Leer (delegado): lee los elementos de todas las colecciones de sitios

SharePoint

- Sitios.FullControl.All (Aplicación): tenga el control total de todas las colecciones de sitios
- Sites.Manage.All (Aplicación): lee y escribe elementos y listas en todas las colecciones de sitios
- Para SharePoint Autenticación solo para aplicaciones: Tomó nota de su SharePoint ID de cliente y secreto de cliente generados al conceder el permiso a SharePoint Solo la aplicación, y tu ID de cliente y tu secreto de cliente se generaron cuando registraste tu SharePoint aplicación con Azure AD.

Note

SharePoint La autenticación solo para aplicaciones es compatible con SharePoint Versión 2013.

- (Opcional) Si estás gateando OneNote documentos y uso Rastreador de identidades, agregó los siguientes permisos:

Microsoft Graph

- GroupMember.Read.All (Aplicación): lee todas las membresías de grupos
- Notes.Read.All (Aplicación): lee todo OneNote cuadernos
- Sites.Read.All (Aplicación): lee los elementos de todas las colecciones de sitios
- User.Read.All (Aplicación): lee todos los perfiles de los usuarios

Note

No se requieren permisos de API para rastrear entidades que utilizan Autenticación básica y SharePoint Autenticación solo mediante aplicaciones.

En SharePoint Servidor, asegúrese de tener:

- Copió tu SharePoint las URL de la instancia y el nombre de dominio de su SharePoint URL. El formato de la URL de host que introduzca es <https://yourcompany/sites/mysite>. La URL debe empezar por https.

Note

(En las instalaciones o en el servidor) Amazon Kendra comprueba si la información del punto final está incluida en AWS Secrets Manager es la misma que la información de punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a proteger contra la [problema de diputado confuso](#), que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, sino que usa Amazon Kendra como proxy para acceder al secreto configurado y realizar la acción. Si posteriormente cambias la información de tu terminal, debes crear un nuevo secreto para sincronizar esta información.

- Si usa SharePoint Autenticación solo mediante aplicaciones para el control de acceso:
 - Copió el SharePoint ID de cliente generado al registrar App Only a nivel de sitio. El formato ClientID es ClientID@TenantId. Por ejemplo, <ffa956f3-8f89-44e7-b0e4-49670756342c @888d0b57-69f1-4fb8-957f-e1f0bedf82fe>.
 - Copió el SharePoint secreto de cliente generado al registrar App Only a nivel de sitio.

Nota: Porque los ID de cliente y los secretos de cliente se generan para sitios individuales solo cuando te registras SharePoint Autenticación de servidor solo para aplicaciones, solo se admite la URL de un sitio SharePoint Autenticación solo en aplicaciones.

Note

SharePoint La autenticación solo para aplicaciones es no compatible con SharePoint Versión 2013.

- Si usa el ID de correo electrónico con dominio personalizado para el control de acceso:
 - Tomó nota del valor de su dominio de correo electrónico personalizado, por ejemplo: «[amazon.com](#)».
- Si usa el ID de correo electrónico con dominio de IDP autorización, copió su:
 - Punto final del servidor LDAP (punto final del servidor LDAP, incluidos el protocolo y el número de puerto). Por ejemplo: `ldap://example.com:389`.
 - Base de búsqueda de LDAP (base de búsqueda del usuario de LDAP). Por ejemplo: `CN=Usuarios, DC=SharePoint, DC=com`.
 - Nombre de usuario de LDAP y contraseña de LDAP.
- Cualquiera de las dos credenciales de autenticación NTLM configuradas o credenciales de autenticación Kerberos configuradas que contienen un nombre de usuario (SharePoint cuenta (nombre de usuario) y contraseña (SharePoint contraseña de la cuenta)).

En tu cuenta de AWS, asegúrate de tener:

- [Creó un Amazon Kendra índice](#), si usa la API, anota el ID del índice.
- [Creó un IAM papel](#) para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al correcto AWS Secrets Manager identificación secreta.

- Almacenó su SharePoint credenciales de autenticación en un AWS Secrets Manager secreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendado reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o secreto, puedes usar la consola para crear un nuevo IAM rol y Secrets Manager secreto cuando conectas tu SharePoint fuente de datos para Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existente IAM rol y Secrets Manager secreto y un identificador de índice.

Instrucciones de conexión

Para conectar Amazon Kendra a tu SharePoint fuente de datos, debe proporcionar detalles de sus SharePoint credenciales para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado SharePoint por Amazon Kendra vea [Requisitos previos \(p. 401\)](#).

Console: SharePoint Online

Para conectar Amazon Kendra a SharePoint En línea

1. Inicia sesión en la [AWS Consola de administración](#) y abra la [Amazon Kendra consola](#).
2. En el panel de navegación de la izquierda, selecciona **Índices**, y, a continuación, elija el índice que desee utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeSharePointconector V2.0y, a continuación, elijaAregar fuente de datos.
5. En elEspecificar los detalles de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:
 - a. EnFuente, paraMétodo de alojamiento—EligeSharePointEn línea.
 - b. URL del sitio específicas para suSharePointrepositorio—Introduzca elSharePointURL de hospedaje. El formato de las URL de host que introduzcas es`https://yourdomain.sharepoint.com/sites/mysite`. La URL debe empezar porhttpsprotocolo. Separa las URL con una línea nueva. Puedes añadir hasta 100 URL.
 - c. Dominio—Introduzca elSharePointdominio. Por ejemplo, el dominio de la URL`https://yourdomain.sharepoint.com/sites/mysiteestu dominio`.
 - d. ParaAutorización, puede optar por utilizar una lista de control de acceso (ACL) para controlar los resultados de la búsqueda en función del nivel de acceso a los documentos del usuario final en suSharePointFuente de datos en línea. La autorización mediante ACL está activada de forma predeterminada.
 - e. Paraautenticación, elige entreBásico,Oauth 2.0,Autenticación exclusiva de la aplicación Azure AD, ySharePointAutenticación solo mediante aplicacionessegún su caso de uso.
 - i. Si usaAutenticación básica, introduzca la siguiente información:
 - ParaAWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para guardar tusSharePointcredenciales de autenticación. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta. Introduzca la siguiente información en la ventana:
 - Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-SharePoint-' se añade automáticamente a tu nombre secreto.
 - Nombre de usuario—Nombre de usuario para suSharePointcuenta.
 - Contraseña—Contraseña para suSharePointcuenta.
 - ii. Si usaAutenticación OAuth 2.0, introduzca la siguiente información:
 - ID de inquilino—ID de inquilino de suSharePointcuenta.
 - ParaAWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para guardar tusSharePointcredenciales de autenticación. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta. Introduzca la siguiente información en la ventana:
 - Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-SharePoint-' se añade automáticamente a tu nombre secreto.

- Nombre de usuario—Nombre de usuario para suSharePointcuenta.
 - Contraseña—Contraseña para suSharePointcuenta.
 - ID de cliente—El ID de cliente de Azure AD generado al registrarseSharePointen Azure AD.
 - Secreto de cliente—El secreto de cliente de Azure AD generado al registrarseSharePointen Azure AD.
- iii. Si usaAutenticación exclusiva de la aplicación Azure AD, introduzca la siguiente información:
- ID de inquilino—ID de inquilino de suSharePointcuenta.
 - Certificado X.509 autofirmado de Azure AD—Certificado para autenticar el conector de Azure AD.
 - ParaAWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para guardar tusSharePointcredenciales de autenticación. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta. Introduzca la siguiente información en la ventana:
 - Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-SharePoint-' se añade automáticamente a tu nombre secreto.
 - ID de cliente—El ID de cliente de Azure AD generado al registrarseSharePointen Azure AD.
 - Clave privada—Una clave privada para autenticar el conector de Azure AD.
- iv. Si usaSharePointAutenticación solo mediante aplicaciones, introduzca la siguiente información:
- ID de inquilino—ID de inquilino de suSharePointcuenta.
 - ParaAWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para guardar tusSharePointcredenciales de autenticación. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta. Introduzca la siguiente información en la ventana:
 - Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-SharePoint-' se añade automáticamente a tu nombre secreto.
 - SharePointID de cliente—LaSharePointID de cliente que generaste al registrar App Only a nivel de inquilino. El formato ClientID es *ID de cliente@TenantId*. Por ejemplo, **ffa956f3-8f89-44e7-b0e4-49670756342c @888d0b57-69f1-4fb8-957f-e1f0bedf82fe**.
 - SharePointsecreto de cliente—LaSharePointsecreto de cliente generado cuando se restablece para App Only a nivel de inquilino.
 - ID de cliente—El ID de cliente de Azure AD generado al registrarseSharePointen Azure AD.
 - Secreto de cliente—El secreto de cliente de Azure AD que se genera al restablecerSharePointa Azure AD.
- f. Rastreador de identidades—(Se activa solo cuando la ACL está habilitada) Elija activarAmazon Kendrarastreador de identidad para sincronizar la información de identidad. Si optas por desactivar el rastreador de identidades, debes cargar la información principal mediante el[PutPrincipalMappingAPI](#).

Note

El mapeo de grupos de Crawl AD está disponible para OAuth 2.0 ySharePointSolo la aplicación Autenticación solo para usuarios.

También puede optar por:

- i. Mapeo de grupos locales de40fawl—Activar para rastrear el mapeo de grupos locales.

- ii. Mapeo de grupos de Crawl AD—Activar para rastrear el mapeo de grupos de Azure Active Directory.
- g. (Opcional)Configurar VPC y grupo de seguridad—Seleccione una VPC para usarla con suSharePointinstancia. Si es así, debe añadirSubredesyGrupos de seguridad de VPC.
- h. IAMpapel—Elige uno existenteIAMrol o crea uno nuevolAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- i. Elija Siguiente.

7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:

- a. EnAlcance de sincronización, elija una de las siguientes opciones:
 - i. Seleccione entidades: elija las entidades que deseé rastrear. Puede seleccionar el rastreoTodosentidades o cualquier combinación deArchivos,Adjuntos,Enlaces Páginas,Eventos,Comentarios, yDatos de lista.
 - ii. EnConfiguración adicional, paraPatrones de expresiones regulares de entidades—Añada patrones de expresiones regulares paraEnlaces,Páginas, yEventospara incluir entidades específicas en lugar de sincronizar todos los documentos.
 - iii. Patrones Regex—Añada patrones de expresiones regulares para incluir o excluir archivos medianteRuta del archivo Nombre de archivo Tipo de archivo,OneNotenombre de sección, yOneNotenombre de páginaen lugar de sincronizar todos tus documentos. Puedes añadir hasta 100.
 - b. ParaModo de sincronizaciónelija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Cuando sincronizas tu fuente de datos conAmazon Kendrapor primera vez, todo el contenido se sincroniza de forma predeterminada.
 - Sincronización completa—Sincronice todo el contenido independientemente del estado de sincronización anterior.
 - Sincronización de documentos nuevos o modificados—Sincronizar solo documentos nuevos o modificados.
 - Sincronización de documentos nuevos, modificados o eliminados—Sincronice solo los documentos nuevos, modificados y eliminados.
 - c. EnCronograma de ejecución de sincronización, paraFrecuencia— Con qué frecuenciaAmazon Kendrase sincronizará con su fuente de datos.
 - d. Elija Siguiente.
8. En elDefinir mapeos de campospágina, introduzca la siguiente información:
- a. ParaEventos Páginas,Archivos,Enlaces,Adjuntos, yComentarios—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

Console: SharePoint Server

Para conectar Amazon Kendra a SharePoint

1. Inicia sesión en la AWS Consola de administración y abra el [Amazon Kendra consola](#).
2. En el panel de navegación de la izquierda, selecciona Índices y, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar su Control de acceso de usuarios ajustes en Ajustes de índice.

3. En el Primeros pasos página, elige Agregar fuente de datos.
 4. En el Agregar fuente de datos página, elige SharePoint conector V2.0 y, a continuación, elija Agregar fuente de datos.
 5. En el Especificar los detalles de la fuente de datos página, introduzca la siguiente información:
 - a. En Nombre y descripción, para Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional) Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, para idioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, para Añadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tus AWS costos.
 - e. Elija Siguiente.
 6. En el Defina el acceso y la seguridad página, introduzca la siguiente información:
 - a. En Fuente, para Método de alojamiento—Elige SharePoint Servidor.
 - b. Escoja SharePoint Versión—Elige entre SharePoint 2013, SharePoint 2016, SharePoint 2019, y SharePoint (Edición de suscripción).
 - c. URL del sitio específicas para su SharePoint repositorio—Introduzca el SharePoint URL de hospedaje. El formato de las URL de host que introduzcas es **https://yourcompany/sites/mysite**. La URL debe empezar por https protocolo. Separa las URL con una línea nueva. Puedes añadir hasta 100 URL.
 - d. Dominio—Introduzca el SharePoint dominio. Por ejemplo, el dominio de la URL **https://yourcompany/sites/mysite estu empresa**
 - e. Ubicación del certificado SSL—Introduzca el Amazon S3 ruta a su archivo de certificado SSL.
 - f. (Opcional) Para Proxy web—Introduzca el nombre del host (sin http:// o https:// protocolo) y el número de puerto utilizado por el protocolo de transporte URL del host. El valor numérico del número de puerto debe estar comprendido entre 0 y 65535.
 - g. Para Autorización—Puede optar por utilizar una lista de control de acceso (ACL) para controlar los resultados de la búsqueda en función del nivel de acceso a los documentos del usuario final en su SharePoint fuente de datos. La autorización mediante ACL está activada de forma predeterminada. Cuando la ACL está desactivada, no se rastrea ninguna información de la ACL y no hay control de acceso ni filtrado de contexto disponible.
- Para SharePoint Servidor que puede elegir entre las siguientes opciones de ACL:
- i. ID de correo electrónico con dominio de IDP—El control de acceso se basará en los identificadores de correo electrónico extraídos de los dominios de correo electrónico obtenidos del proveedor de identidad (IDP) subyacente. Usted proporciona los detalles de la conexión del IDP en su Secrets Manager secreto durante autenticación.
 - ii. ID de correo electrónico con dominio personalizado—El control de acceso se basará en las identificaciones de correo electrónico. Desea proporcionar el valor del dominio de

correo electrónico. Por ejemplo, «amazon.com». El dominio de correo electrónico se utilizará para crear la ID de correo electrónico para el control de acceso. Debe introducir su dominio de correo electrónico utilizandoAregar dominio de correo electrónico.

- iii. Dominio\ Usuario con dominio—El control de acceso se estructurará mediante un formato de dominio\ ID de usuario. Debes proporcionar un nombre de dominio válido. Por ejemplo:«sharepoint 2019»para construir el control de acceso.
- h. Paraautenticación, elige entreSharePointAutenticación solo mediante aplicaciones,Autenticación NTLM, yAutenticación Kerberossegún su caso de uso.
- i. Introduzca la siguiente información para ambosAutenticación NTLMyAutenticación Kerberos:

ParaAWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para guardar tusSharePointcredenciales de autenticación. Si eliges crear un secreto nuevo, unAWS Secrets Managerse abre una ventana secreta. Introduzca la siguiente información en la ventana:

- Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-SharePoint-' se añade automáticamente a tu nombre secreto.
- Nombre de usuario—Nombre de usuario para suSharePointcuenta.
- Contraseña—Contraseña para suSharePointcuenta.

Si usalD de correo electrónico con dominio de IDP, introduce también tu:

- Terminal de servidor LDAP—Punto final del servidor LDAP, incluidos el protocolo y el número de puerto. Por ejemplo:[ldap://example.com:389](http://example.com:389).
 - Base de búsqueda LDAP—Base de búsqueda del usuario de LDAP. Por ejemplo:CN=Usuarios, DC=SharePoint, DC=com.
 - Nombre de usuario de LDAP—Su nombre de usuario de LDAP.
 - Contraseña LDAP—Su contraseña de LDAP.
- ii. Introduzca la siguiente información paraSharePointAutenticación solo mediante aplicaciones.

ParaAWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para guardar tusSharePointcredenciales de autenticación. Si eliges crear un secreto nuevo, unAWS Secrets Managerse abre una ventana secreta. Introduzca la siguiente información en la ventana:

- Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-SharePoint-' se añade automáticamente a tu nombre secreto.
- ID de cliente—LaSharePointID de cliente que generaste al registrar App Only a nivel de sitio. El formato ClientID es ClientID@TenantId. Por ejemplo,ffa956f3-8f89-44e7-b0e4-49670756342c @888d0b57-69f1-4fb8-957f-e1f0bedf82fe.
- SharePointsecreto de cliente—LaSharePointsecreto de cliente generado cuando se restablece para App Only a nivel de sitio.

Nota:Porque los ID de cliente y los secretos de cliente se generan para sitios individuales solo cuando te registrasSharePointAutenticación de servidor solo para aplicaciones, solo se admite la URL de un sitioSharePointAutenticación solo en aplicaciones.

Si usalD de correo electrónico con dominio de IDP, introduce también tu:

- Terminal de servidor LDAP—Punto final del servidor LDAP, incluidos el protocolo y el número de puerto. Por ejemplo:`ldap://example.com:389`.
 - Base de búsqueda LDAP—Base de búsqueda del usuario de LDAP. Por ejemplo:`CN=Usuarios, DC=SharePoint, DC=com`.
 - Nombre de usuario de LDAP—Su nombre de usuario de LDAP.
 - Contraseña LDAP—Su contraseña de LDAP.
- i. Rastreador de identidades—(Se activa solo cuando la ACL está habilitada) Elija habilitar Amazon Kendra rastreador de identidad para sincronizar la información de identidad. Si optas por desactivar el rastreador de identidades, debes cargar la información principal mediante el [PutPrincipalMapping API](#). También puedes elegir:
- i. Mapeo de grupos locales de Crawl—Activar para rastrear el mapeo de grupos locales.
 - ii. (ParalD de correo electrónico con dominio de IDP únicamente) Mapeo de grupos de Crawl AD—Activar para rastrear el mapeo de Active Directory.
- j. (Opcional) Configurar VPC y grupo de seguridad—Seleccione una VPC para usarla con su SharePoint instancia. Si es así, debe añadir Subredes y Grupos de seguridad de VPC.
- k. IAM papel—Elige uno existente IAM rol o crea uno nuevo IAM rol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAM los roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

- i. Elija Siguiente.
7. En el Configurar los ajustes de sincronización página, introduzca la siguiente información:
- a. En Alcance de sincronización, elija una de las siguientes opciones:
 - i. Seleccione entidades: elija las entidades que deseé rastrear. Puede seleccionar el rastreoTodos entidades o cualquier combinación de Archivos, Adjuntos, Enlaces, Páginas, Eventos, y Datos de lista.
 - ii. En Configuración adicional, para Patrones de expresiones regulares de entidades—Añade patrones de expresiones regulares para Enlaces, Páginas, y Eventos para incluir entidades específicas en lugar de sincronizar todos los documentos.
 - iii. Patrones Regex—Añade patrones de expresiones regulares para incluir o excluir archivos mediante Ruta del archivo Nombre de archivo, Tipo de archivo, OneNote nombre de sección, y OneNote nombre de página en lugar de sincronizar todos tus documentos. Puedes añadir hasta 100.
 - b. Para Modo de sincronización elija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Cuando sincronizas tu fuente de datos con Amazon Kendra por primera vez, todo el contenido se sincroniza de forma predeterminada.
 - Sincronización completa—Sincronice todo el contenido independientemente del estado de sincronización anterior.
 - Sincronización de documentos nuevos o modificados—Sincronice solo los documentos nuevos y modificados.
 - Sincronización de documentos nuevos, modificados o eliminados—Sincronice solo los documentos nuevos, modificados y eliminados.
 - c. En Cronograma de ejecución de sincronización, para Frecuencia—Con qué frecuencia Amazon Kendra se sincronizará con su fuente de datos.
 - d. Elija Siguiente.
8. En el Definir mapeos de campos página, introduzca la siguiente información:

- a. ParaEventos Páginas,Archivos,Enlaces,Adjuntos, yDatos de lista—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon KendraaSharePoint

Debe especificar un JSON de[esquema de fuente de datos](#)utilizando el[TemplateConfigurationAPI](#).
Debe proporcionar la siguiente información:

- Metadatos de punto final del repositorio—Especifique eltenantID domainysiteUrlsde tuSharePointinstancia.
- Propiedades adicionales del repositorio—Especifique:
 - (ParaSharePointSolo servidor)s3bucketNamey3certificateNameque utilizó para almacenar su certificado X.509 autofirmado de Azure AD.
 - Tipo de autenticación (auth_Type) que está utilizando, ya seaOAuth2,OAuth2App,OAuth2Certificate,Basic.
 - Versión (version) que está utilizando, ya seaServeroOnline. Si usaServerpuede especificar con más detalle elonPremVersiontan2013,2016,2019, oSubscriptionEdition.
- Teclea—EspecificarTEMPLATEcomo el tipo cuando llamasCreateDataSource.
- Fuente de datos—Especifique la fuente de datos comoSHAREPOINTV2.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación que creó en suSharePointcuenta.

Si usaSharePointEn línea, puede elegir entre Basic, OAuth 2.0, Azure AD solo para aplicaciones ySharePointAutenticación solo en aplicaciones. La siguiente es la estructura JSON mínima que debe estar en su secreto para cada opción de autenticación:

- Autenticación básica

```
{  
    "username": "SharePoint account user name",  
    "password": "SharePoint account password"  
}
```

- Autenticación OAuth 2.0

```
{  
    "clientId": "client id generated when registering SharePoint with Azure AD",  
    "clientSecret": "client secret generated when registering SharePoint with Azure AD",  
    "userName": "SharePoint account user name",  
    "password": "SharePoint account password"  
}
```

- Autenticación exclusiva de la aplicación Azure AD

```
{  
    "clientId": "client id generated when registering SharePoint with Azure AD",  
    "privateKey": "private key to authorize connection with Azure AD"  
}
```

- SharePointAutenticación solo mediante aplicaciones

```
{  
    "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"  
}
```

Si usaSharePointServidor, puedes elegir entreSharePointAutenticación solo para aplicaciones, autenticación NTLM y autenticación Kerberos. La siguiente es la estructura JSON mínima que debe estar en su secreto para cada opción de autenticación:

- SharePointAutenticación solo mediante aplicaciones

```
{  
    "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",  
}
```

- SharePointAutenticación solo para aplicaciones con autorización de dominio procedente del IDP

```
{  
    "siteUrlsHash": "Hash representation of SharePoint site URLs",  
    "clientId": "client id generated when registering SharePoint for App Only at  
    Site Level",  
    "clientSecret": "client secret generated when registering SharePoint for App  
    Only at Site Level",  
    "ldapUrl": "LDAP Account url eg. ldap://example.com:389",  
    "baseDn": "LDAP Account base dn eg. CN=Users,DC=sharepoint,DC=com",  
    "ldapUser": "LDAP account user name",  
    "ldapPassword": "LDAP account password"  
}
```

- (Solo servidor) Autenticación NTLM o Kerberos

```
{  
    "siteUrlsHash": "Hash representation of SharePoint site URLs",  
    "username": "SharePoint account user name",  
    "password": "SharePoint account password"  
}
```

- (Solo servidor) Autenticación NTLM o Kerberos con autorización de dominio desde IDP

```
{  
    "siteUrlsHash": "Hash representation of SharePoint site URLs",  
    "userName": "SharePoint account user name",  
    "password": "SharePoint account password",  
    "ldapUrl": "ldap://example.com:389",  
}
```

```
    "baseDn": "CN=Users,DC=sharepoint,DC=com",
    "ldapUser": "LDAP account user name",
    "ldapPassword": "LDAP account password"
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- Modo de sincronización—Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:
 - FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice
 - FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice
 - CHANGE_LOG para rastrear de forma incremental solo el contenido nuevo y modificado cada vez que la fuente de datos se sincronice con el índice
- Habilitar el rastreador de identidades: especifique si desea activar el rastreador de identidades. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el [PutPrincipalMapping API](#). Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado en contexto de usuario](#).

Note

El rastreo de identidades solo está disponible cuando configuras `crawlAcl=true`.

- IAM papel—Especificare `RoleArn` cuando llamas `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el SharePoint conector y Amazon Kendra. Para obtener más información, consulte [IAM funciones para SharePoint fuentes de datos](#).

También puede añadir las siguientes funciones opcionales:

- Documentos específicos para indexar—Puede utilizar un SharePoint consulta para especificar los documentos que desea de una o más bases de conocimiento, incluidas las bases de conocimiento privadas. El acceso a las bases de conocimiento lo determina el usuario que utilice para conectarse a la SharePoint instancia. Para obtener más información, consulte [Specifying documents to index with a query](#) (Especificar documentos a indexar con una consulta).
- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir los artículos de conocimiento, los catálogos de servicios, los incidentes y sus anexos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elija mapear su SharePoint campos de fuente de datos para su Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).

- Nube privada virtual (VPC)—Especificar `VpcConfiguration` cuando llamas `CreateDataSource`. Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).
- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para SharePoint fuentes de datos](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [Microsoft SharePoint esquema de plantilla \(p. 206\)](#).

Notas

- El conector solo admite mapeos de campos personalizados para Archivosentidad.
- Para todos SharePoint Versiones de servidor, el token de ACL debe estar en minúsculas. Para Correo electrónico con dominio de IDPyID de correo electrónico con dominio personalizado ACL, por ejemplo:`user@sharepoint2019.com`. Para Dominio\ Usuario con dominio ACL, por ejemplo:`sharepoint2013\ usuario`.
- El conector no admite el modo de registro de cambios/Sincronización de contenido nuevo o modificado por SharePoint 2013.
- Si el nombre de una entidad tiene un % carácter en su nombre, el conector omitirá estos archivos debido a las limitaciones de la API.
- OneNotesolo se puede rastrear mediante el conector mediante un ID de inquilino y con OAuth 2.0 o SharePoint Autenticación solo en aplicaciones activada para SharePoint En línea.
- El conector rastrea la primera sección de un OneNote documento utilizando únicamente su nombre predeterminado, incluso si se cambia el nombre del documento.
- El conector rastrea los enlaces en SharePoint 2019, SharePoint Edición en línea y por suscripción, solo si Páginas y Archivos se seleccionan como entidades que se van a rastrear, además de Enlaces.
- El conector rastrea los enlaces en SharePoint 2013 y SharePoint 2016 si Enlaces se selecciona como una entidad que se va a rastrear.
- El conector rastrea los archivos adjuntos de la lista y los comentarios solo cuando Datos de lista también se selecciona como entidad que se va a rastrear.
- El conector rastrea los archivos adjuntos de eventos solo cuando Eventos también se selecciona como entidad que se va a rastrear.

Equipos de Microsoft

Microsoft Teams es una herramienta de colaboración empresarial para mensajería, reuniones e intercambio de archivos. Si eres usuario de Microsoft Teams, puedes usar Amazon Kendra para indexar su fuente de datos de Microsoft Teams.

Puedes conectarte a Amazon Kendra a su fuente de datos de Microsoft Teams mediante [Amazon Kendra console](#) y el [Template Configuration API](#).

Para solucionar problemas de Amazon Kendra Conector de fuentes de datos de Microsoft Teams, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 415\)](#)
- [Requisitos previos \(p. 415\)](#)
- [Instrucciones de conexión \(p. 417\)](#)

- [Más información \(p. 420\)](#)

Características admitidas

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Rastreador de identidades
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados

Requisitos previos

Antes de poder usar Amazon Kendra para indexar su fuente de datos de Microsoft Teams, realice estos cambios en su Microsoft Teams y AWS cuentas.

En Microsoft Teams, asegúrate de tener:

- Creé una cuenta de Microsoft Teams en Office 365.
- Tomó nota de su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
- Creó una aplicación OAuth en el portal de Azure y anotó el ID de inquilino, el ID del cliente y el secreto del cliente o las credenciales del cliente. Consulte [Tutorial de Microsoft](#) para obtener más información.
- Se agregaron los permisos necesarios. Puedes añadir todos los permisos o limitar el alcance seleccionando menos permisos en función de las entidades que quieras que se rastreen. A continuación se muestra la tabla de permisos por entidad correspondiente:

Entidad	Permisos necesarios para la sincronización de datos	Permisos necesarios para la sincronización de identidades
Publicación del canal	<ul style="list-style-type: none">• ChannelMessage.Leer.Todo• Agrupar. Leer todo• Usuario.Leer• Usuario.Leed.All	TeamMember.Leer.Todo
Conexión al canal	<ul style="list-style-type: none">• ChannelMessage.Leer.Todo• Agrupar. Leer todo• Usuario.Leer• Usuario.Leed.All	TeamMember.Lee.Todo
Canal Wiki	<ul style="list-style-type: none">• Agrupar. Leer todo• Usuario.Leer• Usuario.Leed.All	TeamMember.Leer.Todo
Mensaje de chat	<ul style="list-style-type: none">• Chatea, lee. Todo• ChatMessage.Leer.Todo• ChatMember.Lee.Todo• Usuario.Leer• Usuario.Leed.All• Agrupar. Leer todo	TeamMember.Lee.Todo

Entidad	Permisos necesarios para la sincronización de datos	Permisos necesarios para la sincronización de identidades
Chat de reunión	<ul style="list-style-type: none"> • Chatea, lee. Todo • ChatMessage.Leer • ChatMember.Lee.Todo • Usuario.Leer • Usuario.Leed.All • Agrupar. Leerlo todo 	TeamMember.Lee.Todo
Archivo adjunto de chat	<ul style="list-style-type: none"> • Chatea, lee. Todo • ChatMessage.Leer • ChatMember.Lee.Todo • Usuario.Leer • Usuario.Leed.All • Agrupar. Leerlo todo 	TeamMember.Lee.Todo
Archivo de reunión	<ul style="list-style-type: none"> • Chatea, lee. Todo • ChatMessage.Lee.Todo • ChatMember.Lee.Todo • Usuario.Leer • Usuario.Leed.All • Agrupar. Leerlo todo • Archivos.Leer.Todos 	TeamMember.Lee.Todo
Reunión del calendario	<ul style="list-style-type: none"> • Chatea, lee. Todo • ChatMessage.Lee.Todo • ChatMember.Lee.Todo • Usuario.Leer • Usuario.Leed.All • Agrupar. Leerlo todo • Archivos.Leer.Todos 	TeamMember.Lee.Todo
Notas de la reunión	<ul style="list-style-type: none"> • Usuario.Leer • Usuario.Leed.All • Agrupar. Leerlo todo • Archivos.Leer.Todos 	TeamMember.Lee.Todo

- Se comprobó que cada documento es único en Microsoft Teams y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellIAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulIAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Has guardado tus credenciales de autenticación de Microsoft Teams en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Microsoft Teams aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa su fuente de datos de Microsoft Teams, debe proporcionar los detalles necesarios de su fuente de datos de Microsoft Teams para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configurado Microsoft Teams paraAmazon Kendra, consulte[Requisitos previos \(p. 415\)](#).

Console

Para conectarAmazon Kendraa Microsoft Teams

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConejor de Microsoft Teamsy, a continuación, elijaAregar fuente de datos.
5. En elEspecificardetalle de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:
 - a. Fuente—Introduzca su ID de inquilino de Microsoft 365. Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
 - b. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar sus credenciales de autenticación de Microsoft Teams. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - i. Introduzca la siguiente información enCrea unAWS Secrets Managerventana secreta:

- A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Microsoft Teams-' se añade automáticamente a tu nombre secreto.
- B. ParáD de cliente y Secreto de cliente—Introduzca los valores de las credenciales de autenticación que generó en su cuenta de Microsoft Teams en el portal de Azure.
 - ii. Seleccione Guardar.
- c. Modelo de pago—Puede elegir un modelo de licencia y pago para su cuenta de Microsoft Teams. Los modelos de pago del modelo A están restringidos a los modelos de licencias y pago que requieren el cumplimiento de las normas de seguridad. Los modelos de pago del modelo B son adecuados para los modelos de licencias y pago que no requieren el cumplimiento de las normas de seguridad.
- d. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadir Subredes y Grupos de seguridad de VPC.
- e. Rastreador de identidades—Cuando el rastreador de identidades está activado, Amazon Kendra sincroniza la información de identidad. Si optas por desactivar el rastreador de identidades, debes cargar la información principal mediante el [PutPrincipalMapping API](#).
- f. IAMpapel—Elige uno existente IAMrol o crea uno nuevo IAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

- g. Elija Siguiente.
7. En el Configurar los ajustes de sincronización página, introduzca la siguiente información:
 - a. Sincronizar contenido—Seleccione el contenido que desea sincronizar.
 - b. Configuración adicional—Si lo desea, puede utilizar esta configuración para indexar cierto contenido en lugar de sincronizar todos los documentos.
 - c. Modo de sincronización—Puede elegir cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos.
 - i. Si eliges la sincronización completa, Amazon Kendra sincronizará todo el contenido de todas las entidades, independientemente del estado de sincronización anterior.
 - ii. Si eliges la sincronización de contenido nuevo o modificado, Amazon Kendra solo sincronizará contenido nuevo o modificado.
 - iii. Si eliges la sincronización de contenido nuevo, modificado o eliminado, Amazon Kendra solo sincronizará contenido nuevo, modificado o eliminado.
8. En el Definir mapeos de campos página, introduzca la siguiente información:
 - a. Campos de fuente de datos predeterminados—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En el Revisar y crear página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en Fuentes de datos página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Microsoft Teams

Debe especificar un JSON de [esquema de fuente de datos](#) utilizando el [TemplateConfiguration API](#).
Debe proporcionar la siguiente información:

- Fuente de datos—Debe especificar la fuente de datos como MSTEAMS.
- ID de inquilino—Puede encontrar su ID de inquilino en las propiedades de su portal de Azure Active Directory o en la aplicación OAuth.
- Teclea—Especificar TEMPLATE como el tipo cuando llamas `CreateDataSource`.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contiene las credenciales de autenticación de su cuenta de Microsoft Teams. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "clientId": "client ID",  
    "clientSecret": "client secret"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendado reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAM papel—Especificar `RoleArn` cuando llamas `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Microsoft Teams y Amazon Kendra. Para obtener más información, consulte [IAM funciones para las fuentes de datos de Microsoft Teams](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—Especificar `VpcConfiguration` cuando llamas `CreateDataSource`. Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).
- Habilitar el rastreador de identidades: especifique si desea activar el rastreador de identidades. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el [PutPrincipalMapping API](#). Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado en contexto de usuario](#).
- Filtros de inclusión y exclusión—Especifique si desea incluir o excluir cierto contenido en Microsoft Teams.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coinciden con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coinciden con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elija asignar los campos de la fuente de datos de Microsoft Teams a su Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields \(Asignación de campos de origen de datos\)](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [Esquema de plantillas de Microsoft Teams \(p. 217\)](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Microsoft Teams, consulte:

- [Busque de forma inteligente en la fuente de datos de Microsoft Teams de su organización con Amazon Kendra conector para Microsoft Teams](#)

Microsoft Yammer

Microsoft Yammer es una herramienta de colaboración empresarial para mensajería, reuniones y uso compartido de archivos. Si es usuario de Microsoft Yammer, puede usar Amazon Kendra para indexar la fuente de datos de Microsoft Yammer.

Puedes conectarte a su fuente de datos de Microsoft Yammer mediante [Amazon Kendra console](#) o el [Template Configuration API](#).

Para solucionar problemas de Amazon Kendra Conector de fuente de datos de Microsoft Yammer, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Características admitidas

- Change log
- Mapeos de campo
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de poder usar Amazon Kendra para indexar su fuente de datos de Microsoft Yammer, realice estos cambios en su Microsoft Yammer y AWS cuentas.

En Microsoft Yammer, asegúrate de tener:

- Creó una cuenta administrativa de Microsoft Yammer.
- Apuntó su nombre de usuario y contraseña de Microsoft Yammer.
- Creó una aplicación OAuth en el portal de Azure y anotó el ID del cliente y el secreto del cliente o las credenciales del cliente. Consulte [Tutorial de Microsoft](#) para obtener más información.
- Se ha comprobado que cada documento es único en Microsoft Yammer y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu cuenta de AWS, asegúrate de tener:

- [Creó un Amazon Kendra índice](#), si usa la API, anota el ID del índice.
- [Creó un IAM papel](#) para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al correcto AWS Secrets Manager identificación secreta.

- Almacenó sus credenciales de autenticación de Microsoft Yammer en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Microsoft Yammer aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa su fuente de datos de Microsoft Yammer, debe proporcionar los detalles necesarios de su fuente de datos de Microsoft Yammer para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configurado Microsoft Yammer paraAmazon Kendra, consulte[Requisitos previos \(p. 420\)](#).

Console

Para conectarAmazon Kendraa Microsoft Yammer

- Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
- En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que desee utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

- En elPrimeros pasospágina, eligeAgregar fuente de datos.
- En elAgregar fuente de datospágina, eligeConector Microsoft Yammy, a continuación, elijaAgregar fuente de datos.
- En elEspecificardetalle de la fuente de datospágina, introduzca la siguiente información:
 - Nombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - Elija Siguiente.
- En elDefinir acceso y la seguridadpágina, introduzca la siguiente información:
 - Fuente—Utilice su URL de Microsoft Yammer.
 - AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar sus credenciales de autenticación de Microsoft Yammer. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - Introduzca la siguiente información enCrea unAWS Secrets Managerventana secreta:

- A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Microsoft Yammer-' se añade automáticamente a tu nombre secreto.
- B. ParaNombre de usuario,Contraseña—Introduzca su nombre de usuario y contraseña de Microsoft Yammer.
- C. ParalD de cliente,Secreto de cliente—Introduzca los valores de las credenciales de autenticación que generó desde su cuenta de Microsoft Yammer en el portal de Azure.
 - ii. Seleccione Guardar.
- c. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesGrupos de seguridad de VPC.
- d. Rastreador de identidades—Elija rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos y guardarla enAmazon Kendratienda principal/de identidad. Esto resulta útil para filtrar el contexto del usuario, donde los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos.
- e. IAMpapel—Elige uno existenteIAMrol o crea uno nuevoIAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- f. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. Desde la fecha—Especifique la fecha para empezar a rastrear los datos en Microsoft Yammer.
 - b. Sincronizar contenido—Seleccione el tipo de contenido que desea indexar. Por ejemplo, mensajes públicos, mensajes privados y archivos adjuntos.
 - c. Configuración adicional—Si lo desea, puede utilizar estas opciones para indexar cierto contenido en lugar de sincronizar todos los documentos. Por ejemplo, puede indexar nombres de comunidades específicas y usar patrones de expresiones regulares para incluir o excluir ciertos archivos.
 - d. Modo de sincronización—Puede elegir cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos.
 - i. Si eliges la sincronización completa,Amazon Kendrasincronizará todo el contenido de todas las entidades, independientemente del estado de sincronización anterior.
 - ii. Si eliges la sincronización de contenido nuevo o modificado,Amazon Kendrasolo sincronizará contenido nuevo o modificado.
 - iii. Si eliges la sincronización de contenido nuevo, modificado o eliminado,Amazon Kendrasolo sincronizará contenido nuevo, modificado o eliminado.
8. En elDefinir mapeos de campospágina, introduzca la siguiente información:
 - a. Campos de fuente de datos predeterminados—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Microsoft Yammer

Debe especificar un JSON de [esquema de fuente de datos](#) utilizando el [TemplateConfiguration API](#). Debe proporcionar la siguiente información:

- Fuente de datos—Debe especificar la fuente de datos como YAMMER.
- Escriba—Especificando TEMPLATE como el tipo cuando llamas `CreateDataSource`.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contiene las credenciales de autenticación de su cuenta de Microsoft Yammer. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "user name",  
    "password": "password",  
    "clientId": "client ID",  
    "clientSecret": "client secret"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendado reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAM papel—Especificando RoleArn cuando llamas `CreateDataSource` para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Microsoft Yammer y Amazon Kendra. Para obtener más información, consulte [IAM funciones para las fuentes de datos de Microsoft Yammer](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—Especificando VpcConfiguration cuando llamas `CreateDataSource`. Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).
- Rastreador de identidades—Especifica si se va a activar Amazon Kendra rastreador de identidad. Si el rastreador de identidades está desactivado, debe cargar la información de identidad o principal mediante el [PutPrincipalMapping API](#). Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario](#).
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir cierto contenido.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coinciden con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coinciden con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elige asignar los campos de la fuente de datos de Microsoft Yammer a su Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields \(Asignación de campos de origen de datos\)](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Microsoft Yammer, consulte:

- [Anunciamos el conector Yammer para Amazon Kendra](#)

Quip

Quip es un software de productividad colaborativa que ofrece capacidades de creación de documentos en tiempo real. Puedes usar Amazon Kendra para indexar sus carpetas, archivos, comentarios de archivos, salas de chat y archivos adjuntos de Quip.

Puedes conectarte a Amazon Kendra a su fuente de datos de Quip mediante el [Amazon Kendra console](#) y el [Quip Configuration API](#).

Para solucionar problemas de Amazon Kendra Conector de fuente de datos Quip, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 424\)](#)
- [Requisitos previos \(p. 424\)](#)
- [Instrucciones de conexión \(p. 425\)](#)
- [Más información \(p. 427\)](#)

Características admitidas

Amazon Kendra El conector de fuente de datos Quip admite las siguientes funciones:

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de poder usar Amazon Kendra para indexar su fuente de datos de Quip, realice estos cambios en su Quip y AWS cuentas.

En Quip, asegúrate de tener:

- Una cuenta de Quip con permisos administrativos.
- Creé credenciales de autenticación de Quip que incluyen un token de acceso personal. Consulte [Documentación de Quip sobre autenticación](#) para obtener más información.
- Copió el dominio de su sitio de Quip. Por ejemplo, <https://quip-company.quipdomain.com/> **browse** donde **dominio quipes** el dominio.
- Se ha comprobado que cada documento es único en Quip y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapelpara su fuente de datos y](#), si usa la API, anotó el ARN dellAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Almacenó sus credenciales de autenticación de Quip en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Quip aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa su fuente de datos de Quip, debe proporcionar los detalles necesarios de su fuente de datos de Quip para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configurado Quip paraAmazon Kendra, consulte[Requisitos previos \(p. 424\)](#).

Console

Para conectarAmazon Kendraa Quip

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConektor Quipy, a continuación, elijaAregar fuente de datos.
5. En elEspecificardetalles de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefina el acceso y la seguridadpágina, introduzca la siguiente información:

- a. Nombre de dominio Quip—Introduzca el Quip que ha copiado de su cuenta de Quip.
- b. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar sus credenciales de autenticación de Quip. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - i. Introduzca la siguiente información enCrea unAWS Secrets Managerventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Quip-' se añade automáticamente a tu nombre secreto.
 - B. Token Quip—Introduzca el token de acceso personal de Quip que creó en su cuenta de Quip.
 - ii. Seleccione Guardar.
- c. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesyGrupos de seguridad de VPC.
- d. IAMpapel—Elige uno existenteIAMrol o crea uno nuevolAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- e. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. Agregue los identificadores de carpetas de Quip para rastrear—Los ID de carpeta de Quip que desea rastrear.
 - b. Configuración adicional (tipos de contenido)—Introduzca los tipos de contenido que desee rastrear.
 - c. Patrones Regex—Patrones de expresiones regulares para incluir o excluir ciertos archivos. Puede añadir hasta 100 patrones.
 - d. EnCronograma de ejecución de sincronización, paraFrecuencia—Elija la frecuencia con la que Amazon Kendra se sincronizará con su fuente de datos.
 - e. Elija Siguiente.
8. En elDefinir mapeos de campospágina, introduzca la siguiente información:
 - a. Seleccione uno de los campos de fuente de datos predeterminados generados a los que desee asignar.Amazon Kendraíndice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon Kendraa Quip

Debe especificar lo siguiente medianteQuipConfiguration API:

- Dominio del sitio Quip—Por ejemplo,<https://quip-company.quipdomain.com/> browsedonde **dominio quipes** el dominio.

- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de su cuenta de Quip. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "accessToken": "token"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMPapel—EspecificarRoleArn cuando llamasCreateDataSourcepara proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para el conector Quip yAmazon Kendra. Para obtener más información, consulte[IAMfunciones para las fuentes de datos de Quip](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfigurationcomo parte de la configuración de la fuente de datos. Consulte[ConfigurandoAmazon Kendrausar una VPC](#).
- Filtros de inclusión y exclusión: especifique si desea incluir o excluir determinados archivos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elia asignar los campos de su fuente de datos de Quip a suAmazon Kendracampos de índice. Para obtener más información, consulte[Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendrastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte[Filtrado de contexto de usuario para fuentes de datos de Quip](#).

Más información

Para obtener más información sobre la integraciónAmazon Kendracon su fuente de datos de Quip, consulte:

- [Busque conocimientos en documentos de Quip con la búsqueda inteligente mediante el conector Quip paraAmazon Kendra](#)

Salesforce

Salesforce es una herramienta de gestión de relaciones con los clientes (CRM) para gestionar los equipos de soporte, ventas y marketing. Puedes usarAmazon Kendrapara indexar sus objetos estándar de Salesforce e incluso objetos personalizados.

Puedes conectarte a Amazon Kendra a su fuente de datos de Salesforce mediante el [Amazon Kendra consola](#), el [TemplateConfiguration API](#), o [Salesforce Configuration API](#).

Amazon Kendra tiene dos versiones del conector de Salesforce. Las funciones compatibles de cada versión incluyen:

Conejor Salesforce V1.0/[SalesforceConfiguration API](#)

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión

Conejor Salesforce V2.0/[TemplateConfiguration API](#)

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Compatibilidad con archivos adjuntos de entidades
- Rastreo de identidad
- Compatibilidad con VPC
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados

Note

Soporte para el conector Salesforce V1.0/SalesforceConfiguration. Está previsto que la API finalice en 2023. Recomendamos migrar a Salesforce Connector V2.0/TemplateConfigurationAPI.

Para solucionar problemas de Amazon Kendra Conector de fuente de datos de Salesforce, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Conejor Salesforce V1.0 \(p. 428\)](#)
- [Conejor Salesforce V2.0 \(p. 433\)](#)

Conejor Salesforce V1.0

Salesforce es una herramienta de gestión de relaciones con los clientes (CRM) para gestionar los equipos de soporte, ventas y marketing. Puedes usar Amazon Kendra para indexar sus objetos estándar de Salesforce e incluso objetos personalizados.

Important

Amazon Kendra utiliza la versión 48 de la API de Salesforce. La API de Salesforce limita la cantidad de solicitudes que puede realizar por día. Si Salesforce supera esas solicitudes, lo vuelve a intentar hasta que pueda continuar.

Note

Soporte para el conector Salesforce V1.0/SalesforceConfiguration. Está previsto que la API finalice en 2023. Recomendamos migrar a Salesforce Connector V2.0/TemplateConfigurationAPI.

Para solucionar problemas de Amazon Kendra Conector de fuente de datos de Salesforce, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 429\)](#)
- [Requisitos previos \(p. 429\)](#)
- [Instrucciones de conexión \(p. 430\)](#)

Características admitidas

Amazon KendraEl conector de fuente de datos de Salesforce admite las siguientes funciones:

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión

Requisitos previos

Antes de poder usarAmazon Kendrapara indexar su fuente de datos de Salesforce, realice estos cambios en su Salesforce yAWS cuentas.

En Salesforce, asegúrese de tener:

- Creé una cuenta de Salesforce y anoté el nombre de usuario y la contraseña que usa para conectarse a Salesforce.
- Creé una cuenta de Salesforce Connected App con OAuth activado y copié la clave de consumidor (ID de cliente) y el secreto de consumidor (secreto de cliente) asignados a tu aplicación Salesforce Connected. Consulte[Documentación de Salesforce sobre aplicaciones conectadas](#)para obtener más información.
- Copió el token de seguridad de Salesforce asociado a la cuenta utilizada para conectarse a Salesforce.
- Copió la URL de la instancia de Salesforce que desea indexar. Por lo general, esto es<company><https://salesforce.com/>. El servidor debe ejecutar una aplicación conectada a Salesforce.
- Se agregaron credenciales a su servidor de Salesforce para un usuario con acceso de solo lectura a Salesforce mediante la clonación delReadOnlyperfil y luego agregar los permisos Ver todos los datos y administrar artículos. Estas credenciales identifican al usuario que realiza la conexión y a la aplicación conectada a Salesforce queAmazon Kendrase conecta a.
- Se comprobó que cada documento es único en Salesforce y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellIAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulIAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Almacenó sus credenciales de autenticación de Salesforce en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda que actualice o rote periódicamente sus credenciales y su secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las

credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto al conectar su fuente de datos de Salesforce aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa su fuente de datos de Salesforce, debe proporcionar los detalles necesarios de su fuente de datos de Salesforce para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configurado Salesforce paraAmazon Kendraver[Requisitos previos \(p. 429\)](#).

Console

Para conectarAmazon Kendraa Salesforce

1. Inicia sesión en elAWSConsola de administración y abra[Amazon Kendraconsola](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enConfiguración del índice.

3. En elCómo empezarpágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConektor Salesforce V1.0y, a continuación, elijaAñadir coneotor.
5. En elEspecificardetalless de la fuente de datospágina, introduzca la siguiente información:
 - a. Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. Idioma predeterminado— Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos anula el idioma seleccionado.
 - d. Añadir etiqueta nueva—Etiquetas para buscar y filtrar sus recursos o realizar un seguimiento de sus costos compartidos.
 - e. Elija Siguiente.
6. En elDefinardel acceso y la seguridadpágina, introduzca la siguiente información:
 - a. URL de Salesforce—Introduzca la URL de la instancia del sitio de Salesforce que deseé indexar.
 - b. ParaTipo de autenticación, elige entreExistente yNuevo para almacenar sus credenciales de autenticación de Salesforce. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - Introduzca la siguiente información enCrea unAWS Secrets Managerventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Salesforce-' se añade automáticamente a tu nombre secreto.
 - B. ParaNombre de usuario,Contraseña,Token de seguridad,Clave para el consumidor,Secreto del consumidor, yURL de autenticación—Introduzca los valores de las credenciales de autenticación que creó en su cuenta de Salesforce.
 - C. EscojaGuardar autenticación.

- c. IAMpapel—Elige uno existenteIAMrol o crea uno nuevoIAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- d. Elija Siguiente.

7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:

- a. ParaRastrea archivos adjuntos—Seleccione esta opción para rastrear todos los objetos, artículos y ficheros adjuntos.
- b. ParaObjetos estándar,Artículos de conocimiento, yFeeds de Chatter—Seleccione las entidades de Salesforce o los tipos de contenido que deseé rastrear.

Note

Debe proporcionar información de configuración para indexar al menos uno de los objetos estándar, los artículos de conocimiento o las fuentes de chat. Si eliges gatearArtículos de conocimiento debe especificar los tipos de artículos de conocimiento que se van a indexar, el nombre de los artículos y si desea indexar los campos estándar de todos los artículos de conocimiento o solo los campos de un tipo de artículo personalizado. Si decide indexar artículos personalizados, debe especificar el nombre interno del tipo de artículo. Puede especificar hasta 10 tipos de artículos.

- c. Frecuencia— Con qué frecuenciaAmazon Kendrase sincronizará con su fuente de datos.
- d. Elija Siguiente.

8. En elDefinir mapeos de campospágina, introduzca la siguiente información:

- a. ParaArtículo de conocimiento estándar,Adjuntos de objetos estándar, yMapeos de campos sugeridos adicionales—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que deseé asignar a su índice.

Note

Un mapeo de índices para_document_bodies obligatorio. No puede cambiar el mapeo entre losSalesforce IDcampo y elAmazon Kendra _document_id campo.

- b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon Kendraa Salesforce

Debe especificar lo siguiente:[SalesforceConfigurationAPI](#):

- URL del servidor: la URL de la instancia del sitio de Salesforce que deseá indexar.

- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de su cuenta de Salesforce. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "authenticationUrl": "OAUTH endpoint that Amazon Kendra connects to get an OAUTH token",  
    "consumerKey": "Application public key generated when you created your Salesforce application",  
    "consumerSecret": "Application private key generated when you created your Salesforce application.",  
    "password": "Password associated with the user logging in to the Salesforce instance",  
    "securityToken": "Token associated with the user account logging in to the Salesforce instance",  
    "username": "User name of the user logging in to the Salesforce instance"  
}
```

Note

Se recomienda que actualice o rote periódicamente sus credenciales y su secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSource para proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para el conector de Salesforce yAmazon Kendra. Para obtener más información, consulte [IAMfunciones para las fuentes de datos de Salesforce](#).
- Debe proporcionar información de configuración para indexar al menos uno de los objetos estándar, los artículos de conocimiento o las fuentes de chat.
 - Objetos estándar—Si eliges gatearObjetos estándar, debe especificar el nombre del objeto estándar y el nombre del campo de la tabla de objetos estándar que contiene el contenido del documento.
 - Artículos de conocimiento—Si eliges gatearArtículos de conocimiento, debe especificar los tipos de artículos de conocimiento que se van a indexar, los estados de los artículos de conocimiento que se van a indexar y si se deben indexar los campos estándar de todos los artículos de conocimiento o solo los campos de un tipo de artículo personalizado.
 - Feeds de Chatter—Si eliges gatearFeeds de Chatter, debe especificar el nombre de la columna en SalesforceFeedItemtabla que contiene el contenido que se va a indexar.

También puede añadir las siguientes funciones opcionales:

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir ciertos archivos adjuntos.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elija asignar los campos de su fuente de datos de Salesforce a suAmazon Kendracampos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).

- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Salesforce](#).

Conejor Salesforce V2.0

Salesforce es una herramienta de gestión de relaciones con los clientes (CRM) para gestionar los equipos de soporte, ventas y marketing. Puedes usar Amazon Kendra para indexar sus objetos estándar de Salesforce e incluso objetos personalizados.

Note

Soporte para el conector Salesforce V1.0/SalesforceConfiguration. Está previsto que la API finalice en 2023. Recomendamos migrar a Salesforce Connector V2.0/TemplateConfigurationAPI.

Para solucionar problemas de Amazon Kendra Conector de fuente de datos de Salesforce, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 433\)](#)
- [Requisitos previos \(p. 433\)](#)
- [Instrucciones de conexión \(p. 434\)](#)
- [Más información \(p. 438\)](#)

Características admitidas

Amazon Kendra El conector de fuente de datos de Salesforce admite las siguientes funciones:

- Mapeos de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Compatibilidad con archivos adjuntos de entidades
- Rastreo de identidad
- Compatibilidad con VPC
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados

Requisitos previos

Antes de poder usar Amazon Kendra para indexar su fuente de datos de Salesforce, realice estos cambios en su Salesforce y AWS cuentas.

En Salesforce, asegúrese de tener:

- Creó una cuenta administrativa de Salesforce y anoté el nombre de usuario y la contraseña que usa para conectarse a Salesforce.
- Copió el token de seguridad de Salesforce asociado a la cuenta utilizada para conectarse a Salesforce.
- Creó una cuenta de Salesforce Connected App con OAuth activado y copió la clave de consumidor (ID de cliente) y el secreto de consumidor (secreto de cliente) asignados a tu aplicación Salesforce Connected. Consulte [Documentación de Salesforce sobre aplicaciones conectadas](#) para obtener más información.

- Copió la URL de la instancia de Salesforce que desea indexar. Por lo general, esto es `<company>https://.salesforce.com/`. El servidor debe ejecutar una aplicación conectada a Salesforce.
- Se agregaron credenciales a su servidor de Salesforce para un usuario con acceso de solo lectura a Salesforce mediante la clonación del Read Only perfil y luego agregar los permisos Ver todos los datos y administrar artículos. Estas credenciales identifican al usuario que realiza la conexión y a la aplicación conectada a Salesforce que Amazon Kendra se conecta a.
- Se comprobó que cada documento es único en Salesforce y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu cuenta de AWS, asegúrate de tener:

- Creó un Amazon Kendra índice, si usa la API, anota el ID del índice.
- Creó un IAM papel para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al correcto AWS Secrets Manager identificación secreta.

- Almacenó sus credenciales de autenticación de Salesforce en un AWS Secrets Manager secreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda que actualice o rote periódicamente sus credenciales y su secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendado reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o secreto, puedes usar la consola para crear un nuevo IAM rol y Secrets Manager secreto al conectar su fuente de datos de Salesforce a Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existente IAM rol y Secrets Manager secreto y un identificador de índice.

Instrucciones de conexión

Para conectar Amazon Kendra a su fuente de datos de Salesforce, debe proporcionar los detalles necesarios de su fuente de datos de Salesforce para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado Salesforce para Amazon Kendra ver [Requisitos previos \(p. 433\)](#).

Console

Para conectar Amazon Kendra a Salesforce:

1. Inicia sesión en el AWS Consola de administración y abra [Amazon Kendra consola](#).
2. En el panel de navegación de la izquierda, selecciona Índices y, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar su Control de acceso de usuarios ajustes en Configuración del índice.

3. En el Cómo empezar página, elige Agregar fuente de datos.
4. En el Agregar fuente de datos página, elige Conector Salesforce V2.0 y, a continuación, elija Añadir conector.

5. En el Especificar los detalles de la fuente de datos página, introduzca la siguiente información:
 - a. Nombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. Idioma predeterminado— Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos anula el idioma seleccionado.
 - d. Elija Siguiente.
6. En el Defina el acceso y la seguridad página, introduzca la siguiente información:
 - a. URL de Salesforce—Introduzca la URL de la instancia del sitio de Salesforce que desea indexar.
 - b. Introduzca un secreto existente o, si crea uno nuevo, unAWS Secrets Manager se abre una ventana secreta.
 - Introduzca la siguiente información en Crea unAWS Secrets Manager ventana secreta:
 - A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Salesforce-' se añade automáticamente a tu nombre secreto.
 - B. Para Nombre de usuario, Contraseña, Token de seguridad, Clave para el consumidor, Secreto del consumidor, y URL de autenticación—Introduzca los valores de las credenciales de autenticación que generó y descargó de su cuenta de Salesforce.
 - C. Escoja Guardar autenticación.
 - c. Rastreador de identidades Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el [PutPrincipalMappingAPI](#).
 - d. IAM papel—Elige uno existente IAM rol o crea uno nuevo IAM rol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAM roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un nuevo rol para evitar errores.

- e. Elija Siguiente.
7. En el Configurar los ajustes de sincronización página, introduzca la siguiente información:
 - a. Para Rastrear archivos adjuntos—Seleccione esta opción para rastrear todos los objetos de Salesforce adjuntos.
 - b. Para Objetos estándar, Objetos estándar con accesorios, y Objeto estándar sin accesorio y Artículos de conocimiento—Seleccione las entidades de Salesforce o los tipos de contenido que deseé rastrear.
 - c. Debe proporcionar información de configuración para indexar al menos uno de los objetos estándar, los artículos de conocimiento o las fuentes de chat. Si eliges gatear Artículos de conocimiento debe especificar los tipos de artículos de conocimiento que se van a indexar. Puede elegir entre publicaciones, archivos, borradores y archivos adjuntos.

Filtro Regex: especifique un patrón de expresiones regulares para incluir elementos específicos del catálogo.

8. Para Configuración adicional:
 - Información de ACL Todas las listas de control de acceso se incluyen de forma predeterminada. Al anular la selección de una lista de control de acceso, todos los archivos de esa categoría se harán públicos.

- Patrones Regex—Añada patrones de expresiones regulares para incluir o excluir determinados archivos. Puede añadir hasta 100 patrones.

Para modo de sincronización en Salesforce v2, elija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Cuando sincronizas tu fuente de datos con Amazon Kendra por primera vez, todo el contenido se sincroniza de forma predeterminada.

- Sincronización completa—Sincronice todo el contenido independientemente del estado de sincronización anterior.
- Sincronización de contenido nuevo, modificado y eliminado—Sincronice únicamente contenido nuevo, modificado y eliminado.

Sincronización de contenido nuevo y modificado—Sincronice solo contenido nuevo y modificado.

9. Elija Siguiente.

10. En el Establecer mapeos de campos página, introduzca la siguiente información:

- Para Artículo de conocimiento estándar, Adjuntos de objetos estándar, y Mapeos de campos sugeridos adicionales—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados que desea asignar a su índice.

Note

Un mapeo de índices para_document_bodies obligatorio. No puede cambiar el mapeo entre los Salesforce ID campo y el Amazon Kendra _document_id campo.

- Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
- Elija Siguiente.

11. En el Revisar y crear página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en Fuentes de datos página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a Salesforce

Debe especificar un JSON de [esquema de fuente de datos](#) utilizando el [TemplateConfiguration API](#). Debe proporcionar la siguiente información:

- Fuente de datos—Especifique la fuente de datos como SALESFORCEV2.
- URL del servidor—Especifique la URL del host de la instancia de Salesforce.
- Tclea—Especificar TEMPLATE como el tipo cuando llamas CreateDataSource.
- Modo de sincronización—Especifique si Amazon Kendra debe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:
 - FORCED_FULL_CRAWL para volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice
 - FULL_CRAWL para rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contiene las credenciales de autenticación de su cuenta de Salesforce. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "authenticationUrl": "OAUTH endpoint that Amazon Kendra connects to get an OAUTH token",  
    "consumerKey": "Application public key generated when you created your Salesforce application",  
    "consumerSecret": "Application private key generated when you created your Salesforce application",  
    "password": "Password associated with the user logging in to the Salesforce instance",  
    "securityToken": "Token associated with the user account logging in to the Salesforce instance",  
    "username": "User name of the user logging in to the Salesforce instance"  
}
```

Note

Se recomienda que actualice o rote periódicamente sus credenciales y su secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMpapel—EspecificarRoleArn cuando llamasCreateDataSource para proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para el conector de Salesforce yAmazon Kendra. Para obtener más información, consulte [IAMfunciones para las fuentes de datos de Salesforce](#).

También puede añadir las siguientes funciones opcionales:

- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir ciertos documentos, cuentas, campañas, casos, contactos, clientes potenciales, oportunidades, soluciones, tareas, grupos, usuarios de chat y archivos de entidades personalizados.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Nube privada virtual (VPC)—EspecificarVpcConfiguration cuando llamasCreateDataSource. Para obtener más información, consulte [ConfigurandoAmazon Kendrautilizar unAmazon VPC \(p. 465\)](#).
- Habilitar el rastreador de identidades: especifique si desea activar el rastreador de identidades. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el[PutPrincipalMapping](#)API. Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado en contexto de usuario](#).
- Mapeos de campo—Elija asignar los campos de su fuente de datos de Salesforce a suAmazon Kendracampos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Salesforce](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [Esquema de plantillas de Salesforce \(p. 233\)](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Salesforce, consulte:

- [Anunciamos el conector de Salesforce actualizado \(V2\) para Amazon Kendra](#)

ServiceNow

ServiceNow proporciona un sistema de administración de servicios basado en la nube para crear y administrar flujos de trabajo a nivel de organización, como los servicios de TI, los sistemas de venta de entradas y el soporte. Puedes usar Amazon Kendra para indexar sus ServiceNow catálogos, artículos de conocimiento, incidentes y sus anexos.

Puedes conectarte a Amazon Kendra a través de tu ServiceNow fuente de datos mediante la [Amazon Kendra consola](#), la [TemplateConfiguration API](#) o la [ServiceNow Configuration API](#).

Amazon Kendra tiene dos versiones del ServiceNow conector. Las funciones compatibles de cada versión incluyen:

ServiceNow conector V1.0/[ServiceNow Configuration API](#)

- Mapeos de campo
- ServiceNow versiones de instancia: Londres, otras
- Patrones de inclusión/exclusión: catálogos de servicios, artículos de conocimiento, archivos adjuntos

ServiceNow conector V2.0/[Template Configuration API](#)

- Mapeos de campo
- Filtrado de contexto de usuario
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados
- ServiceNow versiones de instancia: Roma, Sandiego, Tokio, otras
- Patrones de inclusión/exclusión: catálogos de servicios, artículos de conocimiento, incidentes, archivos adjuntos

Note

Soporte para ServiceNow conector V1.0/ServiceNow Configuration API. Está previsto que la API finalice en 2023. Recomendamos migrar a o usar ServiceNow conector V2.0/Template Configuration API.

Para solucionar problemas de Amazon Kendra ServiceNow conector de fuente de datos, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [ServiceNow conector V1.0 \(p. 439\)](#)
- [ServiceNow conector V2.0 \(p. 443\)](#)

- [Especificar documentos para indexarlos con una consulta \(p. 449\)](#)

ServiceNowconector V1.0

ServiceNowproporciona un sistema de administración de servicios basado en la nube para crear y administrar flujos de trabajo a nivel de organización, como los servicios de TI, los sistemas de venta de entradas y el soporte. Puedes usarAmazon Kendrapara indexar susServiceNowcatálogos, artículos de conocimiento y sus anexos.

Note

Soporte paraServiceNowconector V1.0/ServiceNowConfigurationEstá previsto que la API finalice en 2023. Recomendamos migrar a o usarServiceNowconector V2.0/TemplateConfigurationAPI.

Para solucionar problemas deAmazon Kendra ServiceNowconector de fuente de datos, consulte[Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 439\)](#)
- [Requisitos previos \(p. 439\)](#)
- [Instrucciones de conexión \(p. 440\)](#)
- [Más información \(p. 443\)](#)

Características admitidas

Amazon Kendra ServiceNowel conector de fuente de datos admite las siguientes funciones:

- ServiceNowversiones de instancia: Londres, otras
- Patrones de inclusión/exclusión: catálogos de servicios, artículos de conocimiento y sus anexos

Requisitos previos

Antes de poder usarAmazon Kendrapara indexar susServiceNowfuente de datos, realice estos cambios en suServiceNowyAWS cuentas.

EnServiceNow, asegúrate de tener:

- Creó unServiceNowcuenta de administrador y han creado unaServiceNowinstancia.
- Copió el anfitrión de suServiceNowURL de instancia. Por ejemplo, si la URL de la instancia es`https://your-domain.service-now.com`, el formato de la URL de host que introduzca es`su-dominiio.service-now.com`.
- Tomó nota de sus credenciales de autenticación básicas, que contienen un nombre de usuario y una contraseña para permitirAmazon Kendrapara conectarse a suServiceNowinstancia.
- Opcional:Se configuró un token de credenciales de OAuth 2.0 que puede identificarAmazon Kendray generar un nombre de usuario, una contraseña, un ID de cliente y un secreto de cliente. El nombre de usuario y la contraseña deben permitir el acceso aServiceNowbase de conocimientos y catálogo de servicios. Consulte[ServiceNowdocumentación sobre la autenticación OAuth 2.0](#)para obtener más información.
- Se agregaron los siguientes permisos:
 - kb_category
 - kb_knowledge

- kb_knowledge_base
- kb_uc_not_read_mtom
- kb_uc_can_read_mtom
- sc_catalog
- sc_category
- sc_cat_item
- sys_attachment
- sys_attachment_doc
- rol de usuario del sistema
- Se ha marcado que cada documento es único en ServiceNow en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu cuenta de AWS, asegúrate de tener:

- [Creó un Amazon Kendra índice](#), si usa la API, anota el ID del índice.
- [Creó un IAM papel](#) para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al correcto AWS Secrets Manager identificación secreta.

- Almacenó sus credenciales de autenticación en un AWS Secrets Manager secreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es importante recomendar reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o secreto, puedes usar la consola para crear un nuevo IAM rol y Secrets Manager secreto cuando conectas tu ServiceNow fuente de datos para Amazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existente IAM rol y Secrets Manager secreto y un identificador de índice.

Instrucciones de conexión

Para conectar Amazon Kendra a tu ServiceNow fuente de datos, debe proporcionar los detalles necesarios de su ServiceNow fuente de datos para que Amazon Kendra pueda acceder a sus datos. Si aún no ha configurado ServiceNow por Amazon Kendra vea [Requisitos previos \(p. 439\)](#).

Console

Para conectar Amazon Kendra a ServiceNow

1. Inicia sesión en la AWS Consola de administración y abra el [Amazon Kendra consola](#).
2. En el panel de navegación de la izquierda, selecciona Índices y, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar su Control de acceso de usuarios ajustes en Ajustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeServiceNowconector V1.0y, a continuación, elijaAregar fuente de datos.
5. En elEspecificardetalle de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefinir acceso y la seguridadpágina, introduzca la siguiente información:
 - a. ServiceNowanfitrón—Introduzca elServiceNowURL del servidor.
 - b. ServiceNowversión—Seleccione suServiceNowversión.
 - c. Elige entreAutenticación básica yAutenticación OAuth 2.0según su caso de uso.
 - d. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para guardar tusServiceNowcredenciales de autenticación. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - i. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-ServiceNow-' se añade automáticamente a tu nombre secreto.
 - ii. Si utiliza la autenticación básica, introduzcaNombre secreto,Nombre de usuario, yContraseña para tuServiceNowcuenta.
Si usa la autenticación OAuth2, introduzca elNombre secreto,Nombre de usuario,Contraseña, ID de cliente, ySecreto del cliente creaste en tuServiceNowcuenta.
 - iii. EscojaGuardar y añadir secreto.
 - e. IAMpapel—Elige uno existenteIAMrol o crea uno nuevoIAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- f. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. Incluir artículos de conocimiento—Elige indexar los artículos de conocimiento.
 - b. Tipo de artículos de conocimiento—Elige entreIncluye solo artículos públicos yIncluye artículos basados enServiceNowconsulta de filtrosegún su caso de uso. Si seleccionasIncluye artículos basados enServiceNowconsulta de filtro, debe introducir unConsulta de filtro copiado de tuServiceNowcuenta.
 - c. Incluir artículos de conocimiento adjuntos—Elige indexar los archivos adjuntos de los artículos de conocimiento. También puede seleccionar tipos de archivos específicos para indexarlos.
 - d. Incluir elementos del catálogo—Elige indexar los elementos del catálogo.
 - e. Incluir archivos adjuntos de elementos del catálogo—Elige indexar los archivos adjuntos de los elementos del catálogo. También puede seleccionar tipos de archivos específicos para indexarlos.

- f. Frecuencia— Con qué frecuencia Amazon Kendra sincronizará con su fuente de datos.
 - g. Elija Siguiente.
8. En el Establecer mapeos de campos página, introduzca la siguiente información:
 - a. Artículos de conocimiento y Catálogo de servicios—Seleccione una de las Amazon Kendra campos de fuente de datos predeterminados generados y mapeos de campos sugeridos adicionales que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
 9. En el Revisar y crear página, compruebe que la información que ha introducido es correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá en Fuentes de datos página después de que se haya agregado correctamente.

API

Para conectar Amazon Kendra a ServiceNow

Debe especificar lo siguiente mediante [ServiceNow Configuration API](#):

- URL de la fuente de datos—Especifique el ServiceNow URL. El punto de conexión del servidor debería tener el siguiente aspecto: **su-dominio.service-now.com**.
- Instancia host de fuente de datos—Especifique el ServiceNow versión de instancia host como cualquiera LONDONo OTHERS.
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contiene las credenciales de autenticación que creó en su ServiceNow cuenta.

Si utiliza la autenticación básica, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "user name",  
    "password": "password"  
}
```

Si utiliza la autenticación OAuth2, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "user name",  
    "password": "password",  
    "clientId": "client id",  
    "clientSecret": "client secret"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Es recomendado reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAM papel—Especificar RoleArn cuando llamas CreateDataSource para proporcionar un IAM rol con permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas

para el ServiceNowconector y Amazon Kendra. Para obtener más información, consulte [IAMfunciones para ServiceNowfuentes de datos](#).

También puede añadir las siguientes funciones opcionales:

- Mapeos de campo—Elija mapear su ServiceNowcampos de fuente de datos para su Amazon Kendracampos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtros de inclusión y exclusión—Especifique si desea incluir o excluir ciertos archivos adjuntos de catálogos y artículos de conocimiento.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Parámetros de indexación—También puede elegir especificar si desea:
 - Indexe artículos de conocimiento y catálogos de servicios, o ambos. Si decide indexar los artículos de conocimiento y los elementos del catálogo de servicios, debe proporcionar el nombre del ServiceNowcampo que está asignado al campo de contenido del documento de índice en Amazon Kendraíndice.
 - Indexe los archivos adjuntos a artículos de conocimiento y elementos del catálogo.
 - Usa un ServiceNowconsulta que selecciona documentos de una o más bases de conocimiento. Las bases de conocimiento pueden ser públicas o privadas. Para obtener más información, consulte [Specifying documents to index with a query](#) (Especificar documentos a indexar con una consulta).

Más información

Para obtener más información sobre la integración Amazon Kendra con tu ServiceNowfuente de datos, consulte:

- [Empezar con Amazon Kendra ServiceNowConector en línea](#)

ServiceNowconector V2.0

ServiceNowproporciona un sistema de administración de servicios basado en la nube para crear y administrar flujos de trabajo a nivel de organización, como los servicios de TI, los sistemas de venta de entradas y el soporte. Puedes usar Amazon Kendrapara indexar sus ServiceNowcatálogos, artículos de conocimiento, incidentes y sus anexos.

Para solucionar problemas de Amazon Kendra ServiceNowconector de fuente de datos, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 444\)](#)
- [Requisitos previos \(p. 444\)](#)
- [Instrucciones de conexión \(p. 445\)](#)
- [Más información \(p. 449\)](#)

Características admitidas

Amazon Kendra ServiceNowel conector de fuente de datos admite las siguientes funciones:

- Mapeos de campo
- Filtrado de contexto de usuario
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])
- Sincronizar todos los documentos/Sincronizar solo documentos nuevos, modificados o eliminados
- ServiceNowversiones de instancia: Roma, Sandiego, Tokio, otras
- Patrones de inclusión/exclusión: catálogos de servicios, artículos de conocimiento, incidentes y sus anexos

Requisitos previos

Antes de poder usarAmazon Kendrapara indexar susServiceNowfuente de datos, realice estos cambios en suServiceNowyAWS cuentas.

EnServiceNow, asegúrate de tener:

- Creé una instancia de desarrollador personal o empresarial y tengo unServiceNowinstancia con una función administrativa.
- Copió el anfitrión de suServiceNowURL de instancia. El formato de la URL de host que introduzca es ***su-dominio.service-now.com***. Necesitas tuServiceNowURL de instancia a la que conectarseAmazon Kendra.
- Credenciales de autenticación básicas configuradas que contienen un nombre de usuario y una contraseña para permitirAmazon Kendrapara conectarse a suServiceNowinstancia.
- Opcional:Se configuró un token de credenciales de OAuth 2.0 que puede identificarAmazon Kendramediante un nombre de usuario, una contraseña, un ID de cliente generado y un secreto de cliente. Consulte[ServiceNowdocumentación sobre la autenticación OAuth 2.0](#)para obtener más información.
- Se ha marcado que cada documento es único enServiceNowy en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tuCuenta de AWS, asegúrate de tener:

- [Creó unAmazon Kendraíndicey](#), si usa la API, anota el ID del índice.
- [Creó unIAMpapel](#)para su fuente de datos y, si usa la API, anotó el ARN dellIAMrol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tulIAMrol para acceder al correctoAWS Secrets Manageridentificación secreta.

- Almacenó suServiceNowcredenciales de autenticación en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tuServiceNowfuente de datos paraAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tuServiceNowfuente de datos, debe proporcionar los detalles necesarios de suServiceNowfuente de datos para queAmazon Kendrapuede acceder a sus datos. Si aún no ha configuradoServiceNowporAmazon Kendraver[Requisitos previos \(p. 444\)](#).

Console

Para conectarAmazon KendraaServiceNow

1. Inicia sesión en elAWSConsola de administración y abra el[Amazon Kendraconsola](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAgregar fuente de datos.
4. En elAgregar fuente de datospágina, eligeServiceNowconector V2.0y, a continuación, elijaAgregar fuente de datos.
5. En elEspecificardetalle de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefinadetalle de la fuente de datos, introduzca la siguiente información:
 - a. ServiceNowanfitrón—Introduzca elServiceNowURL del servidor. El formato de la URL de host que introduzca es`su-dominio.service-now.com`.
 - b. ServiceNowversión—Seleccione suServiceNowversión.
 - c. Elige entreAutenticación básica yAutenticación OAuth 2.0según su caso de uso.
 - d. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para guardar tusServiceNowcredenciales de autenticación. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta. Introduzca la siguiente información en la ventana:
 - i. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-ServiceNow-' se añade automáticamente a tu nombre secreto.
 - ii. Si utiliza la autenticación básica, introduzcaNombre secreto,Nombre de usuario, yContraseña para tuServiceNowcuenta.

Si usa la autenticación OAuth2.0, introduzca elNombre secreto,Nombre de usuario,Contraseña, ID de cliente, ySecreto del clientecreate en tuServiceNowcuenta.

- iii. EscojaGuardar y añadir secreto.

- e. (Opcional)Configurar VPC y grupo de seguridad—Seleccione una VPC para usarla con suServiceNowinstancia.
- f. Rastreador de identidades—Elija activarAmazon Kendrarastreador de identidad para sincronizar la información de identidad. Si optas por desactivar el rastreador de identidades, debes cargar la información principal mediante el[PutPrincipalMapping API](#).
- g. IAMpapel—Elige uno existenteIAMrol o crea uno nuevolAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- h. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. ParaArtículos de conocimiento, elija una de las siguientes opciones:
 - Artículos de conocimiento—Elija indexar los artículos de conocimiento.
 - Adjuntos de artículos de conocimiento—Elija indexar los archivos adjuntos de los artículos de conocimiento.
 - Tipo de artículos de conocimiento—Elige entreSolo artículos públicosArtículos de conocimiento basados enServiceNowconsulta de filtrosegún su caso de uso. Si seleccionasIncluye artículos basados enServiceNowconsulta de filtro, debe introducir unConsulta de filtrocopiado de tuServiceNowcuenta. Los ejemplos de consultas de filtro incluyen:*workflow_state=Borraft^EQ, kb_knowledge_base=dfc19531bf2021003f07e2c1ac07 TEXTO NO ESTÁ VACÍO^EQ, article_type=text^active=True^EQ*.
 - Incluye artículos basados en un filtro de descripción breve—Especifique patrones de expresiones regulares para incluir o excluir artículos específicos.
 - b. ParaElementos del catálogo de servicios:
 - Elementos del catálogo de servicios—Elija indexar los elementos del catálogo de servicios.
 - Adjuntos de elementos del catálogo de servicios—Elija indexar los archivos adjuntos de los elementos del catálogo de servicios.
 - Elementos del catálogo de servicios activos—Elija indexar los elementos del catálogo de servicios activos.
 - Elementos inactivos del catálogo de servicios—Elija indexar los elementos inactivos del catálogo de servicios.
 - Consulta de filtro—Elija incluir elementos del catálogo de servicios en función de un filtro definido en suServiceNowinstancia. Los ejemplos de consultas de filtro incluyen:*descripción_corta como acceso^category=280952237b1300054b6a3549dbe5dd4^eq, El nombre comienza con Service^active=true^EQ*.
 - Incluya elementos del catálogo de servicios basados en un filtro de descripción breve: especifique un patrón de expresiones regulares para incluir elementos específicos del catálogo.
 - c. ParaIncidentes:
 - Incidentes—Elija indexar los incidentes del servicio.
 - Archivos adjuntos del incidente—Elija indexar los archivos adjuntos de los incidentes.
 - Incidentes activos—Elija indexar los incidentes activos.
 - Incidentes inactivos—Elija indexar los incidentes inactivos.

- Tipo de incidente activo—Elige entreTodos los incidentes,Incidentes abiertos,Incidentes abiertos: sin asignar, ylncidencias resueltasen función de su caso de uso.
 - Consulta de filtro—Elija incluir los incidentes en función de un filtro definido en suServiceNowinstancia. Los ejemplos de consultas de filtro incluyen:*Short_descriptionLikeTest^Urgencia=3^Estado=1^EQ,Priority=2^Categoria=S*
 - Incluya los incidentes en función del filtro de descripción breve: especifique un patrón de expresiones regulares para incluir incidentes específicos.
- d. ParaConfiguración adicional:
- Información de ACL—Las listas de control de acceso de las entidades que ha seleccionado se incluyen de forma predeterminada. Al anular la selección de una lista de control de acceso, todos los archivos de esa categoría se harán públicos. Las opciones de ACL se desactivan automáticamente para las entidades no seleccionadas. Para los artículos públicos no se aplica la ACL.
 - Patrones de expresiones regulares de adjunto—Añada patrones de expresiones regulares para incluir o excluir ciertos archivos adjuntos de catálogos, artículos de conocimiento e incidentes. Puede añadir hasta 100 patrones.
- e. ParaModo de sincronizaciónelija cómo desea actualizar el índice cuando cambie el contenido de la fuente de datos. Cuando sincronizas tu fuente de datos conAmazon Kendrapor primera vez, todo el contenido se sincroniza de forma predeterminada.
- Sincronización completa—Sincronice todo el contenido independientemente del estado de sincronización anterior.
 - Sincronización de contenido nuevo, modificado o eliminado—Sincronice únicamente contenido nuevo, modificado y eliminado.
- f. EnCronograma de ejecución de sincronización, paraFrecuencia— Con qué frecuenciaAmazon Kendrase sincronizará con su fuente de datos.
- g. Elija Siguiente.
8. En elEstablecer mapeos de campospágina, introduzca la siguiente información:
- a. Artículos de conocimiento,Catálogo de servicios,Adjuntos, ylncidentes—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon KendraaServiceNow

Debe especificar un JSON de[esquema de fuente de datos](#)utilizando el[TemplateConfiguration](#)API.
Debe proporcionar la siguiente información:

- Fuente de datos—Especifique la fuente de datos comoSERVICENOWV2.
- URL del servidor—Especifique elServiceNowversión de instancia host. Por ejemplo,*SU-dominio.service-now.com*.
- Tipo de autenticación—Especifique el tipo de autenticación, sibasicAuthoAuth2para tuServiceNowinstancia.

- ServiceNowversión de instancia—Especifique elServiceNowinstancia que está utilizando, ya seaTokyo,Sandiego,Rome, o0thers.
- Teclea—EspecificarTEMPLATEcomo el tipo cuando llamasCreateDataSource.
- Modo de sincronización—Especifique siAmazon Kendradebe actualizar el índice sincronizando todos los documentos o solo los documentos nuevos, modificados o eliminados. Puedes elegir entre:
 - FORCED_FULL_CRAWLpara volver a rastrear todo el contenido y reemplazar el contenido existente cada vez que la fuente de datos se sincronice con el índice
 - FULL_CRAWLpara rastrear de forma incremental solo el contenido nuevo, modificado y eliminado cada vez que la fuente de datos se sincronice con el índice
- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación que creó en suServiceNowcuenta.

Si utiliza la autenticación básica, el secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "username": "user name",  
    "password": "password"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- (Opcional): También puedes configurar el token OAuth2. Las credenciales de OAuth2 se almacenan como una cadena JSON enSecrets Managersecreto.

```
{  
    "username": "user name",  
    "password": "password",  
    "clientId": "client id",  
    "clientSecret": "client secret"  
}
```

- IAMpapel—EspecificarRoleArncuando llamasCreateDataSourcepara proporcionar unIAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para elServiceNowconector yAmazon Kendra. Para obtener más información, consulte[IAMfunciones paraServiceNowfuentes de datos](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfigurationcuando llamasCreateDataSource. Para obtener más información, consulte[ConfigurandoAmazon Kendrautilizar unAmazon VPC \(p. 465\)](#).
- Filtros de inclusión y exclusión—Puede especificar si desea incluir o excluir ciertos archivos adjuntos mediante los nombres de los archivos y los tipos de archivos de los artículos de conocimiento, los catálogos de servicios y los incidentes.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión,

solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Habilitar el rastreador de identidades: especifique si desea activar el rastreador de identidades. Si el rastreador de identidades está desactivado, debe cargar la información principal mediante el [elPutPrincipalMappingAPI](#). Rastrear la información de identidad de los usuarios y grupos con acceso a ciertos documentos resulta útil para filtrar el contexto de los usuarios. Los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos. Para obtener más información, consulte [Filtrado en contexto de usuario](#).
- Parámetros de indexación—También puede elegir especificar si desea:
 - Indexe artículos de conocimiento, catálogos de servicios e incidentes o todos ellos. Si opta por indexar artículos de conocimiento, elementos del catálogo de servicios e incidentes, debe proporcionar el nombre del ServiceNowcampo que está asignado al campo de contenido del documento de índice en Amazon Kendraíndice.
 - Indexe los archivos adjuntos a artículos de conocimiento, elementos del catálogo de servicios e incidentes.
 - Incluya artículos de conocimiento, elementos del catálogo de servicios e incidentes basados en la short descriptionpatrón de filtro.
 - Elija filtrar los elementos e incidentes del catálogo de servicios activos e inactivos.
 - Elija filtrar los incidentes según el tipo de incidente.
 - Elija qué entidades deben tener su ACL rastreada.
 - Puede utilizar un ServiceNowconsulta para especificar los documentos que desea de una o más bases de conocimiento, incluidas las bases de conocimiento privadas. El acceso a las bases de conocimiento lo determina el usuario que utilice para conectarse a la ServiceNowinstancia. Para obtener más información, consulte [Specifying documents to index with a query](#) (Especificar documentos a indexar con una consulta).
- Mapeos de campo—Elija mapear su ServiceNowcampos de fuente de datos para su Amazon Kendra campos de índice. Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para ServiceNowfuentes de datos](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [ServiceNowesquema de plantilla \(p. 267\)](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con tu ServiceNow fuente de datos, consulte:

- [Empezar con Amazon Kendra Anunciando la actualización ServiceNowconector \(V2\) para Amazon Kendra](#)

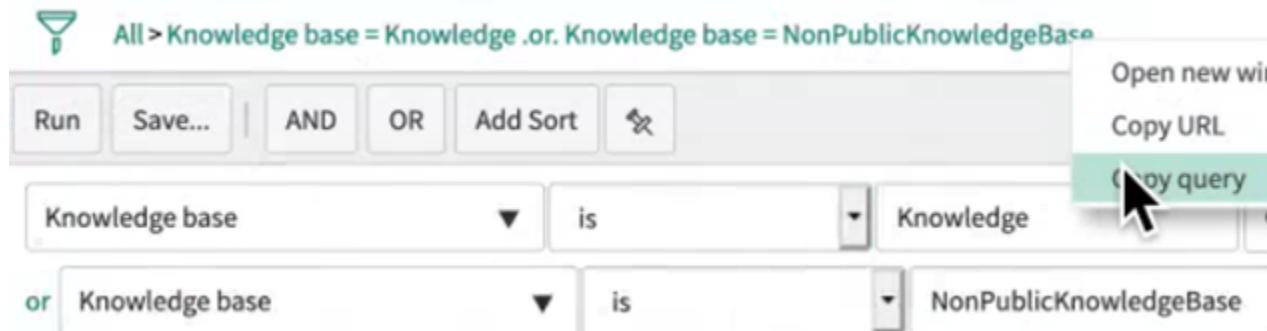
Especificar documentos para indexarlos con una consulta

Puede utilizar un ServiceNowconsulta para especificar los documentos que desea incluir en un Amazon Kendraíndice. Al utilizar una consulta, puede especificar varias bases de conocimiento, incluidas las bases de conocimiento privadas. El acceso a las bases de conocimiento lo determina el usuario que utilice para conectarse a la ServiceNowinstancia.

Para crear una consulta, utilice el ServiceNow generador de consultas. Puede utilizar el generador para crear la consulta y comprobar que la consulta devuelve la lista correcta de documentos.

Para crear una consulta mediante ServiceNow consola

1. Inicie sesión en la consola ServiceNow.
2. En el menú de la izquierda, selecciona Conocimiento, entonces Artículos, y la elección Todos.
3. En la parte superior de la página, selecciona el ícono del filtro.
4. Utilice el generador de consultas para crear la consulta.
5. Cuando la consulta esté completa, haga clic con el botón derecho en la consulta y elija Copiar consulta para copiar la consulta desde el generador de consultas. Guarde esta consulta para utilizarla en Amazon Kendra.



Asegúrese de no cambiar ningún parámetro de consulta al copiar la consulta. Si no se reconoce alguno de los parámetros de la consulta, ServiceNow trata el parámetro como vacío y no lo usa para filtrar los resultados.

Slack

Slack es una aplicación de comunicaciones empresariales que permite a los usuarios enviar mensajes y archivos adjuntos a través de varios canales públicos y privados. Puedes usar Amazon Kendra para indexar tus canales públicos y privados de Slack, enviar y archivar mensajes, archivos y archivos adjuntos, mensajes directos y grupales. También puedes elegir contenido específico para filtrar.

Puedes conectarte a Amazon Kendra a tu fuente de datos de Slack mediante el [Amazon Kendra console](#) y el [Slack Configuration API](#).

Para solucionar problemas de su Amazon Kendra Conector de fuente de datos de Slack, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 450\)](#)
- [Requisitos previos \(p. 451\)](#)
- [Instrucciones de conexión \(p. 452\)](#)
- [Más información \(p. 455\)](#)

Características admitidas

El conector de fuente de datos de Slack admite las siguientes funciones:

- Change log
- Mapeos de campo

- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de poder usar Amazon Kendra para indexar tu fuente de datos de Slack, realiza estos cambios en tu Slack y AWS cuentas.

En Slack, asegúrate de tener:

- Creó un token OAuth de usuario de Slack Bot o un token de OAuth de usuario de Slack. Puede elegir cualquiera de los dos tokens para conectarse a Amazon Kendra a tu fuente de datos de Slack. Consulte [Documentación de Slack sobre los tokens de acceso](#) para obtener más información.

Note

Si utilizas el token del bot como parte de tus credenciales de Slack, no puedes indexar los mensajes directos ni los mensajes de grupo, y debes añadir el token del bot al canal que quieras indexar.

- Apuntó el ID de tu equipo de espacio de trabajo de Slack en la URL de la página principal de tu espacio de trabajo de Slack. Por ejemplo, <https://app.slack.com/client/T0123456789/...> donde **T0123456789** es el identificador del equipo.
- Se agregaron los siguientes permisos [OAuth scopes/ read]:
 - canales:historial
 - canales:leer
 - grupos:historial
 - grupos:leer
 - im:history
 - im:leer
 - mpim:historia
 - mpim:leer
 - team:read
 - perfil de usuario: leer
 - usuarios:leer
 - emoji:leer
 - archivos:leer
 - grupos de usuarios: leer
- Has comprobado que cada documento es único en Slack y en otras fuentes de datos que piensas usar para el mismo índice. Cada fuente de datos que deseas utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En tu Cuenta de AWS, asegúrate de tener:

- [Creó un Amazon Kendra índice](#), si usa la API, anota el ID del índice.
- [Creó un IAM papel](#) para su fuente de datos y, si usa la API, anotó el ARN del IAM rol.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al correcto AWS Secrets Manager identificación secreta.

- Has guardado tus credenciales de autenticación de Slack en unAWS Secrets Managersecreto y, si usa la API, anota el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene unIAMrol o secreto, puedes usar la consola para crear un nuevoIAMrol ySecrets Managersecreto cuando conectas tu fuente de datos de Slack aAmazon Kendra. Si utiliza la API, debe proporcionar el ARN de una existenteIAMrol ySecrets Managersecreto y un identificador de índice.

Instrucciones de conexión

Para conectarAmazon Kendraa tu fuente de datos de Slack, debes proporcionar los detalles necesarios de tu fuente de datos de Slack para queAmazon Kendrapuede acceder a sus datos. Si aún no has configurado Slack paraAmazon Kendra, consulte[Requisitos previos \(p. 451\)](#).

Console

Para conectarAmazon Kendraa Slack

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).
2. En el panel de navegación de la izquierda, seleccionaÍndicesy, a continuación, elija el índice que deseé utilizar de la lista de índices.

Note

Puede elegir configurar o editar suControl de acceso de usuariosajustes enAjustes de índice.

3. En elPrimeros pasospágina, eligeAregar fuente de datos.
4. En elAregar fuente de datospágina, eligeConector Slacky, a continuación, elijaAregar fuente de datos.
5. En elEspecificardetalless de la fuente de datospágina, introduzca la siguiente información:
 - a. EnNombre y descripción, paraNombre de la fuente de datos—Introduzca un nombre para la fuente de datos. Puede incluir guiones pero no espacios.
 - b. (Opcional)Descripción—Introduzca una descripción opcional para la fuente de datos.
 - c. EnIdioma, paraldioma predeterminado—Un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. EnEtiquetas, paraAñadir etiqueta nueva—Etiquetas para buscar y filtrar tus recursos o realizar un seguimiento de tusAWScostos.
 - e. Elija Siguiente.
6. En elDefinad el acceso y la seguridadpágina, introduzca la siguiente información:
 - a. ID del equipo de Slack Workspace—El ID de equipo de tu espacio de trabajo de Slack.
 - b. AWS Secrets Managersecreto—Elige un secreto existente o crea uno nuevoSecrets Managersecreto para almacenar tus credenciales de autenticación de Slack. Si eliges crear un nuevo secreto, unAWS Secrets Managerse abre una ventana secreta.
 - i. Introduzca la siguiente información en elCrea unAWS Secrets Managerventana secreta:

- A. Nombre secreto—Un nombre para tu secreto. El prefijo 'AmazonKendra-Slack-' se añade automáticamente a tu nombre secreto.
- B. ParaToken de Slack—Introduce los valores de las credenciales de autenticación que creaste en tu cuenta de Slack.
 - ii. Seleccione Guardar.
- c. Nube privada virtual (VPC)—Puede optar por utilizar una VPC. Si es así, debe añadirSubredesyGrupos de seguridad de VPC.
- d. IAMpapel—Elige uno existenteIAMrol o crea uno nuevoIAMrol para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elijaCrear un nuevo rolpara evitar errores.

- e. Elija Siguiente.
7. En elConfigurar los ajustes de sincronizaciónpágina, introduzca la siguiente información:
 - a. Seleccione el tipo de contenido que desea rastrear—Las entidades o tipos de contenido de Slack que quieras rastrear.
 - b. Registro de cambios—Seleccione esta opción para actualizar el índice en lugar de sincronizar todos los archivos.
 - c. Patrones de expresiones regulares—Patrones de expresiones regulares para incluir o excluir ciertos archivos. Puede añadir hasta 100 patrones.
 - d. Selecciona la fecha de inicio del rastreo— ¿En qué fecha?Amazon Kendraempezará a rastrear tus datos desde.
 - e. EnCronograma de ejecución de sincronización, paraFrecuencia— Con qué frecuenciaAmazon Kendrase sincronizará con su fuente de datos.
 - f. Elija Siguiente.
8. En elDefinir mapeos de campospágina, introduzca la siguiente información:
 - a. ParaMapeos de campos de Slack—Seleccione una de lasAmazon Kendracampos de fuente de datos predeterminados generados que desea asignar a su índice.
 - b. Añadir campo—Para añadir campos de fuente de datos personalizados para crear un nombre de campo de índice para mapearlo y el tipo de datos del campo.
 - c. Elija Siguiente.
9. En elRevisar y crearpágina, compruebe que la información que ha introducido es correcta y, a continuación, seleccioneAregar fuente de datos. También puedes optar por editar tu información desde esta página. Su fuente de datos aparecerá enFuentes de datospágina después de que se haya agregado correctamente.

API

Para conectarAmazon Kendraa Slack

Debe especificar lo siguiente medianteSlackConfigurationAPI:

- ID del equipo de Slack Workspace—El ID del equipo de Slack que copiaste de la URL de la página principal de Slack.
- Lista de entidades para indexar—SiAmazon Kendradebe indexar tus canales públicos y privados, y tus mensajes grupales y directos.
- Fecha de rastreo—La fecha en la que empezarás a rastrear los datos de tu equipo de espacio de trabajo de Slack. La fecha debe seguir este formato:yyyy-mm-dd.

- Nombre secreto de recurso de Amazon (ARN)—Proporcione el nombre de recurso de Amazon (ARN) de unSecrets Managersecreto que contiene las credenciales de autenticación de tu cuenta de Slack. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "slackToken": "token"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. Esnose recomienda reutilizar las credenciales y los secretos en las fuentes de datos y en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMPapel—EspecificarRoleArn cuando llamasCreateDataSource para proporcionar un IAMrol con permisos para acceder a suSecrets Managersecreto y para llamar a las API públicas requeridas para el conector de Slack yAmazon Kendra. Para obtener más información, consulte[IAMfunciones para las fuentes de datos de Slack](#).

También puede añadir las siguientes funciones opcionales:

- Nube privada virtual (VPC)—EspecificarVpcConfiguration como parte de la configuración de la fuente de datos. Consulte[ConfigurandoAmazon Kendra usar una VPC](#).
- Registro de cambios—SiAmazon Kendradebe utilizar el mecanismo de registro de cambios en la fuente de datos de Slack para determinar si un documento debe agregarse, actualizarse o eliminarse del índice.

Note

Usa el registro de cambios si no quiereshazAmazon Kendrapara escanear todos los documentos. Si el registro de cambios es grande, podría tardarAmazon Kendramenos tiempo para escanear los documentos de la fuente de datos de Slack que para procesar el registro de cambios. Si sincronizas tu fuente de datos de Slack con tu índice por primera vez, se escanean todos los documentos.

- Filtros de inclusión y exclusión—Especifique si desea incluir o excluir ciertos canales públicos y privados, mensajes grupales y privados y mensajes archivados y de bots. Si utilizas un token de bot como parte de tus credenciales de autenticación de Slack, debes añadir el token de bot al canal que quieres indexar. No puedes indexar los mensajes directos ni los mensajes de grupo mediante un token de bot.

Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.

- Mapeos de campo—Elige asignar los campos de tu fuente de datos de Slack a tuAmazon Kendracampos de índice. Para obtener más información, consulte[Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario—Amazon Kendrarastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte[Filtrado de contexto de usuario para fuentes de datos de Slack](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con tu fuente de datos de Slack, consulta:

- [Descubre el conocimiento en los espacios de trabajo de Slack con la búsqueda inteligente mediante la Amazon Kendra Conector Slack](#)

Zendesk

Zendesk es un sistema de gestión de relaciones con los clientes que ayuda a las empresas a automatizar y mejorar las interacciones de atención al cliente. Puede utilizarlos Amazon Kendra para indexar sus tickets de soporte de Zendesk, los comentarios de los tickets, los archivos adjuntos de los tickets, los artículos del centro de ayuda, los comentarios de los artículos, los temas de la comunidad de guías, las publicaciones de la comunidad y los comentarios de las publicaciones de la comunidad.

Puedes filtrar por nombre de organización si quieres indexar los tickets que solo se encuentran dentro de una organización específica. También puede elegir establecer una fecha de rastreo para empezar a rastrear datos de Zendesk.

Puede conectarse Amazon Kendra a su fuente de datos de Zendesk mediante la [Amazon Kendra consola](#) y la [Template Configuration API](#).

Para solucionar problemas con el conector de fuente de datos de Amazon Kendra Zendesk, consulte [Solución de problemas de fuentes de datos \(p. 668\)](#).

Temas

- [Características admitidas \(p. 455\)](#)
- [Requisitos previos \(p. 455\)](#)
- [Instrucciones de conexión \(p. 456\)](#)
- [Más información \(p. 459\)](#)

Características admitidas

Amazon Kendra El conector de fuente de datos de Zendesk admite las siguientes funciones:

- Change log
- Mapeo de campo
- Filtrado de contexto de usuario
- Filtros de inclusión/exclusión
- Virtual Private Cloud (VPC) (Nube virtual privada [VPC])

Requisitos previos

Antes de que pueda utilizarla Amazon Kendra para indexar su fuente de datos de Zendesk, realice estos cambios en su cuenta y AWS en su cuenta de Zendesk.

En Zendesk, asegúrese de tener:

- Creé una cuenta administrativa de Zendesk Suite (Professional/Enterprise).
- Tomó nota de la URL de su host de Zendesk. Por ejemplo, [https://{}sub-domain \(https://{}host/\).zendesk.com/](https://{}sub-domain (https://{}host/).zendesk.com/).

Note

(En las instalaciones o en el servidor) Amazon Kendra comprueba si la información del punto final incluida en AWS Secrets Manager es la misma que la información del punto final especificada en los detalles de configuración de la fuente de datos. Esto ayuda a evitar el [confuso problema del adjunto](#), que es un problema de seguridad en el que un usuario no tiene permiso para realizar una acción, sino que Amazon Kendra lo utiliza como proxy para acceder al secreto configurado y realizar la acción. Si posteriormente cambias la información de tu terminal, debes crear un nuevo secreto para sincronizar esta información.

- Se generó un token de credenciales de OAuth 2.0 que contiene un ID de cliente, un secreto de cliente, un nombre de usuario y una contraseña. Consulte la [documentación de Zendesk sobre la generación de tokens de OAuth 2.0](#) para obtener más información.
- Se agregó el siguiente ámbito de OAuth 2.0:
 - leer
- Opcional: se ha instalado un certificado SSL para Amazon Kendra permitir la conexión.
- Se comprobó que cada documento es único en Zendesk y en otras fuentes de datos que piensa utilizar para el mismo índice. Cada fuente de datos que deseé utilizar para un índice no debe contener el mismo documento en todas las fuentes de datos. Los ID de los documentos son globales para un índice y deben ser únicos para cada índice.

En suCuenta de AWS, asegúrese de tener:

- [Creó un Amazon Kendra índice](#) y, si utilizaba la API, anotó el ID del índice.
- [Creó una IAM función](#) para la fuente de datos y, si utilizaba la API, anotó el ARN de la IAM función.

Note

Si cambias el tipo de autenticación y las credenciales, debes actualizar tu IAM rol para acceder al identificador AWS Secrets Manager secreto correcto.

- Almacenó sus credenciales de autenticación de Zendesk en AWS Secrets Manager secreto y, si utilizaba la API, anotó el ARN del secreto.

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. No se recomienda reutilizar las credenciales y los secretos en las fuentes de datos ni en las versiones 1.0 y 2.0 del conector (cuando corresponda).

Si no tiene un IAM rol o un secreto existentes, puede usar la consola para crear un IAM rol y un Secrets Manager secreto nuevos al conectar su fuente de datos de Zendesk. Amazon Kendra Si utiliza la API, debe proporcionar el ARN de un IAM rol y un Secrets Manager secreto existentes, así como un identificador de índice.

Instrucciones de conexión

Para conectarse Amazon Kendra a su fuente de datos de Zendesk, debe proporcionar los detalles necesarios de su fuente de datos de Zendesk para Amazon Kendra poder acceder a sus datos. Si aún no ha configurado Zendesk paraAmazon Kendra, consulte[Requisitos previos \(p. 455\)](#).

Console

Para conectarse Amazon Kendra a Zendesk

1. Inicie sesión en la AWS Management Console y abra la [consola de Amazon Kendra](#).

2. En el panel de navegación de la izquierda, elija Índices y, a continuación, el índice que deseé utilizar de la lista de índices.

Note

Puede configurar o editar los ajustes del control de acceso de usuario en la configuración del índice.

3. En la página de introducción, selecciona Agregar fuente de datos.
4. En la página Agregar fuente de datos, elija Zendesk connector y, a continuación, elija Agregar fuente de datos.
5. En la página Especificar detalles de la fuente de datos, introduzca la siguiente información:
 - a. En Nombre y descripción, en Nombre de fuente de datos: introduzca un nombre para la fuente de datos. Puede incluir guiones, pero no espacios.
 - b. Descripción (opcional): introduzca una descripción opcional para la fuente de datos.
 - c. En Idioma, en Idioma predeterminado: un idioma para filtrar los documentos para el índice. A menos que especifique lo contrario, el idioma predeterminado es el inglés. El idioma especificado en los metadatos del documento anula el idioma seleccionado.
 - d. En Etiquetas, seleccione Agregar nueva etiqueta: etiquetas para buscar y filtrar sus recursos o realizar un seguimiento de sus AWS costos.
 - e. Elija Siguiente.
6. En la página Definir acceso y seguridad, introduzca la siguiente información:
 - a. URL de Zendesk: introduzca su URL de Zendesk.
 - b. AWS Secrets Manager secreto: elija un secreto existente o cree uno nuevo Secrets Manager para almacenar sus credenciales de autenticación de Zendesk. Si decide crear un nuevo secreto, se abre una ventana AWS Secrets Manager secreta.
 - i. Introduzca la siguiente información en la ventana Crear un AWS Secrets Manager secreto:
 - A. Nombre secreto: un nombre para tu secreto. El prefijo 'AmazonKendra-Zendesk' se añade automáticamente a tu nombre secreto.
 - B. Para ID de cliente, secreto de cliente, nombre de usuario y contraseña: introduzca los valores de las credenciales de autenticación que creó en su cuenta de Zendesk.
 - ii. Seleccione Guardar.
 - c. Nube privada virtual (VPC): puedes elegir usar una VPC. Si es así, debe añadir subredes y grupos de seguridad de VPC.
 - d. IAMrol: elija un IAM rol existente o cree uno nuevo IAM para acceder a las credenciales del repositorio e indexar el contenido.

Note

IAMlos roles utilizados para los índices no se pueden usar para las fuentes de datos. Si no está seguro de si un rol existente se usa para un índice o preguntas frecuentes, elija Crear un rol nuevo para evitar errores.

- e. Elija Siguiente.
7. En la página Configurar los ajustes de sincronización, introduzca la siguiente información:
 - a. Seleccione entidades o tipos de contenido: las entidades de Zendesk o los tipos de contenido que desea rastrear.
 - b. Registro de cambios: seleccione esta opción para actualizar el índice en lugar de sincronizar todos los archivos.
 - c. Nombre de la organización: introduzca los nombres de las organizaciones de Zendesk para filtrar la sincronización.

- d. Fecha de inicio de sincronización: la fecha a partir de la cual deseas indexar el contenido.
 - e. Patrones de expresiones regulares: patrones de expresiones regulares para incluir o excluir ciertos archivos. Puede añadir hasta 100 patrones.
 - f. En Sincronizar programación de ejecución por frecuencia: elija la frecuencia con la Amazon Kendra que se sincronizará con la fuente de datos.
 - g. Elija Siguiente.
8. En la página Establecer mapeos de campos, introduzca la siguiente información:
 - a. Para entradas, comentarios de entradas, adjuntos de comentarios de entradas, artículo, comentario de artículo, adjunto de comentario de artículo, tema de la comunidad, publicación comunitaria, comentario de publicación comunitaria: seleccione uno de los campos de fuente de datos predeterminados Amazon Kendra generados que desee asignar a su índice.
 - b. Agregar campo: para agregar campos de fuente de datos personalizados para crear un nombre de campo de índice al que asignar y el tipo de datos del campo.
 - c. Elija Siguiente.
 9. En la página Revisar y crear, compruebe que la información que ha introducido sea correcta y, a continuación, seleccione Agregar fuente de datos. También puedes optar por editar tu información desde esta página. La fuente de datos aparecerá en la página Fuentes de datos cuando se haya agregado correctamente.

API

Para conectarse Amazon Kendra a Zendesk

Debe especificar un JSON del [esquema de la fuente de datos](#) mediante la [TemplateConfigurationAPI](#). Debe proporcionar la siguiente información:

- Fuente de datos: especifique la fuente de datos comoZENDESK.
- URL del host: proporcione la URL del host de Zendesk como parte de la configuración de la conexión o de los detalles del punto final del repositorio. Por ejemplo, <https://yoursubdomain.zendesk.com>.
- Tipo: especifique TEMPLATE como el tipo cuando llameCreateDataSource.
- Nombre de recurso secreto de Amazon (ARN): proporcione el nombre de recurso de Amazon (ARN) de un Secrets Manager secreto que contenga las credenciales de autenticación de su cuenta de Zendesk. El secreto se almacena en una estructura JSON con las siguientes claves:

```
{  
    "hostUrl": "https://yoursubdomain.zendesk.com",  
    "clientId": "client ID",  
    "clientSecret": "Zendesk client secret",  
    "userName": "Zendesk user name",  
    "password": "Zendesk password"  
}
```

Note

Se recomienda actualizar o rotar periódicamente las credenciales y el secreto. Proporcione solo el nivel de acceso necesario para su propia seguridad. No se recomienda reutilizar las credenciales y los secretos en las fuentes de datos ni en las versiones 1.0 y 2.0 del conector (cuando corresponda).

- IAMrol: especifique RoleArn cuándo llama CreateDataSource para proporcionar a un IAM rol los permisos para acceder a su Secrets Manager secreto y para llamar a las API públicas requeridas para el conector de Zendesk y. Amazon Kendra Para obtener más información, consulte las [IAMfunciones de las fuentes de datos de Zendesk](#).

También puede añadir las siguientes funciones opcionales:

- Virtual Private Cloud (VPC): especifique VpcConfiguration cuándo llama CreateDataSource Para obtener más información, consulte [Configurando Amazon Kendra para utilizar un Amazon VPC \(p. 465\)](#).
- Registro de cambios: si se Amazon Kendra debe utilizar el mecanismo de registro de cambios de la fuente de datos de Zendesk para determinar si un documento debe agregarse, actualizarse o eliminarse del índice.

Note

Usa el registro de cambios si no Amazon Kendra quieras escanear todos los documentos. Si el registro de cambios es grande, puede llevar Amazon Kendra menos tiempo escanear los documentos de la fuente de datos de Zendesk que procesar el registro de cambios. Si sincroniza su fuente de datos de Zendesk con su índice por primera vez, se escanean todos los documentos.

- Filtros de inclusión y exclusión: especifique si desea incluir o excluir:
 - Tickets de soporte, comentarios de tickets y/o adjuntos de comentarios de tickets
 - Artículos del centro de ayuda, archivos adjuntos y comentarios de artículos
 - Guía temas, publicaciones o comentarios de la comunidad
- Note

La mayoría de las fuentes de datos utilizan patrones de expresión regular, que son patrones de inclusión o exclusión denominados filtros. Si especifica un filtro de inclusión, solo se indexa el contenido que coincide con el filtro de inclusión. Los documentos que no coincidan con el filtro de inclusión no se indexan. Si especifica un filtro de inclusión y exclusión, los documentos que coincidan con el filtro de exclusión no se indexarán, aunque coincidan con el filtro de inclusión.
- Asignaciones de campos: elija asignar los campos de la fuente de datos de Zendesk a los campos de índice. Amazon Kendra Para obtener más información, consulte [Mapping data source fields](#) (Asignación de campos de origen de datos).
- Filtrado de contexto de usuario: Amazon Kendra rastrea la lista de control de acceso (ACL) de la fuente de datos de forma predeterminada. La información de la ACL se utiliza para filtrar los resultados de la búsqueda en función del acceso del usuario o su grupo a los documentos. Para obtener más información, consulte [Filtrado de contexto de usuario para fuentes de datos de Zendesk](#).

Para obtener una lista de otras claves JSON importantes para configurar, consulte [Esquema de plantillas de Zendesk \(p. 275\)](#).

Más información

Para obtener más información sobre la integración Amazon Kendra con su fuente de datos de Zendesk, consulte:

- [Descubra información de Zendesk con la búsqueda Amazon Kendra inteligente](#)

Asignación de campos de origen de datos

Puede asignar campos de documentos o contenido de la fuente de datos a los campos del índice. Por ejemplo, si tiene un campo en su fuente de datos denominado «departamento» que contiene la información del departamento de un documento, puede asignarlo a un campo de índice denominado «Departamento». De esta forma, puede utilizar el campo al consultar documentos.

También puedes mapear Amazon Kendra campos reservados o comunes, como `_created_at`. Si su fuente de datos tiene un campo llamado «fecha de creación», puede asignarlo al equivalente Amazon Kendra campo reservado llamado `_created_at`. Para obtener más información sobre Amazon Kendra campos reservados, consulte [Atributos o campos del documento](#).

Puede asignar campos a la mayoría de las fuentes de datos. Puede crear asignaciones de campos para los siguientes orígenes de datos:

- Adobe Experience Manager
- Al aire libre
- Amazon FSx
- Amazon RDS/Aurora
- Amazon Kendra Rastreador web
- Amazon WorkDocs
- Box
- Confluence
- Dropbox
- GitHub
- Unidades de Workspace de Google
- Gmail
- Jira
- Microsoft Exchange
- Microsoft OneDrive
- Microsoft SharePoint
- Equipos de Microsoft
- Microsoft Yammer
- Quip
- Salesforce
- ServiceNow
- Slack
- Zendesk

Si almacena los documentos en un depósito de S3 o en una fuente de datos de S3, especifique los campos mediante un archivo de metadatos JSON. Para obtener más información, consulte [Conector de fuente de datos S3](#).

La asignación de los campos de origen de datos a un campo de índice es un proceso de tres pasos:

1. Cree un índice. Para obtener más información, consulte [Crear un índice](#).
2. Actualice el índice para añadir campos.
3. Cree una fuente de datos e incluya asignaciones de campos para asignar los campos reservados y cualquier campo personalizado a Amazon Kendra campos de índice.

Para actualizar el índice para añadir campos personalizados, utilice la consola para editar las asignaciones de campos de la fuente de datos y añadir un campo personalizado o utilice la [UpdateIndex API](#). Puede añadir un total de 500 campos personalizados a su índice.

Para los orígenes de datos de la base de datos, si el nombre de la columna de la base de datos coincide con el nombre de un campo reservado, el campo y la columna se asignan automáticamente.

Con el[UpdateIndexAPI](#), puedes añadir campos reservados y personalizados usando[DocumentMetadataConfigurationUpdates](#).

En el siguiente ejemplo de JSON se utiliza DocumentMetadataConfigurationUpdates para agregar al índice un campo denominado «Department».

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

Al crear el campo, tiene la opción de configurar cómo se usa el campo para la búsqueda. Puede elegir entre las siguientes opciones:

- Visualizable: determina si el campo se devuelve en la respuesta a la consulta. El valor predeterminado es `true`.
- Tabla facetaria: indica que el campo se puede utilizar para crear facetas. El valor predeterminado es `false`.
- Se puede buscar: determina si el campo se utiliza en la búsqueda. El valor predeterminado es `true` para los campos de cadena y `false` para los campos de número y fecha.
- Clasificable: indica que el campo se puede utilizar para ordenar la respuesta de una consulta. Solo se puede configurar para campos de fecha, número y cadena. No se puede configurar para los campos de lista de cadenas.

En el siguiente ejemplo de JSON se utiliza DocumentMetadataConfigurationUpdates para agregar al índice un campo denominado «Department» y marcarlo como facetable.

```
"DocumentMetadataConfigurationUpdates": [  
    {  
        "Name": "Department",  
        "Type": "STRING_VALUE",  
        "Search": {  
            "Facetable": true  
        }  
    }  
]
```

UsoAmazon Kendracampos de documentos reservados o comunes

Con el[UpdateIndexAPI](#), puede crear campos reservados o comunes mediante[DocumentMetadataConfigurationUpdates](#) especificando elAmazon Kendranombre de campo de índice reservado para asignarlo a su atributo/nombre de campo de documento equivalente. También puede crear campos personalizados. Si utiliza un conector de fuente de datos, la mayoría incluye asignaciones de campos que asignan los campos del documento de fuente de datos aAmazon Kendracampos de índice. Si utiliza la consola, actualiza los campos seleccionando la fuente de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de mapeos de campos para configurar la fuente de datos.

Puede configurar el[Search](#) objeto para establecer un campo como visualizable, facetable, buscable y ordenable. Puede configurar el[Relevance](#) objeto para establecer el orden de clasificación de un campo, la duración del aumento o el período de tiempo que se aplicará a los valores de aumento, frescura, valor

de importancia y valores de importancia mapeados a valores de campo específicos. Si utiliza la consola, puede establecer la configuración de búsqueda de un campo seleccionando la opción de faceta en el menú de navegación. Para configurar el ajuste de relevancia, seleccione la opción de buscar en el índice en el menú de navegación, introduzca una consulta y utilice las opciones del panel lateral para ajustar la relevancia de la búsqueda. No puede cambiar el tipo de campo una vez creado el campo.

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

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

Para los campos personalizados, estos campos se crean mediante `DocumentMetadataConfigurationUpdates` con el `UpdateIndexAPI`, igual que cuando se crea un campo reservado o común. Debe establecer el tipo de datos adecuado para el campo personalizado. Si utiliza la consola, actualiza los campos seleccionando la fuente de datos, seleccionando la acción de edición y, a continuación, pasando a la sección de mapeos de campos para configurar la fuente de datos. Algunas fuentes de datos no admiten la adición de campos nuevos o personalizados. No puede cambiar el tipo de campo una vez creado el campo.

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

- Fecha
- Número
- Cadena
- Lista de cadenas

Si ha añadido documentos al índice utilizando `BatchPutDocument` API, `Attributes` enumera los campos/atributos de sus documentos y puede crear campos mediante el `DocumentAttribute` objeto.

Para documentos indexados desde un Amazon S3 fuente de datos, los campos se crean mediante un [Archivo de metadatos JSON](#) que incluye la información de los campos.

Si utiliza una base de datos compatible como fuente de datos, puede configurar los campos mediante la [opción de mapeos de campos](#).

Añadir documentos en idiomas distintos del inglés

Puede indexar documentos en varios idiomas. Si no especificas un idioma, Amazon Kendra indexa los documentos en inglés de forma predeterminada. Incluye el código de idioma de un documento en los metadatos del documento como campo. Consulte [Mapeos de campo y Atributos personalizados](#) para obtener más información sobre el `_language_code` campo para un documento.

Puede especificar el código de idioma de todos sus documentos en su fuente de datos cuando llame [CreateDataSource](#). Si un documento no tiene un código de idioma especificado en un campo de metadatos, el documento se indexa con el código de idioma especificado para todos los documentos en el nivel de la fuente de datos. En la consola, solo puede indexar documentos en un idioma compatible a nivel de fuente de datos. Ir a [Fuentes de datos](#), luego el [Especificificar los detalles de la fuente de datos](#) página, y elige un idioma del menú desplegable `Idioma`.

También puede buscar o consultar documentos en un idioma compatible. Para obtener más información, consulte [Búsqueda en idiomas](#).

Se admiten los siguientes idiomas y sus códigos (inglés o no se admite de forma predeterminada si no especificas un idioma). Esta tabla incluye los idiomas que Amazon Kendra admite la búsqueda semántica completa, así como los idiomas que solo admiten la coincidencia simple de palabras clave. Los idiomas que admiten la búsqueda semántica completa se marcan con un asterisco y aparecen en negrita en la siguiente tabla. El inglés (idioma predeterminado) también es compatible con la búsqueda semántica completa.

Nombre del idioma	Código de idioma
Árabe	ar
armenio	hy
vasco	eu
Bengalí	bn
Búlgaro	bg
catalán	ca
Chino: simplificado y tradicional*	zh
Checo	cs
Danés	da
Neerlandés	nl
Finés	fi
Francés: incluye francés (Canadá) *	fr
gallego	gl

Nombre del idioma	Código de idioma
Alemán*	de
Griego	el
Hindi	hi
Húngaro	hu
Indonesio	id
irlandés	ga
Italiano	it
Japonés*	ja
Coreano*	ko
Letón	lv
Lituano	lt
Noruego	no
Persa	fa
Portugués	pt
Portugués (Brasil) *	pt-BR
Rumano	ro
Ruso	ru
Sorani	ckb
Español: incluye español (Méjico) *	es
Sueco	sv
Turco	tr

*Se admite la búsqueda semántica para el idioma.

Para los idiomas que admiten la búsqueda semántica, se admiten las siguientes funciones.

- Relevancia del documento más allá de la simple coincidencia de palabras clave.
- Preguntas frecuentes más allá de la simple coincidencia de palabras clave.
- Extraer respuestas de documentos basándose en Amazon Kendra's comprensión lectora.
- Grupos de confianza (muy alto, alto, medio y bajo) de los resultados de la búsqueda.

Para los idiomas que no admiten la búsqueda semántica, se admite la coincidencia simple de palabras clave por motivos de relevancia del documento y preguntas frecuentes.

[sinónimos](#)(incluidos sinónimos personalizados),[aprendizaje incremental y retroalimentación](#), y [sugerencias de consultas](#) se admiten en inglés (idioma predeterminado).

ConfigurandoAmazon Kendrautilizar unAmazon VPC

Amazon Kendrapuede conectarse a suAmazon Virtual Private Cloudpara indexar el contenido almacenado en fuentes de datos o bases de datos que se ejecutan en su nube privada. Al crear un conector de fuente de datos o un conector de base de datos, proporciona identificadores de subred y grupo de seguridad para la subred que contiene la fuente de datos o la base de datos.Amazon Kendrautiliza esta información para crear una interfaz de red elástica que se utiliza para comunicarse de forma segura con la fuente de datos o la base de datos.

Si su fuente de datos o base de datos no se ejecuta en unAmazon VPC, puede conectar su fuente de datos o base de datos a suAmazon VPCmediante una red privada virtual (VPN). Al crear la cuenta de Amazon, se obtiene una VPC predeterminada. Para obtener información sobre la configuración de una VPN, consulte la[AWS VPNDocumentación](#).

Para utilizar una VPC, debe notificar a Amazon Kendra el identificador de la subred a la que pertenece la base de datos y los identificadores de cualquier grupo de seguridad que Amazon Kendra debe utilizar para acceder a la subred. Por ejemplo, si utiliza el puerto predeterminado para una base de datos MySQL, los grupos de seguridad deben permitirAmazon Kendrapara acceder al puerto 3306 del host que ejecuta la base de datos.

Selección de una subred para su fuente de datos o base de datos

Utilice subredes privadas únicamente en la configuración de VPC de su fuente de datos o base de datos. Si tu instancia de RDS está en una subred pública de tu VPC, no puedes usar esa subred directamente para sincronizar la fuente de datos o la base de datos. En su lugar, cree una subred privada que tenga acceso saliente a una gateway NAT en la subred pública. Al configurar la VPC para la fuente de datos o la base de datos, especifique esa subred privada. Para un origen de datos de base de datos configurado con una VPC, las subredes deben estar en uno de los siguientes ID de zona de disponibilidad:

- Oeste de EE. UU. (Oregón) —usw2-az1, usw2-az2, usw2-az3
- EE.UU. Este (N. Virginia) —use1-az1, use1-az2, use1-az4
- Europa (Irlanda) —euw1-az1, uew1-az2, euw1-az3

Los identificadores de subredes y grupos de seguridad se configuran en el panel de control de Amazon VPC. Para ver los identificadores, abra la consola de Amazon VPC como se indica a continuación:

Para ver los identificadores de subred

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon VPC en <https://console.aws.amazon.com/vpc/>.
2. En el panel de navegación, elija Subnets (Subredes).
3. En la lista de subredes, elija la subred que contiene el servidor de base de datos.
4. Anote el identificador en la pestaña de descripción en la pestaña de descripción Subnet ID (ID de subred).

Para ver los identificadores de grupos de seguridad

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon VPC en <https://console.aws.amazon.com/vpc/>.
2. En el panel de navegación, elija Security Groups (Grupos de seguridad).

3. En la lista de grupos de seguridad, elija el grupo para el que desea el identificador.
4. Anote el identificador en la pestaña de descripción en la pestaña de descripción Group ID (ID de grupo).

Si Amazon Kendra debe enrutar la conexión entre dos o más subredes, puede proporcionar varias subredes. Por ejemplo, si la subred que contiene el servidor de base de datos no tiene direcciones IP, Amazon Kendra puede conectarse a una subred con direcciones IP gratuitas y enrutar la conexión a la primera subred. Si enumera varias subredes, las subredes deben poder comunicarse entre sí. Cada subred debe estar asociada a una tabla de enrutamiento que proporcione acceso a Internet saliente mediante un dispositivo de traducción de direcciones de red (NAT).

También puede proporcionar varios grupos de seguridad. El efecto combinado de los grupos de seguridad debe permitir Amazon Kendra acceder a la fuente de datos o al servidor de base de datos que haya especificado en la configuración de conexión.

Conexión a una base de datos en una VPC

El siguiente ejemplo muestra cómo conectar una base de datos MySQL que se ejecuta en una VPC. En el ejemplo se supone que está empezando por la VPC predeterminada y que necesita crear una base de datos MySQL. Si ya tiene una VPC, asegúrese de que esté configurada como se muestra. Si tiene una base de datos MySQL, puede utilizarla en lugar de crear una nueva.

Temas

- [Paso 1: Configurar una VPC \(p. 466\)](#)
- [Paso 2: Configurar la seguridad \(p. 467\)](#)
- [Paso 3: Crear una base de datos \(p. 467\)](#)
- [Paso 4: Crear un conector de fuente de datos de base de datos \(p. 467\)](#)

Paso 1: Configurar una VPC

Configure su VPC para tener una subred privada y un grupo de seguridad para Amazon Kendra para obtener acceso a una base de datos MySQL que se ejecute en la subred. Las subredes proporcionadas en la configuración de VPC deben estar en Oeste de EE. UU. (Oregón), Este de EE. UU. (Norte de Virginia), Europa (Irlanda).

Para configurar una VPC

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon VPC en <https://console.aws.amazon.com/vpc/>.
2. En el panel de navegación, elija Route Tables (Tablas de enrutamiento) y, a continuación, elija Create route table (Crear tabla de enrutamiento).
3. En el campo Name tag (Nombre etiqueta), introduzca **Private subnet route table**. En VPC field (Campo de VPC), elija su VPC y, a continuación, elija Create (Crear). Elija Close (Cerrar) para volver a la lista de tablas de enrutamiento.
4. En el panel de navegación, elija NAT Gateways (Gateways NAT), Create NAT Gateway (Crear gateway NAT).
5. En el campo Subnet, elija la subred que es la subred pública y anote el ID de subred.
6. Si no tiene una dirección IP elástica, elija Create New EIP (Crear nueva EIP), elija Create a NAT Gateway (Crear gateway NAT) y, a continuación, elija Close (Cerrar).
7. En el panel de navegación, elija Route Tables (Tablas de enrutamiento).
8. En la lista de tablas de enrutamiento, elija la tabla de enrutamiento de la subred privada creada en el paso 3. En Actions (Acciones), elija Edit (Editar).

9. Seleccione Add route (Añadir ruta). Añada el destino 0.0.0.0/0 para permitir todo el tráfico saliente a Internet. En Target (Destino), elija NAT Gateway (Gateway NAT) y luego, el gateway creado en el paso 4. Elija Save routes (Guardar rutas) y, a continuación, elija Close (Cerrar).
10. En el menú Actions (Acciones), elija Edit subnet associations (Editar asociaciones de subred).
11. Elija las subredes que quiere que sean privadas. No elija la subred con la gateway NAT que ha indicado anteriormente.

Paso 2: Configurar la seguridad

A continuación, configure los grupos de seguridad para su base de datos.

Para crear grupos de seguridad

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon VPC en <https://console.aws.amazon.com/vpc/>.
2. En la descripción de la VPC, anote el CIDR de IPv4.
3. En el panel de navegación, elija Grupos de seguridad y luego elige Crear grupo de seguridad.
4. En Security group name (Nombre del grupo de seguridad), introduzca **DataSourceInboundSecurityGroup**. Proporcione una descripción y, a continuación, elija su VPC en la lista. Elija Create (Crear) y, a continuación, Close (Cerrar).
5. Elija la pestaña Inbound (Reglas entrantes).
6. Elija Edit rules (Editar reglas) y, a continuación, Add Rule (Añadir regla).
7. En una base de datos, escriba el número de puerto para Range Port (Rango de puertos). Por ejemplo, para MySQL es **3306**, y, para HTTPS, es **443**. Para Source (Origen), escriba el enrutamiento entre dominios sin clases (CIDR) de la VPC. Elija Save (Guardar) y, a continuación, elija Close (Cerrar).

El grupo de seguridad permite que cualquier persona de la VPC se conecte a la base de datos y permite conexiones salientes a Internet.

Paso 3: Crear una base de datos

Cree una base de datos para almacenar sus documentos o puede utilizar la base de datos existente. Consulte [Uso de una fuente de datos de base de datos](#) para obtener una lista de bases de datos que Amazon Kendrasoportes.

Para obtener instrucciones sobre cómo crear una base de datos MySQL, consulte [Cómo empezar con una fuente de datos de base de datos MySQL \(consola\) usando Amazon RDS](#).

Paso 4: Crear un conector de fuente de datos de base de datos

Tras configurar la VPC y crear la base de datos, puede crear un conector de fuente de datos para la base de datos. Consulte [Uso de una fuente de datos de base de datos](#).

Asegúrese de configurar la VPC, las subredes privadas que creó en la VPC y el grupo de seguridad que creó en la VPC para la base de datos.

Para obtener instrucciones sobre cómo crear una fuente de datos para una base de datos MySQL, consulte [Cómo empezar con una fuente de datos de base de datos MySQL \(consola\) usando Amazon RDS](#).

Eliminar un índice, una fuente de datos o documentos cargados por lotes

En esta sección se muestra cómo eliminar un índice, un repositorio de fuentes de datos con los documentos del índice o los documentos del índice que se han cargado por lotes.

Temas

- [Eliminar un índice \(p. 468\)](#)
- [Eliminar una fuente de datos \(p. 469\)](#)
- [Eliminar documentos subidos por lotes \(p. 470\)](#)

Eliminar un índice

Puede eliminar un índice desde el Amazon Kendra momento en que ya no lo utilice. Por ejemplo, elimina un índice cuando:

- Ya no utilizas el índice y quieras reducir los cargos a tu AWS cuenta. Un Amazon Kendra índice acumula cargos mientras está en ejecución, independientemente de si se realizan consultas en el índice o no.
- Desea volver a configurar el índice para una edición diferente de Amazon Kendra. Elimine el índice existente y, a continuación, cree uno nuevo con la edición diferente.
- Has alcanzado el número máximo de índices de tu cuenta y no quieras superar tu cuota. Elimine un índice existente y añada uno nuevo. Para obtener información sobre el número máximo de índices que puede crear, consulte [Cuotas](#).

Para eliminar un índice, utilice la consola AWS Command Line Interface, el AWS CloudFormation script o la `DeleteIndex` API. Al eliminar un índice, se elimina el índice y todas las fuentes de datos y los datos del documento asociados. Al eliminar un índice, los documentos originales no se eliminan del almacenamiento.

La eliminación de un índice es una operación asíncrona. Al empezar a eliminar un índice, el estado del índice cambia a `DELETING`. Permanece en el `DELETING` estado hasta que se elimine toda la información relacionada con el índice. Una vez que se elimina el índice, ya no aparece en los resultados de una llamada a la `ListIndices` API. Si llamas a la `DescribeIndex` API con el identificador del índice eliminado, recibirás `ResourceNotFoundException` una excepción.

Para eliminar un índice (consola)

1. Inicie sesión AWS Management Console y abra la Amazon Kendra consola en <https://console.aws.amazon.com/kendra/>.
2. En el panel de navegación, seleccione Índices y, a continuación, el índice que deseé eliminar.
3. Elija Eliminar para eliminar el índice seleccionado.

Para eliminar un índice (CLI)

- En la AWS CLI utilice el siguiente comando. El comando está formateado para Linux y macOS. Si utiliza Windows, sustituya el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

```
aws kendra delete-index \  
  --id index-id
```

Eliminar una fuente de datos

Se elimina una fuente de datos cuando se desea eliminar del Amazon Kendra índice la información que contiene la fuente de datos. Por ejemplo, elimine una fuente de datos cuando:

- Una fuente de datos no está configurada correctamente. Elimine la fuente de datos, espere a que la fuente de datos termine de eliminarse y, a continuación, vuelva a crearla.
- Ha migrado documentos de una fuente de datos a otra. Elimine la fuente de datos original y vuelva a crearla en la nueva ubicación.
- Ha alcanzado el límite de fuentes de datos para un índice. Elimine una de las fuentes de datos existentes y añada una nueva. Para obtener más información sobre la cantidad de fuentes de datos que puede crear, consulte[Cuotas \(p. 663\)](#).

Para eliminar una fuente de datos, utilice la consola, la AWS Command Line Interface (AWS CLI), la [DeleteDataSource API](#) o un AWS CloudFormation script. Al eliminar una fuente de datos, se quita del índice toda la información sobre la fuente de datos. Si solo desea detener la sincronización de la fuente de datos, cambie el programa de sincronización de la fuente de datos para que se ejecute bajo demanda.

Eliminar una fuente de datos es una operación asíncrona. Al empezar a eliminar una fuente de datos, el estado de la fuente de datos cambia aDELETING. Permanece en ese DELETING estado hasta que se elimine la información relacionada con la fuente de datos. Una vez eliminada la fuente de datos, ya no aparece en los resultados de una llamada a la [ListDataSourcesAPI](#). Si llamas a la [DescribeDataSourceAPI](#) con el identificador de la fuente de datos eliminada, recibirás una ResourceNotFound excepción.

Note

Eliminar una fuente de datos completa o volver a sincronizar el índice después de eliminar documentos específicos de una fuente de datos puede tardar hasta una hora o más, según la cantidad de documentos que deseé eliminar.

Para eliminar una fuente de datos (consola)

1. Inicie sesión AWS Management Console y abra la Amazon Kendra consola en <https://console.aws.amazon.com/kendra/>.
2. En el panel de navegación, elija Índices y, a continuación, el índice que contiene la fuente de datos que desea eliminar.
3. En el panel de navegación, elija Data source (Orígenes de datos).
4. Elija la fuente de datos que deseé eliminar.
5. Elija Eliminar para eliminar la fuente de datos.

Para eliminar una fuente de datos (CLI)

- En la AWS Command Line Interface utilice el siguiente comando. El comando está formateado para Linux y macOS. Si utiliza Windows, sustituya el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

```
aws kendra delete-data-source \  
  --id data-source-id \  
  --index-id index-id
```

Al eliminar una fuente de datos, Amazon Kendra elimina toda la información almacenada sobre la fuente de datos. Amazon Kendra elimina todos los datos del documento almacenados en el índice y todos los historiales de ejecución y las métricas asociados a la fuente de datos. La eliminación de una fuente de datos no elimina los documentos originales del almacenamiento.

Los documentos de la fuente de datos pueden incluirse en el recuento de documentos devuelto por la `DescribeIndex` API mientras Amazon Kendra se elimina una fuente de datos. Los documentos de la fuente de datos pueden aparecer en los resultados de la búsqueda mientras Amazon Kendra se elimina la fuente de datos.

Amazon Kendra libera los recursos de una fuente de datos en cuanto llamas a la `DeleteDataSource` API o decides eliminar la fuente de datos en la consola. Si va a eliminar la fuente de datos para reducir la cantidad de fuentes de datos por debajo de su límite, puede crear una nueva fuente de datos de inmediato.

Si va a eliminar una fuente de datos y luego crear otra fuente de datos para los datos del documento, espere a que se elimine la primera fuente de datos antes de sincronizar la nueva fuente de datos.

Puede eliminar una fuente de datos con la que se esté sincronizando Amazon Kendra. Se detiene la sincronización y se elimina la fuente de datos. Si intenta iniciar una sincronización cuando se elimina la fuente de datos, se produce una `ConflictException` excepción.

No puede eliminar una fuente de datos si el índice asociado está en ese `DELETING` estado. Al eliminar un índice, se eliminan todas las fuentes de datos del índice. Puede empezar a eliminar un índice mientras la fuente de datos de ese índice esté en ese `DELETING` estado.

Si tiene dos fuentes de datos que apuntan a los mismos documentos, como dos fuentes de datos que apuntan al mismo Amazon S3 depósito, es posible que los documentos del índice no sean coherentes si se elimina una de las fuentes de datos. Cuando dos fuentes de datos hacen referencia a los mismos documentos, solo se almacena una copia de los datos del documento en el índice. Al eliminar una fuente de datos, se eliminan los datos de índice de los documentos. La otra fuente de datos no sabe que los documentos se han eliminado, por lo que no Amazon Kendra volverá a indexarlos correctamente la próxima vez que se sincronice. Cuando tenga dos fuentes de datos que apunten a la misma ubicación del documento, debe eliminar ambas fuentes de datos y, a continuación, volver a crear una.

Eliminar documentos subidos por lotes

Puede eliminar documentos directamente de un índice mediante la [BatchDeleteDocument](#) API. No puedes eliminar documentos directamente desde la consola. Si usa la consola, puede eliminar documentos específicos del repositorio de fuentes de datos y volver a sincronizarlos con el índice o eliminar todo el conector de la fuente de datos.

Eliminar documentos de un índice mediante `BatchDeleteDocument` una operación asíncrona. Después de llamar a la `BatchDeleteDocument` API, la [BatchGetDocumentStatus](#) utilizarás para supervisar el proceso de eliminación de tus documentos. Cuando se elimina un documento del índice, `NOT_FOUND` se Amazon Kendra devuelve como estado.

Note

Eliminar documentos de un índice utilizando `BatchDeleteDocument` puede tardar hasta una hora o más, según la cantidad de documentos que desee eliminar.

Para eliminar documentos cargados por lotes de un índice (CLI)

- En la AWS Command Line Interface utilice el siguiente comando. El comando está formateado para Linux y macOS. Si utiliza Windows, sustituya el carácter de continuación de línea de Unix (\) por un signo de intercalación (^).

```
aws kendra batch-delete-document \
```

```
--index-id index-id \
--document-id-list 'doc-id-1' 'doc-id-2'
```

Enriquecer sus documentos durante la ingestión

Puede modificar los campos o atributos de los metadatos del contenido y del documento durante el proceso de ingesta del documento. Con Amazon Kendra la función de enriquecimiento personalizado de documentos, puede crear, modificar o eliminar los atributos y el contenido del documento al introducir sus documentos. Amazon Kendra Esto significa que puedes manipular e ingerir sus datos según lo necesite.

Esta función le permite controlar la forma en que se tratan y se Amazon Kendra ingieren sus documentos. Por ejemplo, puede eliminar la información de identificación personal en los metadatos del documento mientras se ingieren los documentos en Amazon Kendra.

Otra forma de utilizar esta función es invocar una función Lambda AWS Lambda para ejecutar el reconocimiento óptico de caracteres (OCR) en las imágenes, la traducción del texto y otras tareas a fin de preparar los datos para la búsqueda o el análisis. Por ejemplo, puede invocar una función para ejecutar OCR en imágenes. La función podría interpretar el texto de las imágenes y tratar cada imagen como un documento textual. Una empresa que recibe encuestas de clientes enviadas por correo y las almacena como imágenes, podría ingerirlas como documentos textuales en Amazon Kendra. A continuación, la empresa puede buscar información valiosa de la encuesta de clientes en Amazon Kendra.

Puede utilizar operaciones básicas para aplicarlas como primer análisis de los datos y, a continuación, utilizar una función Lambda para aplicar operaciones más complejas a los datos. Por ejemplo, puede utilizar una operación básica para eliminar simplemente todos los valores del campo de metadatos del documento «Customer_ID» y, a continuación, aplicar una función Lambda para extraer texto de las imágenes del texto de los documentos.

Cómo funciona Custom Document Enrichment

El proceso general de Custom Document Enrichment es el siguiente:

1. Custom Document Enrichment se configura al crear o actualizar el origen de datos o indexar los documentos directamente en Amazon Kendra.
2. Amazon Kendra aplica configuraciones en línea o lógica básica para modificar los datos. Para obtener más información, consulte [the section called “Operaciones básicas para cambiar los metadatos” \(p. 473\)](#).
3. Si elige configurar la manipulación avanzada de datos, Amazon Kendra puede aplicarlo en sus documentos originales, sin procesar o en los documentos estructurados y analizados. Para obtener más información, consulte [the section called “Funciones lambda: extraer y cambiar metadatos o contenido” \(p. 478\)](#).
4. Los documentos modificados se ingieren en Amazon Kendra.

En cualquier momento de este proceso, si la configuración no es válida, Amazon Kendra arroja un error.

Cuando llamas a las [BatchPutDocument API](#) [CreateDataSourceUpdateDataSource](#), proporcionas tu configuración personalizada de enriquecimiento de documentos. Si se llama a BatchPutDocument, se debe configurar Custom Document Enrichment con cada solicitud. Si utiliza la consola, seleccione el índice y, a continuación, seleccione Document enrichments (Enriquecimientos de documentos) para configurar Custom Document Enrichment.

Si utilizas el enriquecimiento de documentos en la consola, puedes configurar solo las operaciones básicas o solo las funciones de Lambda o ambas, como puedes hacer con la API. Puede seleccionar Siguiente

en los pasos de la consola para optar por no configurar las operaciones básicas y solo las funciones de Lambda, incluida la opción de aplicarlas a los datos originales (previa a la extracción) o estructurados (posterior a la extracción). Solo puede guardar las configuraciones si completa todos los pasos de la consola. Las configuraciones de sus documentos no se guardarán si no completa todos los pasos.

Operaciones básicas para cambiar los metadatos

Puede manipular los campos y el contenido del documento mediante la lógica básica. Esto incluye la eliminación de valores de un campo, la modificación de los valores de un campo mediante una condición o la creación de un campo. Para manipulaciones avanzadas que van más allá de lo que puedes manipular con la lógica básica, invoque una función Lambda. Para obtener más información, consulte [the section called "Funciones lambda: extraer y cambiar metadatos o contenido" \(p. 478\)](#).

Para aplicar la lógica básica, especifique el campo de destino que desea manipular mediante el [DocumentAttributeTarget](#) objeto. Proporcione la clave de atributo. Por ejemplo, la clave «Department» es un campo o atributo que contiene todos los nombres de departamento asociados a los documentos. También puede especificar un valor que se utilizará en el campo de destino si se cumple una condición determinada. La condición se establece mediante el [DocumentAttributeCondition](#) objeto. Por ejemplo, si el campo «source_URL» contiene «financiero» en su valor URI, rellene previamente el campo de destino «Department» con el valor objetivo «Finance» del documento. También puede eliminar los valores del atributo del documento de destino.

Para aplicar la lógica básica mediante la consola, seleccione el índice y, a continuación, seleccione Document enrichments(Enriquecimiento de documentos) en el menú de navegación. Vaya a Configurar operaciones básicas para aplicar manipulaciones básicas a los campos y al contenido del documento.

El siguiente es un ejemplo del uso de la lógica básica para eliminar todos los números de identificación del cliente del campo del documento denominado «Customer_ID».

Ejemplo 1: Eliminar los números de identificación del cliente asociados a los documentos

Datos antes de aplicar la manipulación básica.

Document_ID	Body_Text	Customer_ID
1	Lorem Ipsum.	CID1234
2	Lorem Ipsum.	CID1235
3	Lorem Ipsum.	CID1236

Datos después de aplicar la manipulación básica.

Document_ID	Body_Text	Customer_ID
1	Lorem Ipsum.	
2	Lorem Ipsum.	
3	Lorem Ipsum.	

El siguiente es un ejemplo del uso de la lógica básica para crear un campo denominado «Departamento» y llenar previamente este campo con los nombres de los departamentos según la información del campo «source_URL». Por ejemplo, si el campo «source_URL» contiene «financial» en su valor URI, rellene previamente el campo de destino «Department» con el valor objetivo «Finance» para el documento.

Ejemplo 2: Crear el campo «Departamento» y rellenarlo previamente con los nombres de los departamentos asociados a los documentos mediante una condición.

Datos antes de aplicar la manipulación básica.

Document_ID	Body_Text	URI de origen
1	Lorem Ipsum.	financial/1
2	Lorem Ipsum.	financial/2
3	Lorem Ipsum.	financial/3

Datos después de aplicar la manipulación básica.

Document_ID	Body_Text	URI de origen	Department
1	Lorem Ipsum.	financial/1	Finance
2	Lorem Ipsum.	financial/2	Finance
3	Lorem Ipsum.	financial/3	Finance

Note

Amazon Kendrano puede crear un campo de documento de destino si aún no se ha creado como campo de índice. Después de crear el campo de índice, puede crear un campo de documento mediante `DocumentAttributeTarget`. Amazon Kendraluego asigna el campo de metadatos del documento recién creado a su campo de índice.

El código siguiente es un ejemplo de configuración de la manipulación básica de datos para eliminar los números de identificación de clientes asociados a los documentos.

Console

Para configurar la manipulación básica de datos para eliminar números de identificación de clientes

1. En el panel de navegación izquierdo, en Indexes (Índices), seleccione Document enrichments (Enriquecimiento de documentos) y luego seleccione Add document enrichment (Añadir enriquecimiento de documentos).
2. En la página Configurar operaciones básicas, elija en el menú desplegable la fuente de datos en la que desea modificar los campos y el contenido del documento. A continuación, elija en el menú desplegable el nombre del campo de documento «Customer_ID», seleccione en el menú desplegable el nombre del campo de índice «Customer_ID» y seleccione en el menú desplegable la acción de destino Delete (Eliminar). A continuación, seleccione Add basic operation (Añadir operación básica).

CLI

Para configurar la manipulación básica de datos para eliminar números de identificación de clientes

```
aws kendra create-data-source \
--name data-source-name \
--index-id index-id \
--role-arn arn:aws:iam::account-id:role/role-name \
```

```
--type S3 \
--configuration '{"S3Configuration":{"BucketName":"'${S3-bucket-name}'"}' \
--custom-document-enrichment-configuration '{"InlineConfigurations":[{"Target": \
{"TargetDocumentAttributeKey":"Customer_ID", "TargetDocumentAttributeValueDeletion": \
true}]}]'
```

Python

Para configurar la manipulación básica de datos para eliminar números de identificación de clientes

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

kendra = boto3.client("kendra")

print("Create a data source with customizations")

# Provide the name of the data source
name = "data-source-name"
# Provide the index ID for the data source
index_id = "index-id"
# Provide the IAM role ARN required for data sources
role_arn = "arn:aws:iam::${account-id}:role/${role-name}"
# Provide the data source connection information
data_source_type = "S3"
S3_bucket_name = "S3-bucket-name"
# Configure the data source with Custom Document Enrichment
configuration = {"S3Configuration":
{
    "BucketName": S3_bucket_name
}
}
custom_document_enrichment_configuration = {"InlineConfigurations": [
{
    "Target": {"TargetDocumentAttributeKey": "Customer_ID",
               "TargetDocumentAttributeValueDeletion": True}
}
]
}

try:
    data_source_response = kendra.create_data_source(
        Name = name,
        IndexId = index_id,
        RoleArn = role_arn,
        Type = data_source_type
        Configuration = configuration
        CustomDocumentEnrichmentConfiguration =
        custom_document_enrichment_configuration
    )

    pprint.pprint(data_source_response)

    data_source_id = data_source_response["Id"]

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

    while True:
        # Get the details of the data source, such as the status
        data_source_description = kendra.describe_data_source(
            Id = data_source_id,
            IndexId = index_id
        )

```

```
status = data_source_description["Status"]
print(" Creating data source. Status: "+status)
time.sleep(60)
if status != "CREATING":
    break

print("Synchronize the data source.")

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

pprint.pprint(sync_response)

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

while True:

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

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

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

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

print("Program ends.")
```

Java

Para configurar la manipulación básica de datos para eliminar números de identificación de clientes

```
package com.amazonaws.kendra;

import java.util.concurrent.TimeUnit;
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;
import software.amazon.awssdk.services.kendra.model.DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.DataSourceStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJob;
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJobStatus;
import software.amazon.awssdk.services.kendra.model.DataSourceType;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceRequest;
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceResponse;
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;
import software.amazon.awssdk.services.kendra.model.IndexStatus;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;
import software.amazon.awssdk.services.kendra.model.S3DataSourceConfiguration;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;
```

```
public class CreateDataSourceWithCustomizationsExample {

    public static void main(String[] args) throws InterruptedException {
        System.out.println("Create a data source with customizations");

        String dataSourceName = "data-source-name";
        String indexId = "index-id";
        String dataSourceRoleArn = "arn:aws:iam::account-id:role/role-name";
        String s3BucketName = "S3-bucket-name"

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

        CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
            .builder()
            .name(dataSourceName)
            .description(experienceDescription)
            .roleArn(experienceRoleArn)
            .type(DataSourceType.S3)
            .configuration(
                DataSourceConfiguration
                    .builder()
                    .s3Configuration(
                        S3DataSourceConfiguration
                            .builder()
                            .bucketName(s3BucketName)
                            .build()
                    ).build()
            )
            .customDocumentEnrichmentConfiguration(
                CustomDocumentEnrichmentConfiguration
                    .builder()
                    .inlineConfigurations(Arrays.asList(
                        InlineCustomDocumentEnrichmentConfiguration
                            .builder()
                            .target(
                                DocumentAttributeTarget
                                    .builder()
                                    .targetDocumentAttributeKey("Customer_ID")
                                    .targetDocumentAttributeValueDeletion(true)
                                    .build()
                            ).build()
                    )).build();
            )

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

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

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

            DataSourceStatus status = describeDataSourceResponse.status();
            System.out.println(String.format("Creating data source. Status: %s",
            status));
            TimeUnit.SECONDS.sleep(60);
            if (status != DataSourceStatus.CREATING) {
```

```
        break;
    }

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

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

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

        TimeUnit.SECONDS.sleep(60);
        if (job.status() != DataSourceSyncJobStatus.SYNCING) {
            break;
        }
    }

    System.out.println("Data source creation with customizations is complete");
}
}
```

Funciones lambda: extraer y cambiar metadatos o contenido

Puede manipular los campos y el contenido del documento mediante las funciones de Lambda. Esto resulta útil si desea ir más allá de la lógica básica y aplicar manipulaciones avanzadas de datos. Por ejemplo, mediante el reconocimiento óptico de caracteres (OCR), que interpreta el texto de las imágenes y trata cada imagen como un documento textual. O bien, recuperar la fecha y hora actual en una zona horaria determinada e insertar la fecha y hora donde haya un valor vacío para un campo de fecha.

Puede aplicar primero la lógica básica y, a continuación, utilizar una función Lambda para seguir manipulando los datos, o viceversa. También puede optar por aplicar solo una función Lambda.

Amazon Kendrapuede invocar una función Lambda para aplicar manipulaciones de datos avanzadas durante el proceso de ingestión como parte de su [CustomDocumentEnrichmentConfiguration Especifica una función que incluye el permiso para ejecutar la función Lambda y acceder a su Amazon S3 bucket para almacenar el resultado de las manipulaciones IAM de datos; consulte las funciones de acceso.](#)

Amazon Kendra puede aplicar una función Lambda en sus documentos originales sin procesar o en los documentos estructurados y analizados. Puede configurar una función Lambda que tome sus datos originales o sin procesar y aplique sus manipulaciones de datos mediante. [PreExtractionHookConfiguration](#) También puede configurar una función Lambda que tome sus documentos estructurados y aplique sus manipulaciones de datos mediante. [PostExtractionHookConfiguration](#) Amazon Kendra extrae los metadatos y el texto del documento para estructurar los documentos. Sus funciones Lambda deben seguir las estructuras obligatorias de solicitud y respuesta. Para obtener más información, consulte [the section called "Contratos de datos para funciones Lambda" \(p. 484\)](#).

Para configurar una función Lambda en la consola, seleccione el índice y, a continuación, seleccione Document enrichments(Enriquecimiento de documentos) en el menú de navegación. Vaya a Configure Lambda functions (Configurar funciones Lambda) para configurar una función Lambda.

Solo puede configurar una función Lambda para PreExtractionHookConfiguration y solo una función Lambda para PostExtractionHookConfiguration. Sin embargo, la función Lambda puede invocar otras funciones que requiere. Puede configurar ambos PreExtractionHookConfiguration y PostExtractionHookConfiguration, o cualquiera de los dos. La función Lambda para PreExtractionHookConfiguration no debe exceder un tiempo de ejecución de 5 minutos y su función Lambda para PostExtractionHookConfiguration no debe exceder un tiempo de ejecución de 1 minuto. La configuración de Custom Document Enrichment tarda más tiempo en ingerir sus documentos de forma natural en Amazon Kendra que si no lo configurara.

Puede configurar Amazon Kendra para invocar una función Lambda solo si se cumple una condición. Por ejemplo, puede especificar una condición que si hay valores de fecha y hora vacíos, Amazon Kendra debería invocar una función que insertara la fecha y hora actuales.

A continuación se muestra un ejemplo de uso de una función Lambda para ejecutar OCR para interpretar texto de imágenes y almacenar este texto en un campo denominado «Document_Image_Text».

Ejemplo 1: Extraer texto de imágenes para crear documentos textuales

Datos antes de aplicar la manipulación avanzada.

Document_ID	Document_Image
1	image_1.png
2	image_2.png
3	image_3.png

Datos después de aplicar la manipulación avanzada.

Document_ID	Document_Image	Document_Image_Text
1	image_1.png	Mailed survey response
2	image_2.png	Mailed survey response
3	image_3.png	Mailed survey response

A continuación se muestra un ejemplo de uso de una función Lambda para insertar la fecha y hora actual para valores de fecha vacíos. Utiliza la condición de que si el valor de un campo de fecha es «null», se sustituye por la fecha y hora actuales.

Ejemplo 2: Sustituir los valores vacíos del campo Last_Updated por la fecha y hora actual.

Datos antes de aplicar la manipulación avanzada.

Document_ID	Body_Text	Last_Updated
1	Lorem Ipsum.	January 1, 2020
2	Lorem Ipsum.	
3	Lorem Ipsum.	July 1, 2020

Datos después de aplicar la manipulación avanzada.

Document_ID	Body_Text	Last_Updated
1	Lorem Ipsum.	January 1, 2020
2	Lorem Ipsum.	December 1, 2021
3	Lorem Ipsum.	July 1, 2020

El siguiente código es un ejemplo de configuración de una función Lambda para la manipulación avanzada de datos en los datos originales y sin procesar.

Console

Para configurar una función Lambda para la manipulación avanzada de datos en los datos originales sin procesar

1. En el panel de navegación izquierdo, en Indexes (Índices), seleccione Document enrichments (Enriquecimiento de documentos) y luego seleccione Add document enrichment (Añadir enriquecimiento de documentos).
2. En la página Configure Lambda functions (Configurar funciones Lambda), en la sección Lambda for pre-extraction (Lambda para preextracción), seleccione en los menús desplegables su ARN de la función Lambda y su bucket de Amazon S3. Agregue su rol de IAM acceso seleccionando la opción para crear un nuevo rol en el menú desplegable. Esto crea los Amazon Kendra permisos necesarios para crear el enriquecimiento del documento.

CLI

Para configurar una función Lambda para la manipulación avanzada de datos en los datos originales sin procesar

```
aws kendra create-data-source \
--name data-source-name \
--index-id index-id \
--role-arn arn:aws:iam::account-id:role/role-name \
--type S3 \
--configuration '{"S3Configuration":{"BucketName":"S3-bucket-name"}' \
--custom-document-enrichment-configuration '{"PreExtractionHookConfiguration": \
{"LambdaArn":"arn:aws:iam::account-id:function/function-name", "S3Bucket":"S3-bucket-name", "RoleArn": "arn:aws:iam:account-id:role/cde-role-name"}}
```

Python

Para configurar una función Lambda para la manipulación avanzada de datos en los datos originales sin procesar

```
import boto3
```

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

kendra = boto3.client("kendra")

print("Create a data source with customizations.")

# Provide the name of the data source
name = "data-source-name"
# Provide the index ID for the data source
index_id = "index-id"
# Provide the IAM role ARN required for data sources
role_arn = "arn:aws:iam::${account-id}:role/${role-name}"
# Provide the data source connection information
data_source_type = "S3"
S3_bucket_name = "S3-bucket-name"
# Configure the data source with Custom Document Enrichment
configuration = {"S3Configuration":
    {
        "BucketName": S3_bucket_name
    }
}
custom_document_enrichment_configuration = {"PreExtractionHookConfiguration":
    {
        "LambdaArn": "arn:aws:iam::account-id:function/function-name",
        "S3Bucket": "S3-bucket-name"
    }
}
"RoleArn": "arn:aws:iam::account-id:role/cde-role-name"
}

try:
    data_source_response = kendra.create_data_source(
        Name = name,
        IndexId = index_id,
        RoleArn = role_arn,
        Type = data_source_type
        Configuration = configuration
        CustomDocumentEnrichmentConfiguration =
    custom_document_enrichment_configuration
    )

    pprint.pprint(data_source_response)

    data_source_id = data_source_response["Id"]

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

    while True:
        # Get the details of the data source, such as the status
        data_source_description = kendra.describe_data_source(
            Id = data_source_id,
            IndexId = index_id
        )
        status = data_source_description["Status"]
        print(" Creating data source. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

    print("Synchronize the data source.")

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

```
)  
  
pprint.pprint(sync_response)  
  
print("Wait for the data source to sync with the index.")  
  
while True:  
  
    jobs = kendra.list_data_source_sync_jobs(  
        Id = data_source_id,  
        IndexId = index_id  
    )  
  
    # For this example, there should be one job  
    status = jobs["History"][0]["Status"]  
  
    print(" Syncing data source. Status: "+status)  
    time.sleep(60)  
    if status != "SYNCING":  
        break  
  
except ClientError as e:  
    print("%s" % e)  
  
print("Program ends.")
```

Java

Para configurar una función Lambda para la manipulación avanzada de datos en los datos originales sin procesar

```
package com.amazonaws.kendra;  
  
import java.util.concurrent.TimeUnit;  
import software.amazon.awssdk.services.kendra.KendraClient;  
import software.amazon.awssdk.services.kendra.model.CreateDataSourceRequest;  
import software.amazon.awssdk.services.kendra.model.CreateDataSourceResponse;  
import software.amazon.awssdk.services.kendra.model.CreateIndexRequest;  
import software.amazon.awssdk.services.kendra.model.CreateIndexResponse;  
import software.amazon.awssdk.services.kendra.model.DataSourceConfiguration;  
import software.amazon.awssdk.services.kendra.model.DataSourceStatus;  
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJob;  
import software.amazon.awssdk.services.kendra.model.DataSourceSyncJobStatus;  
import software.amazon.awssdk.services.kendra.model.DataSourceType;  
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceRequest;  
import software.amazon.awssdk.services.kendra.model.DescribeDataSourceResponse;  
import software.amazon.awssdk.services.kendra.model.DescribeIndexRequest;  
import software.amazon.awssdk.services.kendra.model.DescribeIndexResponse;  
import software.amazon.awssdk.services.kendra.model.IndexStatus;  
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsRequest;  
import software.amazon.awssdk.services.kendra.model.ListDataSourceSyncJobsResponse;  
import software.amazon.awssdk.services.kendra.model.S3DataSourceConfiguration;  
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobRequest;  
import software.amazon.awssdk.services.kendra.model.StartDataSourceSyncJobResponse;  
  
public class CreateDataSourceWithCustomizationsExample {  
  
    public static void main(String[] args) throws InterruptedException {  
        System.out.println("Create a data source with customizations");  
  
        String dataSourceName = "data-source-name";  
        String indexId = "index-id";  
        String dataSourceRoleArn = "arn:aws:iam::account-id:role/role-name";  
        String s3BucketName = "S3-bucket-name"
```

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

CreateDataSourceRequest createDataSourceRequest = CreateDataSourceRequest
    .builder()
    .name(dataSourceName)
    .description(experienceDescription)
    .roleArn(experienceRoleArn)
    .type(DataSourceType.S3)
    .configuration(
        DataSourceConfiguration
            .builder()
            .s3Configuration(
                S3DataSourceConfiguration
                    .builder()
                    .bucketName(s3BucketName)
                    .build()
            ).build()
    )
    .customDocumentEnrichmentConfiguration(
        CustomDocumentEnrichmentConfiguration
            .builder()
            .preExtractionHookConfiguration(
                HookConfiguration
                    .builder()
                    .lambdaArn("arn:aws:iam::account-id:function/function-
name")
                    .s3Bucket("S3-bucket-name")
                    .build()
            ).roleArn("arn:aws:iam::account-id:role/cde-role-name")
            .build();
    );

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

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

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

    DataSourceStatus status = describeDataSourceResponse.status();
    System.out.println(String.format("Creating data source. Status: %s",
status));
    TimeUnit.SECONDS.sleep(60);
    if (status != DataSourceStatus.CREATING) {
        break;
    }
}

System.out.println(String.format("Synchronize the data source %s",
dataSourceId));
StartDataSourceSyncJobRequest startDataSourceSyncJobRequest =
StartDataSourceSyncJobRequest
    .builder()
    .indexId(indexId)
    .id(dataSourceId)
```

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

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

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

        TimeUnit.SECONDS.sleep(60);
        if (job.status() != DataSourceSyncJobStatus.SYNCING) {
            break;
        }
    }

    System.out.println("Data source creation with customizations is complete");
}
}
```

Contratos de datos para funciones Lambda

Sus funciones Lambda para la manipulación avanzada de datos interactúan con contratos de datos de Amazon Kendra. Los contratos son las estructuras de solicitud y respuesta obligatorias de sus funciones Lambda. Si sus funciones Lambda no siguen estas estructuras, Amazon Kendra arroja un error.

La función Lambda para PreExtractionHookConfiguration debería esperar la siguiente estructura de solicitud:

```
{
    "version": <str>,
    "dataBlobStringEncodedInBase64": <str>, //In the case of a data blob
    "s3Bucket": <str>, //In the case of an S3 bucket
    "s3ObjectKey": <str>, //In the case of an S3 bucket
    "metadata": <Metadata>
}
```

La estructura de metadata, que incluye la estructura de CustomDocumentAttribute, es la siguiente:

```
{
    "attributes": [<CustomDocumentAttribute>]
}

CustomDocumentAttribute
{
    "name": <str>,
    "value": <CustomDocumentAttributeValue>
```

```
}
```

```
CustomDocumentAttributeValue
{
    "stringValue": <str>,
    "integerValue": <int>,
    "longValue": <long>,
    "stringListValue": list<str>,
    "dateValue": <str>
}
```

La función Lambda para `PreExtractionHookConfiguration` debe cumplir la siguiente estructura de respuesta:

```
{
```

```
    "version": <str>,
    "dataBlobStringEncodedInBase64": <str>, //In the case of a data blob
    "s3ObjectKey": <str>, //In the case of an S3 bucket
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

La función Lambda para `PostExtractionHookConfiguration` debería esperar la siguiente estructura de solicitud:

```
{
```

```
    "version": <str>,
    "s3Bucket": <str>,
    "s3ObjectKey": <str>,
    "metadata": <Metadata>
}
```

La función Lambda para `PostExtractionHookConfiguration` debe cumplir la siguiente estructura de respuesta:

```
PostExtractionHookConfiguration Lambda Response
{
    "version": <str>,
    "s3ObjectKey": <str>,
    "metadataUpdates": [<CustomDocumentAttribute>]
}
```

El documento modificado se carga en su bucket de Amazon S3. El documento modificado debe seguir el formato que se muestra en [the section called “Formato del documento estructurado” \(p. 485\)](#).

Formato del documento estructurado

Amazon Kendra carga su documento estructurado en el bucket de Amazon S3 determinado. El documento estructurado sigue este formato:

```
Kendra document
{
    "textContent": <TextContent>
}

TextContent
{
    "documentBodyText": <str>
```

}

Ejemplo de una función Lambda que se adhiere a los contratos de datos

El siguiente código de Python es un ejemplo de una función Lambda que aplica manipulación avanzada de los campos de metadatos `_authors`, `_document_title` y el contenido del cuerpo de los documentos originales o sin procesar.

En el caso del contenido del cuerpo que reside en un bucket de Amazon S3

```
import json
import boto3

s3 = boto3.client("s3")

# Lambda function for advanced data manipulation
def lambda_handler(event, context):
    # Get the value of "S3Bucket" key name or item from the given event input
    s3_bucket = event.get("s3Bucket")
    # Get the value of "S3ObjectKey" key name or item from the given event input
    s3_object_key = event.get("s3ObjectKey")

    content_object_before_CDE = s3.get_object(Bucket = s3_bucket, Key = s3_object_key)
    content_before_CDE = content_object_before_CDE["Body"].read().decode("utf-8");
    content_after_CDE = "CDEInvolved " + content_before_CDE

    # Get the value of "metadata" key name or item from the given event input
    metadata = event.get("metadata")
    # Get the document "attributes" from the metadata
    document_attributes = metadata.get("attributes")

    s3.put_object(Bucket = s3_bucket, Key = "dummy_updated_kendra_document",
    Body=json.dumps(content_after_CDE))
    return {
        "version": "v0",
        "s3ObjectKey": "dummy_updated_kendra_document",
        "metadataUpdates": [
            {"name": "_document_title", "value": {
                "stringValue": "title_from_pre_extraction_lambda"}},
            {"name": "_authors", "value": {"stringListValue": ["author1", "author2"]}}
        ]
    }
```

En el caso del contenido del cuerpo que reside en un blob de datos

```
import json
import boto3
import base64

# Lambda function for advanced data manipulation
def lambda_handler(event, context):

    # Get the value of "dataBlobStringEncodedInBase64" key name or item from the given
    event input
    data_blob_string_encoded_in_base64 = event.get("dataBlobStringEncodedInBase64")
    # Decode the data blob string in UTF-8
    data_blob_string =
    base64.b64decode(data_blob_string_encoded_in_base64).decode("utf-8")
    # Get the value of "metadata" key name or item from the given event input
    metadata = event.get("metadata")
```

```
# Get the document "attributes" from the metadata
document_attributes = metadata.get("attributes")

new_data_blob = "This should be the modified data in the document by pre processing
lambda ".encode("utf-8")
return {
    "version": "v0",
    "dataBlobStringEncodedInBase64": base64.b64encode(new_data_blob).decode("utf-8"),
    "metadataUpdates": [
        {"name": "_document_title", "value": {
            "stringValue": "title_from_pre_extraction_lambda"}},
        {"name": "_authors", "value": {"stringListValue": ["author1", "author2"]}}
    ]
}
```

El siguiente código de Python es un ejemplo de una función Lambda que aplica manipulación avanzada de los campos de metadatos `_authors`, `_document_title` y el contenido del cuerpo de los documentos estructurados o analizados.

```
import json
import boto3
import time

s3 = boto3.client("s3")

# Lambda function for advanced data manipulation
def lambda_handler(event, context):

    # Get the value of "S3Bucket" key name or item from the given event input
    s3_bucket = event.get("s3Bucket")
    # Get the value of "S3ObjectKey" key name or item from the given event input
    s3_key = event.get("s3ObjectKey")
    # Get the value of "metadata" key name or item from the given event input
    metadata = event.get("metadata")
    # Get the document "attributes" from the metadata
    document_attributes = metadata.get("attributes")

    kendra_document_object = s3.get_object(Bucket = s3_bucket, Key = s3_key)
    kendra_document_string = kendra_document_object['Body'].read().decode('utf-8')
    kendra_document = json.loads(kendra_document_string)
    kendra_document["textContent"]["documentBodyText"] = "Changing document body to a short
sentence."

    s3.put_object(Bucket = s3_bucket, Key = "dummy_updated_kendra_document",
Body=json.dumps(kendra_document))

    return {
        "version" : "v0",
        "s3ObjectKey": "dummy_updated_kendra_document",
        "metadataUpdates": [
            {"name": "_document_title", "value": {"stringValue": "title_from_post_extraction_lambda"}},
            {"name": "_authors", "value": {"stringListValue": ["author1", "author2"]}}
        ]
    }
```

Búsqueda en un índice

Para buscar un Amazon Kendraíndice, se utiliza el [Consulta API](#). El [Query API](#) devuelve información sobre los documentos indexados que utilizas en tu aplicación. En esta sección se muestra cómo realizar una consulta, realizar filtros e interpretar la respuesta que se obtiene del [Query API](#).

Para buscar documentos que haya indexado con Amazon Kendrapor Amazon Lex, utilizar [AMAZONA.KendraSearchIntent](#). Para ver un ejemplo de configuración Amazon Kendra con Amazon Lex, consulte [Creación de un bot de preguntas frecuentes para un Amazon Kendraíndice](#).

Temas

- [Consultar un índice \(p. 488\)](#)
- [Navegar por un índice \(p. 500\)](#)
- [Incluye resultados de búsqueda \(p. 502\)](#)
- [Búsqueda tabular de HTML \(p. 504\)](#)
- [Sugerencias de consulta \(p. 507\)](#)
- [Corrector ortográfico de consultas \(p. 519\)](#)
- [Filtrado y búsqueda de facetas \(p. 520\)](#)
- [Filtrar según el contexto del usuario \(p. 525\)](#)
- [Respuestas a consultas y tipos de respuestas \(p. 538\)](#)
- [Ajustar y ordenar las respuestas \(p. 544\)](#)

Consultar un índice

Cuando buscas en tu índice, Amazon Kendra utiliza toda la información que ha proporcionado sobre sus documentos para determinar los documentos más relevantes para los términos de búsqueda introducidos. Algunos de los artículos que Amazon Kendra considera que son:

- El texto o el cuerpo del documento.
- Título del documento.
- Campos de texto personalizados que ha marcado como en los que se pueden realizar búsquedas.
- El campo de fecha que ha indicado debe usarse para determinar la «frescura» de un documento.
- Cualquier otro campo que pueda proporcionar información relevante.

Amazon Kendratambién puede filtrar la respuesta en función de cualquier filtro de campo o atributo que haya configurado para la búsqueda. Por ejemplo, si tienes un campo personalizado llamado «departamento», puedes filtrar la respuesta para que solo devuelva documentos de un departamento denominado «legal». Para obtener más información, consulte [Campos o atributos personalizados](#).

Los resultados de búsqueda devueltos se ordenan según la relevancia que Amazon Kendradetermina para cada documento. Los resultados se paginan para que puedas mostrar una página a la vez a tu usuario.

Para buscar documentos que haya indexado con Amazon Kendrapor Amazon Lex, utilizar [AMAZONA.KendraSearchIntent](#). Para ver un ejemplo de configuración Amazon Kendra con Amazon Lex, consulte [Creación de un bot de preguntas frecuentes para un Amazon Kendraíndice](#).

El siguiente ejemplo muestra cómo buscar en un índice. Amazon Kendradetermina el tipo de resultado de la búsqueda (respuesta, documento, pregunta-respuesta) que mejor se adapta a la consulta.

Para obtener información sobre las respuestas a la consulta, consulte [Respuestas a consultas y tipos de respuestas \(p. 538\)](#).

Requisitos previos

Antes de usar el[ConsultaAPI](#) para consultar un índice:

- Configure los permisos necesarios para un índice y conéctese a su fuente de datos o cargue sus documentos por lotes. Para obtener más información, consulte[IAMpapeles](#). Utiliza el nombre de recurso de Amazon del rol cuando llama a la API para crear un índice y un conector de fuentes de datos o para cargar documentos por lotes.
- Configure una de lasAWS Command Line Interface, un SDK o vaya aAmazon Kendraconsola. Para obtener más información, consulte[Configuración de Amazon Kendra](#).
- Cree un índice y conéctese a una fuente de datos de documentos o cargue documentos por lotes. Para obtener más información, consulte[Crear un índice](#)y[Creación de un conector de fuente de datos](#).

Búsqueda en un índice (consola)

Puede utilizar elAmazon Kendraconsola para buscar y probar el índice. Puede realizar consultas y ver los resultados.

Para buscar en un índice con la consola

- Inicia sesión en elAWS Management Consoley abre elAmazon Kendraconsola en<http://console.aws.amazon.com/kendra/>.
- En el panel de navegación, elijaÍndices.
- Elige tu índice.
- En el menú de navegación, selecciona la opción para buscar en tu índice.
- Introduzca una consulta en el cuadro de texto y, a continuación, pulse Entrar.
- Amazon Kendradevuelve los resultados de la búsqueda.

También puede obtener el identificador de la consulta para la búsqueda seleccionando el icono de la bombilla en el panel lateral.

Búsqueda en un índice (SDK)

Para buscar en un índice con Python o Java

- En el ejemplo siguiente se busca en un índice. Cambiar el valor dequerya tu consulta de búsqueda yindex_idoindexIdal identificador de índice del índice en el que desea buscar.

También puede obtener el identificador de consulta de la búsqueda como parte de los elementos de respuesta cuando llama al[ConsultaAPI](#).

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.documentElement().text();
                    System.out.println(String.format("Title: %s", documentTitle));
                    String documentExcerpt = item.documentElement().text();
                    System.out.println(documentExcerpt);
            }
        }
    }
}
```

```
        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");
}
```

Búsqueda en un índice (Postman)

Puedes usar [Cartero](#) para consultar y probar su Amazon Kendra índice.

Para buscar en un índice mediante Postman

1. Crea una nueva colección en Postman y establece el tipo de solicitud en CORREO.
2. Introduzca la URL del punto final. Por ejemplo, `https://kendra.<region>.amazonaws.com`.
3. Seleccione el Autorización tabla e introduce la siguiente información.
 - Teclea—Seleccione AWS firma.
 - AccessKey—Introduzca la clave de acceso generada al crear un IAM usuario.
 - SecretKey—Introduzca la clave secreta generada al crear un IAM usuario.
 - AWSRegión—Introduzca la región de su índice. Por ejemplo, US-west-2.
 - Nombre del servicio—Entrar kendra. Esto distingue entre mayúsculas y minúsculas, por lo que debe ser en minúsculas.

Warning

Si introduce un nombre de servicio incorrecto o no utiliza minúsculas, se generará un error una vez que seleccione Enviar para enviar la solicitud: «La credencial debe estar asignada al servicio 'kendra' correcto».

También debe comprobar que ha introducido la clave de acceso y la clave secreta correctas.

4. Seleccione el Encabezado tabla e introduce la siguiente información de clave y valor.
 - Clave:X-Amz-Target
Valor:com.amazonaws.kendra.AWSKendraFrontendService.Consulta
 - Clave:Codificación de contenido
Valor:amz-1.0
5. Seleccione el Cuerpo tabla y haz lo siguiente.
 - Elige el crudo Tipo JSON para el cuerpo de la solicitud.
 - Introduzca un JSON que incluya su ID de índice y el texto de la consulta.

```
{
    "IndexId": "index-id",
    "QueryText": "enter a query here"
}
```

Warning

Si tu JSON no usa la indentación correcta, se produce un error:»SerializationException». Comprueba la indentación en tu JSON.

6. SeleccioneEnviar(cerca de la parte superior derecha).

Búsqueda con sintaxis de consulta avanzada

Puede crear consultas que sean más específicas que las consultas de palabras clave simples o de lenguaje natural mediante el uso de operadores o sintaxis de consulta avanzados. Esto incluye rangos, booleanos, caracteres comodín y mucho más. Mediante el uso de operadores, puede dar más contexto a la consulta y refinar aún más los resultados de la búsqueda.

Amazon Kendra admite los siguientes operadores.

- Booleano: lógica para limitar o ampliar la búsqueda. Por ejemplo,amazon AND sports limita la búsqueda para buscar únicamente documentos que contengan ambos términos.
- Paréntesis: lee los términos de consulta anidados en orden de prioridad. Por ejemplo,(amazon AND sports) NOT rainforest lee(amazon AND sports) delante de NOT rainforest.
- Intervalos: valores de fecha o rango numérico. Los rangos pueden ser inclusivos, exclusivos o ilimitados. Por ejemplo, puede buscar documentos que se actualizaron por última vez entre el 1 de enero de 2020 y el 31 de diciembre de 2020, incluidas estas fechas.
- Campos: utiliza un campo específico para limitar la búsqueda. Por ejemplo, puede buscar documentos que tengan la palabra «Estados Unidos» en el campo «ubicación».
- Comodín: coinciden parcialmente con una cadena de texto. Por ejemplo, Cloud* podría coincidir CloudFormation. Amazon Kendra actualmente solo admite caracteres comodín al final.
- Comillas exactas: coinciden exactamente con una cadena de texto. Por ejemplo, los documentos que contienen "Amazon Kendra" "pricing".

Puede utilizar una combinación de cualquiera de los operadores anteriores.

Tenga en cuenta que el uso excesivo de operadores o consultas muy complejas podría afectar a la latencia de las consultas. Los comodines son algunos de los operadores más caros en términos de latencia. Una regla general es que cuantos más términos y operadores utilice, mayor será el impacto potencial en la latencia. Otros factores que afectan a la latencia son el tamaño promedio de los documentos indexados, el tamaño del índice, cualquier filtrado de los resultados de búsqueda y la carga total de Amazon Kendra índice.

Booleano

Puede combinar o excluir palabras mediante los operadores booleanos AND, OR, NOT.

Los siguientes son ejemplos del uso de operadores booleanos.

amazon AND sports

Devuelve los resultados de búsqueda que contienen los términos «amazon» y «sports» en el texto, como Amazon Prime Video sports u otro contenido similar.

sports OR recreation

Devuelve los resultados de búsqueda que contienen los términos «deportes» o «recreación», o ambos, en el texto.

amazon NOT rainforest

Devuelve los resultados de búsqueda que contienen el término «amazonia», pero no el término «selva tropical» en el texto. Esto es para buscar documentos sobre la empresa Amazon, no sobre la selva amazónica.

Paréntesis

Puede consultar las palabras anidadas en orden de prioridad mediante paréntesis. Los paréntesis indican que Amazon Kendra cómo debe leerse una consulta.

Los siguientes son ejemplos del uso de operadores de paréntesis.

(amazon AND sports) NOT rainforest

Devuelve documentos que contienen los términos «amazonía» y «deportes» en el texto, pero no el término «selva tropical». Esto es para buscar videos deportivos u otro contenido similar en Amazon Prime, no sobre deportes de aventura en la selva amazónica. Los paréntesis ayudan a indicar que `amazon AND sports` debe leerse antes de `NOT rainforest`. La consulta no debe leerse como `amazon AND (sports NOT rainforest)`.

(amazon AND (sports OR recreation)) NOT rainforest

Devuelve documentos que contienen los términos «deportes» o «recreación», o ambos, y el término «Amazon». Pero no incluye el término «selva tropical». Esto es para buscar videos deportivos o recreativos en Amazon Prime, no sobre deportes de aventura en la selva amazónica. Los paréntesis ayudan a indicar que `sports OR recreation` debe leerse antes de combinarse con 'amazon', que se lee antes de `NOT rainforest`. La consulta no debe leerse como `amazon AND (sports OR (recreation NOT rainforest))`.

Gamas

Puede utilizar un rango de valores para filtrar los resultados de la búsqueda. Especifica un atributo y los valores del rango. Puede ser de fecha o de tipo numérico.

Los intervalos de fechas deben tener los siguientes formatos:

- Epoch
- YYYY
- aaaa-mm
- aaaa-mm-dd
- aaaa-mm-dd't'hh

También puede especificar si desea incluir o excluir los valores inferior y superior del rango.

Los siguientes son ejemplos del uso de operadores de rango.

_processed_date:>2019-12-31 AND _processed_date:<2021-01-01

Devuelve los documentos que se procesaron en 2020 (más del 31 de diciembre de 2019 y menos del 1 de enero de 2021).

_processed_date:>=2020-01-01 AND _processed_date:<=2020-12-31

Devuelve los documentos que se procesaron en 2020 (igual o superior al 1 de enero de 2020 e inferior o igual al 31 de diciembre de 2020).

_document_likes:<1

Devuelve documentos con cero «me gusta» o sin comentarios de los usuarios (menos de 1 «me gusta»).

Puede especificar si un rango debe tratarse como inclusivo o exclusivo de los valores de rango dados.

Inclusivo

_last_updated_at:[2020-01-01 TO 2020-12-31]

Devuelve los documentos actualizados por última vez en 2020, incluidos los días 1 de diciembre de 2020 y 31 de diciembre de 2020.

Exclusivo

_last_updated_at:{2019-12-31 TO 2021-01-01}

Devuelve los documentos que se actualizaron por última vez en 2020; no incluye los días 31 de diciembre de 2019 y 1 de enero de 2021.

< and >Para rangos ilimitados que no son ni inclusivos ni exclusivos, simplemente use los operadores. Por ejemplo, _last_updated_at:>2019-12-31 AND _last_updated_at:<2021-01-01 .

Campos

Puede limitar la búsqueda para que solo devuelvan los documentos que cumplan con un valor en un campo específico. El campo puede ser de cualquier tipo.

Los siguientes son ejemplos del uso de operadores de contexto a nivel de campo.

status:"Incomplete" AND financial_year:2021

Devuelve los documentos del ejercicio financiero de 2021 con su estado como incompletos.

(sports OR recreation) AND country:"United States" AND level:"professional"

Devuelve documentos que hablan sobre deportes o actividades recreativas profesionales en los Estados Unidos.

Caracteres comodín

Puede ampliar la búsqueda para incluir variantes de palabras y frases mediante el operador comodín. Esto resulta útil cuando se buscan variantes de nombres. Amazon Kendra actualmente solo admite caracteres comodín al final. El número de caracteres de prefijo de un comodín al final debe ser superior a dos.

Los siguientes son ejemplos del uso de operadores comodín.

Cloud*

Devuelve documentos que contienen variantes como CloudFormation y CloudWatch.

kendra*aws

Devuelve documentos que contienen variantes, como kendra.amazonaws.

kendra*aws*

Devuelve documentos que contienen variantes, como kendra.amazonaws.com

Cotizaciones exactas

Puede utilizar comillas para buscar una coincidencia exacta de un fragmento de texto.

Los siguientes son ejemplos del uso de comillas.

"Amazon Kendra" "pricing"

Devuelve documentos que contienen la frase 'Amazon Kendra' y el término «precios». Los documentos deben contener ambos 'Amazon Kendra'y precios 'para obtener los resultados.

"Amazon Kendra" "pricing" cost

Devuelve documentos que contienen la frase 'Amazon Kendra'y el término «fijación de precios» y, opcionalmente, el término «coste». Los documentos deben contener ambos 'Amazon Kendra'y precios 'para mostrar los resultados, pero es posible que no incluyan necesariamente el «coste».

Sintaxis de consulta no válida

Amazon Kendraemite una advertencia si hay problemas con la sintaxis de la consulta o si su consulta no es compatible actualmente conAmazon Kendra. Para obtener más información, consulte[Documentación de la API para advertencias de consultas](#).

Las siguientes consultas son ejemplos de sintaxis de consulta no válida.

_last_updated_at:<2021-12-32

Fecha no válida. El día 32 no existe en el calendario gregoriano, que es utilizado porAmazon Kendra.

_view_count:ten

Valor numérico no válido. Se deben utilizar dígitos para representar valores numéricos.

nonExistentField:123

Búsqueda de campo no válida. El campo debe existir para poder utilizar la búsqueda de campos.

Product:[A TO D]

Intervalo no válido. Se deben utilizar valores numéricos o fechas para los rangos.

OR Hello

Booleano no válido. Los operadores deben usarse con términos y colocarse entre términos.

Búsqueda en idiomas

Puede buscar documentos en un idioma compatible. Introduce el código de idioma en el[AttributeFilter](#)para devolver los documentos filtrados en el idioma elegido. Puede escribir la consulta en un idioma compatible.

Si no especificas un idioma,Amazon Kendraconsulta documentos en inglés de forma predeterminada. Para obtener más información sobre los idiomas admitidos, incluidos sus códigos, consulte[Añadir documentos en idiomas distintos del inglés](#).

Para buscar documentos en un idioma compatible en la consola, seleccione su índice y, a continuación, seleccione la opción de buscar su índice en el menú de navegación. Elija el idioma en el que desea devolver los documentos seleccionando la configuración de búsqueda y, a continuación, seleccionando un idioma del menú desplegable.Idioma.

Los siguientes ejemplos muestran cómo buscar documentos en español.

Para buscar un índice en español en la consola

1. Inicia sesión en elAWS Management Consoley abre elAmazon Kendraconsola en<http://console.aws.amazon.com/kendra/>.
2. En el menú de navegación, seleccionaÍndicesy elige tu índice.
3. En el menú de navegación, selecciona la opción para buscar en tu índice.
4. En la configuración de búsqueda, seleccionaldiomasdesplegable y elige español.

5. Introduzca una consulta en el cuadro de texto y, a continuación, pulse Entrar.
6. Amazon Kendra devuelve los resultados de la búsqueda en español.

Para buscar un índice en español mediante la CLI, Python o Java

- En el siguiente ejemplo, se busca un índice en español. Cambiar el valor searchString a tu consulta de búsqueda y al valor indexID al identificador del índice en el que deseas buscar. El código de idioma del español es es. Puedes sustituirlo por el código de su propio idioma.

CLI

```
{  
    "EqualsTo": {  
        "Key": "_language_code",  
        "Value": {  
            "StringValue": "es"  
        }  
    }  
}
```

Python

```
import boto3  
import pprint  
  
kendra = boto3.client("kendra")  
  
# Provide the index ID  
index_id = "index-id"  
# Provide the query text  
query = "search-string"  
  
# Includes the index ID, query text, and language attribute filter  
response = kendra.query(  
    QueryText = query,  
    IndexId = index_id,  
    AttributeFilter = {  
        "EqualsTo": {  
            "Key": "_language_code",  
            "Value": {  
                "StringValue": "es"  
            }  
        }  
    }  
)  
  
print ("\nSearch results|Resultados de la búsqueda: " + query + "\n")  
  
for query_result in response["ResultItems"]:  
  
    print("-----")  
    print("Type: " + str(query_result["Type"]))  
  
    if query_result["Type"]=="ANSWER" or query_result["Type"]=="QUESTION_ANSWER":  
        answer_text = query_result["DocumentExcerpt"]["Text"]  
        print(answer_text)  
  
    if query_result["Type"]=="DOCUMENT":  
        if "DocumentTitle" in query_result:  
            document_title = query_result["DocumentTitle"]["Text"]  
            print("Title: " + document_title)  
        document_text = query_result["DocumentExcerpt"]["Text"]  
        print(document_text)
```

```
print("-----\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResponse;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String query = "searchString";
        String indexId = "indexID";

        QueryRequest queryRequest = QueryRequest.builder()
            .queryText(query)
            .indexId(indexId)
            .attributeFilter(
                AttributeFilter.builder()
                    .withEqualsTo(
                        DocumentAttribute.builder()
                            .withKey("_language_code")
                            .withValue("es")
                            .build())
                    .build())
            .build();
        .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.documentElement().text();
                    System.out.println(String.format("Title: %s", documentTitle));
                    String documentExcerpt = item.documentExcerpt().text();
                    System.out.println(String.format("Excerpt: %s",
                        documentExcerpt));
                    break;
                default:
                    System.out.println(String.format("Unknown query result type:
%s", item.type()));
            }
            System.out.println("-----\n");
        }
    }
}
```

}

Recuperación de pasajes

Puede utilizar el [RetrieveAPI](#) como recuperador para sistemas de recuperación y generación aumentada (RAG).

Los sistemas RAG utilizan inteligencia artificial generativa para crear aplicaciones de respuesta a preguntas. Los sistemas RAG constan de un recuperador y modelos de lenguaje grande (LLM). Ante una consulta, el recuperador identifica los fragmentos de texto más relevantes de un corpus de documentos y los envía al LLM para proporcionar la respuesta más útil. Luego, el LLM analiza los fragmentos o pasajes de texto relevantes y genera una respuesta completa para la consulta.

El [RetrieveLa API](#) analiza fragmentos de texto o extractos que se denominan [pasajes](#) y devuelve los pasajes principales que son más relevantes para la consulta.

Al igual que el [QueryAPI](#), la [RetrieveLa API](#) también busca información relevante mediante la búsqueda semántica. La búsqueda semántica tiene en cuenta el contexto de la consulta de búsqueda, además de toda la información disponible de los documentos indexados. Sin embargo, de forma predeterminada, el [QueryLa API](#) solo devuelve fragmentos de hasta 100 palabras simbólicas. Con el [RetrieveAPI](#), puede recuperar pasajes más largos de hasta 200 palabras simbólicas y hasta 100 pasajes semánticamente relevantes. Esto no incluye las respuestas de tipo pregunta-respuesta o preguntas frecuentes de tu índice. Los pasajes son extractos de texto que se pueden extraer semánticamente de varios documentos y de varias partes del mismo documento. Si, en casos extremos, sus documentos no producen ningún pasaje, utilice el [RetrieveAPI](#), también puede utilizar la [QueryLa API](#) y sus tipos de respuestas.

También puede hacer lo siguiente con [RetrieveAPI](#):

- Anule el aumento a nivel del índice
- Filtrar según los campos o atributos del documento
- Filtra según el acceso del usuario o su grupo a los documentos

También puedes incluir ciertos campos en la respuesta que podrían proporcionar información adicional útil.

El [RetrieveActualmente](#), la API no admite todas las funciones compatibles con [QueryAPI](#). No se admiten las siguientes funciones: realizar consultas mediante [sintaxis de consulta avanzada](#), [correcciones ortográficas sugeridas](#) para consultas, [facetado](#), [sugerencias de consulta](#) para completar automáticamente las consultas de búsqueda, [aprendizaje incremental](#), y [cubos de confianza](#). Tenga en cuenta que no todas las funciones se aplican a [RetrieveAPI](#). Cualquier versión futura de [RetrieveLa API](#) se documentará en esta guía.

El [RetrieveLa API](#) comparte el número de [unidades de capacidad de consulta](#) que configuraste para tu índice. Para obtener más información sobre lo que se incluye en una sola unidad de capacidad y la capacidad base predeterminada de un índice, consulte [Capacidad de ajuste](#).

Note

No puede añadir capacidad si está utilizando el [Amazon Kendra Edición para desarrolladores](#); solo se puede añadir capacidad cuando se usa [Amazon Kendra Edición empresarial](#). Para obtener más información sobre lo que se incluye en las ediciones para desarrolladores y empresas, consulte [Amazon Kendra Ediciones](#).

El siguiente es un ejemplo del uso de [RetrieveAPI](#) para recuperar los 100 pasajes más relevantes de los documentos de un índice para la consulta "how does amazon kendra work?"

Python

```
import boto3
import pprint

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the query text
query = "how does amazon kendra work?"
# You can retrieve up to 100 relevant passages
# You can paginate 100 passages across 10 pages, for example
page_size = 10
page_number = 10

result = kendra.retrieve(
    IndexId = index_id,
    QueryText = query,
    PageSize = page_size,
    PageNumber = page_number)

print("\nRetrieved passage results for query: " + query + "\n")

for retrieve_result in result["ResultItems"]:

    print("-----")
    print("Title: " + str(retrieve_result["DocumentTitle"]))
    print("URI: " + str(retrieve_result["DocumentURI"]))
    print("Passage content: " + str(retrieve_result["Content"]))
    print("-----\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.RetrieveRequest;
import software.amazon.awssdk.services.kendra.model.RetrieveResult;
import software.amazon.awssdk.services.kendra.model.RetrieveResultItem;

public class RetrievePassageExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indxId = "index-id";
        String query = "how does amazon kendra work?";
        Integer pgSize = 10;
        Integer pgNumber = 10;

        RetrieveRequest retrieveRequest = retrieveRequest
            .builder()
            .indexId(indxId)
            .queryText(query)
            .pageSize(pgSize)
            .pageNumber(pgNumber)
            .build();

        RetrieveResult retrieveResult = kendra.retrieve(retrieveRequest);

        System.out.println(String.format("\nRetrieved passage results for query: %s",
            query));
        for(RetrieveResultItem item: retrieveResult.resultItems()) {
            System.out.println("-----");
```

```
        System.out.println(String.format("Title: %s", documentTitle));
        System.out.println(String.format("URI: %s", documentURI));
        System.out.println(String.format("Passage content: %s", content));
        System.out.println("-----\n");
    }
}
```

Navegar por un índice

Puede buscar documentos por sus atributos o facetas sin tener que escribir una consulta de búsqueda. Amazon Kendra Navegación por índicespuede ayudar a los usuarios a descubrir documentos al navegar libremente por un índice sin tener en cuenta una consulta específica. Esto también ayuda a los usuarios a explorar ampliamente un índice como punto de partida en su búsqueda.

La búsqueda de índices solo se puede utilizar para buscar por atributo o faceta del documento con un tipo de clasificación. No puede buscar en un índice completo mediante el Explorador de índices. Si falta el texto de la consulta, entoncesAmazon Kendrasolicita un filtro de atributos del documento o una faceta y un tipo de clasificación.

Para permitir la navegación por índices mediante el[ConsultaAPI](#), debes incluir[AttributeFilteroFaceta](#), [ySortingConfiguration](#). Para permitir la navegación por índices en la consola, selecciona tu índice enIndicesen el menú de navegación y, a continuación, selecciona la opción para buscar en tu índice. En el cuadro de búsqueda, pulsa elEntrarclave dos veces. Selecciona el menú desplegableFiltrar resultados de búsquedapara elegir un filtro y seleccionar el menú desplegableClasificarpara elegir un tipo de clasificación.

A continuación se muestra un ejemplo de cómo explorar un índice de documentos en español en orden descendente según la fecha de creación del documento.

CLI

```
aws kendra query \
--index-id "index-id" \
--attribute-filter '{
    "EqualsTo": {
        "Key": "_language_code",
        "Value": {
            "StringValue": "es"
        }
    }
}' \
--sorting-configuration '{
    "DocumentAttributeKey": "_created_at",
    "SortOrder": "DESC"
}'
```

Python

```
import boto3

kendra = boto3.client("kendra")

# Must include the index ID, the attribute filter, and sorting configuration
response = kendra.query(
    IndexId = "index-id",
    AttributeFilter = {
        "EqualsTo": {
```

```
        "Key": "_language_code",
        "Value": {
            "StringValue": "es"
        }
    },
    SortingConfiguration = {
        "DocumentAttributeKey": "_created_at",
        "SortOrder": "DESC"})
}

print("\nSearch results|Resultados de la búsqueda: \n")

for query_result in response["ResultItems"]:

    print("-----")
    print("Type: " + str(query_result["Type"]))

    if query_result["Type"]=="ANSWER" or query_result["Type"]=="QUESTION_ANSWER":
        answer_text = query_result["DocumentExcerpt"]["Text"]
        print(answer_text)

    if query_result["Type"]=="DOCUMENT":
        if "DocumentTitle" in query_result:
            document_title = query_result["DocumentTitle"]["Text"]
            print("Title: " + document_title)
        document_text = query_result["DocumentExcerpt"]["Text"]
        print(document_text)

    print("-----\n\n")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.QueryRequest;
import software.amazon.awssdk.services.kendra.model.QueryResult;
import software.amazon.awssdk.services.kendra.model.QueryResultItem;

public class SearchIndexExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();
        QueryRequest queryRequest = QueryRequest.builder()
            .withIndexId("index-id")
            .withAttributeFilter(AttributeFilter.builder()
                .withEqualsTo(DocumentAttribute.builder()
                    .withKey("_language_code")
                    .WithValue(DocumentAttributeValue.builder()
                        .withStringValue("es")
                        .build())
                    .build())
                .build())
            .withSortingConfiguration(SortingConfiguration.builder()
                .withDocumentAttributeKey("_created_at")
                .withSortOrder("DESC")
                .build())
            .build());
    }

    QueryResult queryResult = kendra.query(queryRequest);
    for (QueryResultItem item : queryResult.getResultItems()) {
        System.out.println("-----");
        System.out.println(String.format("Type: %s", item.getType()));

        switch (item.getType()) {
            case QueryResultType.QUESTION_ANSWER:
```

```
case QueryResultType.ANSWER:  
    String answerText = item.getDocumentExcerpt().getText();  
    System.out.println(answerText);  
    break;  
case QueryResultType.DOCUMENT:  
    String documentTitle = item.getDocumentTitle().getText();  
    System.out.println(String.format("Title: %s", documentTitle));  
    String documentExcerpt = item.getDocumentExcerpt().getText();  
    System.out.println(String.format("Excerpt: %s", documentExcerpt));  
    break;  
default:  
    System.out.println(String.format("Unknown query result type: %s",  
item.getType()));  
}  
}  
System.out.println("-----\n");  
}  
}
```

Incluye resultados de búsqueda

Puede incluir ciertos documentos en los resultados de la búsqueda cuando los usuarios realicen determinadas consultas. Esto ayuda a que los resultados sean más visibles y prominentes para los usuarios. Los resultados destacados se separan de la lista habitual de resultados y se muestran en la parte superior de la página de búsqueda. Puedes experimentar con diferentes documentos para diferentes consultas o asegurarte de que ciertos documentos tengan la visibilidad que se merecen.

Asigna consultas específicas a documentos específicos para incluirlos en los resultados. Si una consulta contiene una coincidencia exacta, se mostrarán uno o más documentos específicos en los resultados de la búsqueda.

Por ejemplo, puedes especificar que, si tus usuarios publican la consulta «nuevos productos 2023», seleccionen los documentos titulados «Novedades» y «Próximamente» para que aparezcan en la parte superior de la página de resultados de búsqueda. Esto ayuda a garantizar que estos documentos sobre los nuevos productos reciban la visibilidad que se merecen.

Amazon Kendra duplica los resultados de búsqueda si ya se ha seleccionado un resultado para que aparezca en la parte superior de la página de resultados de búsqueda. Un resultado destacado no vuelve a clasificarse como el primer resultado si ya aparece por encima de todos los demás resultados.

Para mostrar ciertos resultados, debe especificar una coincidencia exacta de una consulta de texto completo, no una coincidencia parcial de una consulta mediante una palabra clave o frase incluida en una consulta. Por ejemplo, si solo especificas la consulta «Kendra» en un conjunto de resultados destacados, consultas como «¿Cómo clasifica Kendra semánticamente los resultados?» no mostrará los resultados destacados. Los resultados destacados están diseñados para consultas específicas, en lugar de consultas con un alcance demasiado amplio. Amazon Kendra gestiona de forma natural las consultas de tipos de palabras clave para clasificar los documentos más útiles en los resultados de búsqueda, evitando que los resultados se muestren excesivamente en función de palabras clave simples.

Si hay determinadas consultas que los usuarios utilizan con frecuencia, puede especificarlas para los resultados destacados. Por ejemplo, si observas tus consultas principales utilizando [Amazon Kendra Analítica](#) y encuentras esas consultas específicas, como «¿Cómo clasifica Kendra semánticamente los resultados?» y la «búsqueda semántica de kendra», se utilizan con frecuencia, por lo que podría ser útil especificar estas consultas para incluir el documento titulado 'Amazon Kendra buscar 101'.

Amazon Kendra trata las consultas de resultados destacados como insensibles a mayúsculas y minúsculas. Amazon Kendra convierte una consulta a minúsculas y reemplaza los espacios en blanco

finales por un espacio único. Amazon Kendra coincide con todos los demás caracteres tal como están cuando especificas tus consultas para los resultados destacados.

Se crea un conjunto de resultados destacados que se asignan a determinadas consultas mediante la [CreateFeaturedResultsSetAPI](#). Si utilizas la consola, seleccionas tu índice y, a continuación, seleccionas Resultados destacados en el menú de navegación para crear un conjunto de resultados destacados. Puede crear hasta 50 conjuntos de resultados destacados por índice, hasta cuatro documentos para destacar por conjunto y hasta 49 textos de consulta por conjunto de resultados destacados. Puede solicitar el aumento de estos límites comunicándose con [Soporte](#).

Puede seleccionar el mismo documento en varios conjuntos de resultados destacados. Sin embargo, no debe utilizar el mismo texto de consulta de coincidencia exacta en varios conjuntos. Las consultas que especifique para los resultados destacados deben ser únicas para cada conjunto de resultados destacados para cada índice.

Puede organizar el orden de los documentos al seleccionar hasta cuatro documentos destacados. Si utilizas la API, el orden en que aparecen los documentos destacados es el mismo que se muestra en los resultados destacados. Si utilizas la consola, solo tienes que arrastrar y soltar el orden de los documentos al seleccionar los documentos para que aparezcan en los resultados.

El control de acceso, en el que ciertos usuarios y grupos tienen acceso a ciertos documentos y otros no, se sigue respetando al configurar los resultados destacados. Esto también es válido para el filtrado de contextos de usuario. Por ejemplo, el usuario A pertenece al grupo empresarial «Becarios», que no debería acceder a los documentos sobre los secretos de la empresa. Si el usuario A introduce una consulta que incluye un documento secreto de la empresa, el usuario A no verá este documento en sus resultados. Esto también se aplica a cualquier otro resultado de la página de resultados de búsqueda. También puede utilizar etiquetas para controlar el acceso a un conjunto de resultados destacados, que es un Amazon Kendra recurso al que se controla el acceso.

A continuación se muestra un ejemplo de cómo crear un conjunto de resultados destacados con las consultas «nuevos productos 2023» y «nuevos productos disponibles» asignadas a los documentos titulados «Novedades» (doc-id-1) y «Próximamente» (doc-id-2).

CLI

```
aws kendra create-featured-results-set \
--featured-results-set-name 'New product docs to feature' \
--description "Featuring What's new and Coming soon docs" \
--index-id index-id \
--query-texts 'new products 2023' 'new products available' \
--featured-documents '[{"Id": "doc-id-1", "Id": "doc-id-2"}]'
```

Python

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

kendra = boto3.client("kendra")

print("Create a featured results set.")

# Provide a name for the featured results set
featured_results_name = "New product docs to feature"
# Provide an optional description 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 = ['doc-id-1', 'doc-id-2']

try:
    featured_results_set_response = kendra.createFeaturedResultsSet(
        FeaturedResultsSetName = featured_results_name,
        Description = 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.describeFeaturedResultsSet(
            Id = featured_results_set_id
        )
        status = featured_results_set_description["Status"]
        print(" Featured results set status: "+status)

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

print("Program ends.")
```

Búsqueda tabular de HTML

Amazon KendraLa función de búsqueda tabular puede buscar y extraer respuestas de tablas incrustadas en documentos HTML. Cuando buscas en tu índice,Amazon Kendraincluye un extracto de una tabla si es relevante para la consulta y proporciona información útil.

Amazon Kendraexamina toda la información del cuerpo del documento, incluida la información útil de las tablas. Por ejemplo, un índice contiene informes comerciales con tablas sobre los costos de operación, los ingresos y otra información financiera. Para la consulta, «¿cuál es el costo operativo anual de 2020 a 2022?» ,Amazon Kendrapuede mostrar un extracto de una tabla que contiene las columnas pertinentes de la tabla «Operaciones (millones de USD)» y «Ejercicio financiero», y filas de la tabla que contienen los valores de los ingresos de 2020, 2021 y 2022. El extracto de la tabla se incluye en el resultado, junto con el título del documento, un enlace al documento completo y cualquier otro campo del documento que decida incluir.

Los extractos de la tabla se pueden mostrar en los resultados de la búsqueda tanto si la información se encuentra en una celda de una tabla como en varias celdas. Por ejemplo,Amazon Kendrapuede mostrar un extracto de la tabla adaptado a cada uno de estos tipos de consultas:

- «tarjeta de crédito con la tasa de interés más alta en 2020»
- «tarjeta de crédito con la tasa de interés más alta de 2020-2022»
- «Las 3 tarjetas de crédito con tasas de interés más altas en 2020-2022»
- «tarjetas de crédito con tasas de interés inferiores al 10%»
- «todas las tarjetas de crédito de bajo interés disponibles»

Amazon Kendraresalta la celda o celdas de la tabla que son más relevantes para la consulta. Las celdas más relevantes con sus filas, columnas y nombres de columnas correspondientes se muestran en el

resultado de la búsqueda. El extracto de la tabla muestra hasta cinco columnas y tres filas, según el número de celdas de la tabla que sean relevantes para la consulta y el número de columnas disponibles en la tabla original. La celda más relevante se muestra en el extracto de la tabla, junto con las siguientes celdas más relevantes.

La respuesta incluye el depósito de confianza (MEDIUM,HIGH,VERY_HIGH) para mostrar la relevancia de la respuesta de la tabla para la consulta. Si el valor de una celda de la tabla es VERY_HIGH de forma confidencial, luego se convierte en la «respuesta principal» y se resalta. Para valores de celdas de tabla que son HIGH en confianza, luego se destacan. Para valores de celdas de tabla que son MEDIUM en confianza, entonces no se destacan. La confianza general de la respuesta de la tabla se devuelve en la respuesta. Por ejemplo, si una tabla contiene principalmente celdas de tabla con HIGH confianza, entonces la confianza general devuelta en la respuesta de la tabla es HIGH confianza.

De forma predeterminada, las tablas no reciben un mayor nivel de importancia ni más peso que otros componentes de un documento. Dentro de un documento, si una tabla solo es ligeramente relevante para una consulta, pero hay un párrafo muy relevante, Amazon Kendra devuelve un extracto del párrafo. Los resultados de la búsqueda muestran el contenido que ofrece la mejor respuesta posible y la información más útil, en el mismo documento o en otros documentos. Si la confianza de una tabla es inferior MEDIUM confianza, entonces el extracto de la tabla no se devuelve en la respuesta.

Para utilizar la búsqueda tabular en un índice existente, debes volver a indexar el contenido.

Amazon Kendra soporta búsquedas tabulares [sinónimos](#) (incluidos sinónimos personalizados). Amazon Kendra solo admite documentos en inglés con tablas HTML que estén dentro de la etiqueta de la tabla.

El siguiente ejemplo muestra un extracto de la tabla incluido en el resultado de la consulta. Para ver un ejemplo de JSON con las respuestas a las consultas, incluidos extractos de tablas, consulte [Respuestas y tipos de consultas](#).

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("Format: " + 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.documentElement().text();
                    System.out.println(String.format("Title: %s", documentTitle));
                    String documentExcerpt = item.documentElement().text();
```

```
        System.out.println(String.format("Excerpt: %s", documentExcerpt));
        break;
    default:
        System.out.println(String.format("Unknown query result type: %s",
item.type()));
    }
}
System.out.println("-----\n");
}
```

Sugerencias de consulta

Amazon Kendra Sugerencias de consulta puede ayudar a los usuarios a escribir sus consultas de búsqueda más rápidamente y a guiar su búsqueda.

Amazon Kendra sugiere consultas relevantes para sus usuarios basándose en una de las siguientes opciones:

- Consultas populares en el historial de consultas o en el registro de consultas
- El contenido de los campos o atributos del documento

Puede establecer sus preferencias para utilizar el historial de consultas o los campos del documento configurando `suggestionTypes` como `Ya sea QUERY o DOCUMENT_ATTRIBUTES` y llamando [GetQuerySuggestions](#). De forma predeterminada, Amazon Kendra usa el historial de consultas para basar las sugerencias. Si los campos del historial de consultas y del documento están activados al llamar [UpdateQuerySuggestionsConfig](#) y no has configurado tu `suggestionTypes` preferencia de usar los campos del documento, luego Amazon Kendra usa el historial de consultas.

Si usa la consola, puede basar las sugerencias de consulta en el historial de consultas o en los campos del documento. Primero selecciona el índice y, a continuación, selecciona Sugerencias de consulta debajo de Enriquecimientos en el menú de navegación. A continuación, seleccione Configurar sugerencias de consulta. Tras configurar las sugerencias de consulta, se le dirigirá a una consola de búsqueda en la que podrá seleccionar la Historial de consultas o Campos de documentos en el panel derecho e introduzca una consulta de búsqueda en la barra de búsqueda.

De forma predeterminada, las sugerencias de consulta que utilizan el historial de consultas y los campos del documento se activan sin coste adicional. Puede desactivar este tipo de sugerencias de consulta en cualquier momento mediante el [UpdateQuerySuggestionsConfig API](#). Para desactivar las sugerencias de consulta basadas en el historial de consultas, defina `mode` a `DISABLED` al llamar [UpdateQuerySuggestionsConfig](#). Para desactivar las sugerencias de consulta basadas en los campos del documento, defina `attributeSuggestionsMode` a `INACTIVE` en la configuración de los campos del documento y, a continuación, llame [UpdateQuerySuggestionsConfig](#). Si utiliza la consola, puede desactivar las sugerencias de consulta en la Configuración de sugerencias de consultas.

Las sugerencias de consulta no distinguen entre mayúsculas y minúsculas. Amazon Kendra convierte el prefijo de consulta y la consulta sugerida a minúsculas, omite todas las comillas simples y dobles y reemplaza varios espacios en blanco por un solo espacio. Amazon Kendra coincide con todos los demás caracteres especiales tal como están. Amazon Kendra no muestra ninguna sugerencia si un usuario escribe menos de dos caracteres o más de 60 caracteres.

Temas

- [Sugerencias de consulta mediante el historial de consultas \(p. 508\)](#)

- [Sugerencias de consulta mediante campos de documentos \(p. 512\)](#)
- [Bloquear determinadas consultas o el contenido de los campos del documento para que no aparezcan en las sugerencias \(p. 515\)](#)

Sugerencias de consulta mediante el historial de consultas

Temas

- [Configuración para seleccionar consultas para sugerencias \(p. 508\)](#)
- [Borra las sugerencias y conserva el historial de consultas \(p. 511\)](#)
- [No hay sugerencias disponibles \(p. 512\)](#)

Puede optar por sugerir consultas relevantes para sus usuarios en función de las consultas populares del historial de consultas o del registro de consultas. Amazon Kendra utiliza todas las consultas que buscan los usuarios y aprende de estas consultas para hacer sugerencias a los usuarios. Amazon Kendra sugiere consultas populares a los usuarios cuando comienzan a escribir su consulta. Amazon Kendra sugiere una consulta si el prefijo o los primeros caracteres de la consulta coinciden con lo que el usuario empieza a escribir como consulta.

Por ejemplo, un usuario comienza a escribir la consulta «próximos eventos». Amazon Kendra ha aprendido del historial de consultas que muchos usuarios han buscado «próximos eventos en 2050» muchas veces. El usuario ve aparecer «próximos eventos en 2050» directamente debajo de la barra de búsqueda, completando automáticamente su consulta de búsqueda. El usuario selecciona esta sugerencia de consulta y el documento «Nuevos eventos: Qué está pasando en 2050» aparece en los resultados de la búsqueda.

Puede especificar cómo Amazon Kendra selecciona las consultas que cumplen los requisitos para sugerirlas a los usuarios. Por ejemplo, puede especificar que una sugerencia de consulta haya sido buscada por al menos 10 usuarios únicos (el valor predeterminado es tres), que se haya buscado en los últimos 30 días y que no contenga ninguna palabra o frase de la [lista de bloqueos](#). Amazon Kendra requiere que la consulta tenga al menos un resultado de búsqueda y contenga al menos una palabra de más de cuatro caracteres.

Configuración para seleccionar consultas para sugerencias

Puede configurar los siguientes parámetros para seleccionar consultas o sugerencias mediante la [UpdateQuerySuggestionsConfig API](#):

- Modo—Las sugerencias de consulta que utilizan el historial de consultas son: ENABLED o LEARN_ONLY. Amazon Kendra activa las sugerencias de consulta de forma predeterminada. LEARN_ONLY desactiva las sugerencias de consulta. Si está desactivado, Amazon Kendra sigue aprendiendo sugerencias, pero no hace sugerencias de consultas a los usuarios.
- Ventana de tiempo de registro de consultas—Qué tan recientes son las consultas en la ventana de tiempo del registro de consultas. La ventana de tiempo es un valor entero para el número de días desde el día actual hasta los días anteriores.
- Consultas sin información de usuario—Definido en TRUE para incluir todas las consultas, o configurarlo como FALSE para incluir solo consultas con información de usuario. Puede utilizar esta configuración si la aplicación de búsqueda incluye información del usuario, como el identificador de usuario, cuando un usuario realiza una consulta. De forma predeterminada, esta configuración no filtra las consultas si no hay información de usuario específica asociada a las consultas. Sin embargo, puede utilizar esta configuración para hacer sugerencias únicamente a partir de consultas que incluyan información del usuario.

- **Usuarios únicos**—El número mínimo de usuarios únicos que deben buscar una consulta para que la consulta pueda sugerirla a los usuarios. Este número es un valor entero.
- **Recuento de consultas**—El número mínimo de veces que se debe buscar una consulta para que pueda sugerirla a los usuarios. Este número es un valor entero.

Esta configuración afecta a la forma en que las consultas se seleccionan como consultas populares para sugerirlas a los usuarios. La forma en que ajustes la configuración dependerá de tus necesidades específicas, por ejemplo:

- Si tus usuarios suelen buscar una vez al mes en promedio, puedes establecer el número de días en la ventana de registro de consultas en 30 días. Al usar esa configuración, captura la mayoría de las consultas recientes de sus usuarios antes de que queden desactualizadas en la ventana de tiempo.
- Si solo una pequeña cantidad de sus consultas incluyen información de usuario y no desea sugerir consultas basadas en un tamaño de muestra pequeño, puede configurar las consultas para que incluyan a todos los usuarios.
- Si define las consultas populares como búsquedas realizadas por al menos 10 usuarios únicos y se buscan al menos 100 veces, establece los usuarios únicos en 10 y el recuento de consultas en 100.

Warning

Es posible que los cambios en la configuración no surtan efecto inmediatamente.

Puede realizar un seguimiento de los cambios de configuración mediante el [DescribeQuerySuggestionsConfig](#) API. El tiempo que tarda en surtir efecto la configuración actualizada depende de las actualizaciones que realices y del número de consultas de búsqueda en tu índice. Amazon Kendra actualiza automáticamente las sugerencias cada 24 horas, después de cambiar una configuración o después de aplicar una [lista de bloqueos](#).

CLI

Para recuperar sugerencias de consultas

```
aws kendra get-query-suggestions \
--index-id index-id \
--query-text "query-text" \
--suggestion-types '[ "QUERY" ]' \
--max-suggestions-count 1 // If you want to limit the number of suggestions
```

Para actualizar las sugerencias de consulta

Por ejemplo, para cambiar la ventana de tiempo del registro de consultas y el número mínimo de veces que se debe buscar una consulta:

```
aws kendra update-query-suggestions-config \
--index-id index-id \
--query-log-look-back-window-in-days 30 \
--minimum-query-count 100
```

Python

Para recuperar sugerencias de consultas

```
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
numSuggestions = 1

try:
    querySuggestionsResponse = kendra.get_query_suggestions(
        IndexId = index_id,
        QueryText = query_text,
        SuggestionTypes = query_suggestions_type,
        MaxSuggestionsCount = numSuggestions
    )

    # Print out the suggestions you received
    if ("Suggestions" in querySuggestionsResponse.keys()):
        for (suggestion: querySuggestionsResponse["Suggestions"]):
            print(suggestion["Value"]["Text"]["Text"]);
    }
}

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

print("Program ends.")
```

Para actualizar las sugerencias de consulta

Por ejemplo, para cambiar la ventana de tiempo del registro de consultas y el número mínimo de veces que se debe buscar una consulta:

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

kendra = boto3.client("kendra")

print("Updating query suggestions settings/configuration for an index.")

# Provide the index ID
index_id = "index-id"

# Configure the settings you want to update
minimum_query_count = 100
query_log_look_back_window_in_days = 30

try:
    kendra.update_query_suggestions_config(
        IndexId = index_id,
        MinimumQueryCount = minimum_query_count,
        QueryLogLookBackWindowInDays = query_log_look_back_window_in_days
    )
```

```
print("Wait for Amazon Kendra to update the query suggestions.")

while True:
    # Get query suggestions description of settings/configuration
    query_sugg_config_response = kendra.describe_querySuggestionsConfig(
        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.")
```

Borra las sugerencias y conserva el historial de consultas

Puede borrar las sugerencias de consulta mediante el [ClearQuerySuggestions](#) API. Al borrar las sugerencias, solo se eliminan las sugerencias de consulta existentes, no las consultas del historial de consultas. Cuando borras las sugerencias, Amazon Kendra aprende nuevas sugerencias en función de las nuevas consultas agregadas al registro de consultas desde el momento en que borraste las sugerencias.

CLI

Para borrar las sugerencias de consulta

```
aws kendra clear-query-suggestions \
--index-id index-id
```

Python

Para borrar las sugerencias de consulta

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Clearing out query suggestions for an index.")

# Provide the index ID
index_id = "index-id"

try:
    kendra.clear_querySuggestions(
        IndexId = index_id
    )

    # Confirm last cleared date-time and that there are no suggestions
    query_sugg_config_response = kendra.describe_querySuggestionsConfig(
        IndexId = index_id
    )
    print("Query Suggestions last cleared at: " +
        str(query_sugg_config_response["LastClearTime"]));
    print("Number of suggestions available from the time of clearing: " +
        str(query_sugg_config_response["TotalSuggestionsCount"]));
```

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

No hay sugerencias disponibles

Si no ves sugerencias para una consulta, podría deberse a uno de los siguientes motivos:

- No hay suficientes consultas en su índice para Amazon Kendra de lo que aprender.
- La configuración de las sugerencias de consultas es demasiado estricta, por lo que la mayoría de las consultas se excluyen de las sugerencias.
- Has borrado sugerencias recientemente, y Amazon Kendra todavía necesita tiempo para que se acumulen nuevas consultas a fin de aprender nuevas sugerencias.

Puede comprobar su configuración actual mediante el [DescribeQuerySuggestionsConfig API](#).

Sugerencias de consulta mediante campos de documentos

Temas

- [Ajustes para seleccionar campos para sugerencias \(p. 512\)](#)
- [Control de usuario en los campos del documento \(p. 515\)](#)

Puede optar por sugerir consultas relevantes para sus usuarios en función del contenido de los campos del documento. En lugar de utilizar el historial de consultas para sugerir otras consultas relevantes populares, puede utilizar la información contenida en un campo de documento que resulta útil para completar automáticamente la consulta. Amazon Kendra busca contenido relevante en los campos establecidos en `enSuggestable` y eso se alinea estrechamente con la consulta del usuario. Luego, Amazon Kendra sugiere este contenido al usuario cuando comience a escribir su consulta.

Por ejemplo, si especificas el campo de título en el que basar las sugerencias y un usuario empieza a escribir la consulta ««Cómo sabe Amazon...»», el título más relevante 'How' Amazon Kendra se podría sugerir 'works' para completar automáticamente la búsqueda. El usuario ve ««Cómo» Amazon Kendra las obras aparecen directamente debajo de la barra de búsqueda, completando automáticamente la consulta de búsqueda. El usuario selecciona esta sugerencia de consulta y el documento ««Cómo»» Amazon Kendra se devuelve 'works' en los resultados de la búsqueda.

Puede utilizar el contenido de cualquier campo del documento de `StringyStringList` para sugerir una consulta configurando el campo `enSuggestable` como parte de la configuración de campos para sugerencias de consultas. También puede utilizar una [lista de bloqueos](#) para que los usuarios no muestren los campos del documento sugeridos que contienen determinadas palabras o frases. Puede utilizar una lista de bloqueo. La lista de bloqueos se aplica tanto si configura las sugerencias de consulta para utilizar el historial de consultas o los campos del documento.

Ajustes para seleccionar campos para sugerencias

Puede configurar los siguientes parámetros para seleccionar los campos del documento para recibir sugerencias mediante [AttributeSuggestionsConfig](#) y llamando al [UpdateQuerySuggestionsConfig API](#) para actualizar la configuración a nivel de índice:

- Modo de sugerencias de campos/atributos—Las sugerencias de consulta mediante campos de documentos son:ACTIVEoINACTIVE.Amazon Kendra activa las sugerencias de consulta de forma predeterminada.
- Campos/atributos sugeridos—Los nombres de los campos o las claves de campo en los que basar las sugerencias. Estos campos deben estar configurados enTRUEporSuggestable, como parte de la configuración de los campos. Puede anular la configuración de los campos en el nivel de consulta y, al mismo tiempo, mantener la configuración en el nivel de índice. Utilice el[GetQuerySuggestionsAPI](#) para cambiarAttributeSuggestionConfiga nivel de consulta. Esta configuración a nivel de consulta puede resultar útil para experimentar rápidamente con el uso de diferentes campos de documentos sin tener que actualizar la configuración a nivel de índice.
- Campos/atributos adicionales—Los campos adicionales que desea incluir en la respuesta a una sugerencia de consulta. Estos campos se utilizan para proporcionar información adicional en la respuesta; sin embargo, no se utilizan para basar las sugerencias.

Warning

Es posible que los cambios en la configuración no surtan efecto inmediatamente. Puede realizar un seguimiento de los cambios de configuración mediante el[DescribeQuerySuggestionsConfigAPI](#). El tiempo que tarda en surtir efecto la configuración actualizada depende de las actualizaciones que realice.Amazon Kendra actualiza automáticamente las sugerencias cada 24 horas, después de cambiar una configuración o después de aplicar una lista de bloques.

CLI

Para recuperar las sugerencias de consulta y anular la configuración de los campos del documento a nivel de consulta en lugar de tener que cambiar la configuración a nivel de índice.

```
aws kendra get-query-suggestions \
--index-id index-id \
--query-text "query-text" \
--suggestion-types '["DOCUMENT_ATTRIBUTES"]' \
--attribute-suggestions-config '{"SuggestionAttributes":' '[{"field/attribute key 1", \
"field/attribute key 2"}]', "AdditionalResponseAttributes":' [{"response field/attribute key 1", "response field/attribute key 2"}]'}' \
--max-suggestions-count 1 // If you want to limit the number of suggestions
```

Para actualizar las sugerencias de consulta

Por ejemplo, para cambiar la configuración de los campos del documento a nivel de índice:

```
aws kendra update-query-suggestions-config \
--index-id index-id \
--attribute-suggestions-config '{"SuggestableConfigList": ' '[{"SuggestableConfig": \
"_document_title", "Suggestable": true}]', "AttributeSuggestionsMode": "ACTIVE"}'
```

Python

Para recuperar las sugerencias de consulta y anular la configuración de los campos del documento a nivel de consulta en lugar de tener que cambiar la configuración a nivel de índice.

```
import boto3
from botocore.exceptions import ClientError

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

```
print("Get query suggestions.")

# Provide the index ID
index_id = "index-id"

# Provide the query text
query_text = "query"

# Provide the query suggestions type
query_suggestions_type = "DOCUMENT_ATTRIBUTES"

# Override fields/attributes configuration at query level
configuration = {"SuggestionAttributes":
    '[{"field/attribute key 1", "field/attribute key 2"}]',
    "AdditionalResponseAttributes":
        '[{"response field/attribute key 1", "response field/attribute key 2"}]'
}

# If you want to limit the number of suggestions
numSuggestions = 1

try:
    query_suggestions_response = kendra.get_query_suggestions(
        IndexId = index_id,
        QueryText = query_text,
        SuggestionTypes = [query_suggestions_type],
        AttributeSuggestionsConfig = configuration,
        MaxSuggestionsCount = numSuggestions
    )

    # Print out the suggestions you received
    if ("Suggestions" in query_suggestions_response.keys()):
        for (suggestion: query_suggestions_response["Suggestions"]):
            print(suggestion["Value"]["Text"]["Text"]);
    }
}

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

print("Program ends.")
```

Para actualizar las sugerencias de consulta

Por ejemplo, para cambiar la configuración de los campos del documento a nivel de índice:

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

kendra = boto3.client("kendra")

print("Updating query suggestions settings/configuration for an index.")

# Provide the index ID
index_id = "index-id"

# Configure the settings you want to update at the index level
configuration = {"SuggestableConfigList":
    '[{"SuggestableConfig": {"_document_title", "Suggestable": true}},',
    "AttributeSuggestionsMode": "ACTIVE"
}

try:
```

```
kendra.update_query_suggestions_config(  
    IndexId = index_id,  
    AttributeSuggestionsConfig = configuration  
)  
  
print("Wait for Amazon Kendra to update the query suggestions.")  
  
while True:  
    # Get query suggestions description of settings/configuration  
    query_sugg_config_response = kendra.describe_query_suggestions_config(  
        IndexId = index_id  
)  
  
    # If status is not UPDATING, then quit  
    status = query_sugg_config_response["Status"]  
    print(" Updating query suggestions config. Status: " + status)  
    if status != "UPDATING":  
        break  
    time.sleep(60)  
  
except ClientError as e:  
    print("%s" % e)  
  
print("Program ends.")
```

Control de usuario en los campos del documento

Puede aplicar el filtrado de contexto de usuario a los campos del documento en los que desee basar las sugerencias de consulta. Esto filtra la información de los campos del documento en función del acceso del usuario o de su grupo a los documentos. Por ejemplo, un becario busca en el portal de la empresa y no tiene acceso a un documento empresarial ultrasecreto. Por lo tanto, las consultas sugeridas basadas en el título del documento ultrasecreto o en cualquier otro campo sugerible no se muestran al becario.

Puede indexar sus documentos con una lista de control de acceso (ACL), que defina a qué usuarios y grupos se les asigna acceso a qué documentos. A continuación, puede aplicar el filtrado de contexto de usuario a los campos de sus documentos para obtener sugerencias de consultas. El filtrado de contexto de usuario que está configurado actualmente para el índice es el mismo filtrado de contexto de usuario que se aplica a la configuración de los campos del documento para las sugerencias de consulta. El filtrado del contexto de usuario forma parte de la configuración de los campos del documento. Usas el [AttributeSuggestionsGetConfigy](#) llama [GetQuerySuggestions](#).

Bloquear determinadas consultas o el contenido de los campos del documento para que no aparezcan en las sugerencias

UNlista de bloquesosse detieneAmazon Kendra desde sugerir ciertas consultas a tus usuarios. Una lista de bloqueo es una lista de palabras o frases que quieres excluir de las sugerencias de consulta. Amazon Kendra excluye las consultas que contienen una coincidencia exacta de las palabras o frases de la lista de bloqueo.

Puede utilizar una lista de bloqueo para protegerse de las palabras o frases ofensivas que suelen aparecer en el historial de consultas o en los campos del documento y que Amazon Kendra podría seleccionarlo como sugerencias. Una lista de bloqueo también puede impedir Amazon Kendra sugerir consultas que contengan información que no esté lista para ser divulgada o anunciada públicamente. Por ejemplo, sus usuarios consultan con frecuencia sobre el próximo lanzamiento de un posible producto nuevo. Sin embargo, no querrás sugerir el producto porque no estás preparado para lanzarlo. Puedes bloquear

las consultas que contienen el nombre y la información del producto para que no aparezcan en las sugerencias.

Puede crear una lista de bloqueo para consultas mediante el[CreateQuerySuggestionsBlockListAPI](#). Coloca cada palabra o frase del bloqueo en una línea independiente de un archivo de texto. A continuación, sube el archivo de texto a su bucket de Amazon S3 y proporciona la ruta o la ubicación del archivo en Amazon S3. Amazon Kendra actualmente admite la creación de una sola lista de bloqueo.

Puedes reemplazar el archivo de texto de tus palabras y frases bloqueadas en tu Amazon S3 balde. Para actualizar la lista de bloques en Amazon Kendra, utilice el[UpdateQuerySuggestionsBlockListAPI](#).

Utilice el[DescribeQuerySuggestionsBlockListAPI](#) para obtener el estado de tu lista de bloqueo. `DescribeQuerySuggestionsBlockList` también puede proporcionarle otra información útil, como la siguiente:

- Cuándo se actualizó tu lista de bloqueo por última vez
- Cuántas palabras o frases hay en tu lista de bloqueo actual
- Mensajes de error útiles al crear una lista de bloqueo

También puede utilizar el[ListQuerySuggestionsBlockListsAPI](#) para obtener una lista de resúmenes de listas de bloques para un índice.

Para eliminar tu lista de bloqueo, usa la[DeleteQuerySuggestionsBlockListAPI](#).

Es posible que las actualizaciones de la lista de bloqueados no surtan efecto de inmediato. Puede realizar un seguimiento de las actualizaciones mediante el[DescribeQuerySuggestionsBlockListAPI](#).

CLI

Para crear una lista de bloqueo

```
aws kendra create-query-suggestions-block-list \
--index-id index-id \
--name "block-list-name" \
--description "block-list-description" \
--source-s3-path "Bucket=bucket-name,Key=query-suggestions/block_list.txt" \
--role-arn role-arn
```

Para actualizar una lista de bloqueo

```
aws kendra update-query-suggestions-block-list \
--index-id index-id \
--name "new-block-list-name" \
--description "new-block-list-description" \
--source-s3-path "Bucket=bucket-name,Key=query-suggestions/new_block_list.txt" \
--role-arn role-arn
```

Para eliminar una lista de bloqueo

```
aws kendra delete-query-suggestions-block-list \
--index-id index-id \
--id block-list-id
```

Python

Para crear una lista de bloqueo

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

kendra = boto3.client("kendra")

print("Create a query suggestions block list.")

# Provide a name for the block list
block_list_name = "block-list-name"
# Provide an optional description for the block list
block_list_description = "block-list-description"
# Provide the IAM role ARN required for query suggestions block lists
block_list_role_arn = "role-arn"

# Provide the index ID
index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "query-suggestions/block_list.txt"
source_s3_path = {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    block_list_response = kendra.create_querySuggestionsBlockList(
        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.describeQuerySuggestionsBlockList(
            Id = block_list_id,
            IndexId = index_id
        )
        # If status is not CREATING, then quit
        status = block_list_description["Status"]
        print("Creating block list. Status: " + status)
        if status != "CREATING":
            break
        time.sleep(60)

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

print("Program ends.")
```

Para actualizar una lista de bloqueo

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

```
import time

kendra = boto3.client("kendra")

print("Update a block list for query suggestions.")

# Provide the block list name you want to update
block_list_name = "new-block-list-name"
# Provide the block list description you want to update
block_list_description = "new-block-list-description"
# Provide the IAM role ARN required for query suggestions block lists
block_list_role_arn = "role-arn"

# Provide the block list ID
block_list_id = "block-list-id"
# Provide the index ID
index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "query-suggestions/new_block_list.txt"
source_s3_path = {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    kendra.update_querySuggestionsBlockList(
        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.describeQuerySuggestionsBlockList(
            Id = block_list_id,
            IndexId = index_id
        )
        # If status is not UPDATING, then the update has finished
        status = block_list_description["Status"]
        print("Updating block list. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

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

print("Program ends.")
```

Para eliminar una lista de bloqueo

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Delete a block list for query suggestions.")

# provide the block list ID
```

```
query_suggestions_block_list_id = "query-suggestions-block-list-id"
# Provide the index ID
index_id = "index-id"

try:
    kendra.delete_querySuggestionsBlockList(
        Id = query_suggestions_block_list_id,
        IndexId = index_id
    )

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

print("Program ends.")
```

Corrector ortográfico de consultas

Amazon Kendra Corrector ortográficosugiere correcciones ortográficas para una consulta. Esto puede ayudarte a reducir al mínimo la aparición de cero resultados de búsqueda y a obtener resultados relevantes. Es posible que tus usuarios recibancero resultados de búsquedade consultas mal escritas sin resultados coincidentes o sin documentos devueltos. O bien, tus usuarios podrían recibirresultados de búsqueda irrelevantesde consultas mal escritas.

El corrector ortográfico está diseñado para sugerir correcciones de palabras mal escritas en función de las palabras que aparecen en los documentos indexados y del grado de coincidencia entre una palabra corregida y una palabra mal escrita. Por ejemplo, si la palabra «estados financieros» aparece en los documentos indexados, podría coincidir estrechamente con la palabra mal escrita «estados financieros» de la consulta «estados financieros de fin de año».

El corrector ortográfico devuelve las palabras deseadas o corregidas que sustituyen a las palabras mal escritas en el texto de consulta original. Por ejemplo, «implementar la búsqueda de kendre» podría devolver «implementar la búsqueda de Kendra». También puedes utilizar las ubicaciones de desplazamiento que se proporcionan en la API para resaltar o poner en cursiva las palabras corregidas devueltas en una consulta de tu aplicación de interfaz. En la consola, las palabras corregidas aparecen resaltadas o en cursiva de forma predeterminada. Por ejemplo, 'desplegando Kendrabuscar'.

En el caso de los términos especializados o específicos de la empresa que aparecen en los documentos indexados, el corrector ortográfico no malinterpreta estos términos como errores ortográficos en la consulta. Por ejemplo, «amazon macie» no se corrige a «amazon mace».

En el caso de las palabras separadas por guiones, como «fin de año», el corrector ortográfico las trata como palabras individuales para sugerir correcciones para estas palabras. Por ejemplo, la corrección sugerida para «fin de año» podría ser «fin de año».

ParaDOCUMENTyQUESTION_ANSWERtipos de respuesta a consultas, el corrector ortográfico sugiere correcciones a las palabras mal escritas en función de las palabras del cuerpo del documento. El cuerpo del documento es más fiable que el título para sugerir correcciones que coincidan estrechamente con las palabras mal escritas. ParaANSWERtipos de respuesta a consultas, el corrector ortográfico sugiere correcciones basadas en las palabras del documento de preguntas y respuestas predeterminado del índice.

Puede activar el corrector ortográfico mediante el[SpellCorrectionConfiguration](#)objeto. Tú configuras`IncludeQuerySpellCheckSuggestions`aTRUE. El corrector ortográfico está activado de forma predeterminada en la consola. Está integrado en la consola de forma predeterminada.

El corrector ortográfico también puede sugerir correcciones ortográficas para consultas en varios idiomas, no solo en inglés. Para obtener una lista de los idiomas compatibles con el corrector ortográfico, consulte[Amazon Kendra idiomas admitidos](#).

Uso del corrector ortográfico de consultas con límites predeterminados

El corrector ortográfico está diseñado con ciertos valores predeterminados o límites. La siguiente es una lista de los límites actuales que se aplican al activar las sugerencias de corrección ortográfica.

- Las correcciones ortográficas sugeridas no se pueden devolver para palabras de menos de tres caracteres o más de 30 caracteres. Para permitir más de 30 caracteres o menos de tres caracteres, póngase en contacto con[Soporte](#).
- Las correcciones ortográficas sugeridas no pueden restringir las sugerencias en función del control de acceso del usuario o de su lista de control de acceso para[filtrado de contexto de usuario](#). Las correcciones ortográficas se basan en todas las palabras de los documentos indexados, independientemente de si las palabras están restringidas a ciertos usuarios o no. Si desea evitar que determinadas palabras aparezcan en las correcciones ortográficas sugeridas para las consultas, no active[SpellCorrectionConfiguration](#).
- Las correcciones ortográficas sugeridas no se pueden devolver para palabras que incluyan números. Por ejemplo, 'how 2 not br8k ubun2'.
- Las correcciones ortográficas sugeridas no pueden utilizar palabras que no aparezcan en los documentos indexados.
- Las correcciones ortográficas sugeridas no pueden utilizar palabras con una frecuencia inferior al 0,01 por ciento en los documentos indexados. Para cambiar el umbral del 0,01%, póngase en contacto con[Soporte](#).

Filtrado y búsqueda de facetas

Puede mejorar los resultados de la búsqueda o la respuesta desde[ConsultaAPI](#) mediante filtros. Los filtros restringen los documentos de la respuesta a los que se aplican directamente a la consulta. Para crear sugerencias de búsqueda por facetas, utilice la lógica booleana para filtrar los atributos específicos del documento de la respuesta o los documentos que no coincidan con criterios específicos. Puede especificar facetas mediante el[Facets](#) parámetro en el[QueryAPI](#).

Para buscar documentos que haya indexado con Amazon Kendra o Amazon Lex, utilizar[AMAZON.KendraSearchIntent](#). Para ver un ejemplo de configuración Amazon Kendra con Amazon Lex, consulte[Creación de un bot de preguntas frecuentes para un Amazon Kendra Índice](#). También puede proporcionar un filtro para la respuesta mediante[AttributeFilter](#). Este es el filtro de consulta en JSON cuando se configura `AMAZON.KendraSearchIntent`. Para proporcionar un filtro de atributos al configurar una intención de búsqueda en la consola, vaya al editor de intenciones y elija Amazon Kendra consulta para proporcionar un filtro de consulta en JSON. Para obtener más información sobre `AMAZON.KendraSearchIntent`, consulte el[Amazon Lex Guía de documentación](#).

Facetas

Las facetas son vistas con alcance de un conjunto de resultados de búsqueda. Por ejemplo, puede proporcionar resultados de búsqueda para ciudades de todo el mundo, donde los documentos se filtran por una ciudad específica a la que están asociados. O bien, puede crear facetas para mostrar los resultados de un autor específico.

Puede utilizar un atributo de documento o un campo de metadatos asociado a un documento como faceta para que los usuarios puedan buscar por categorías o valores dentro de esa faceta. También puede mostrar facetas anidadas en los resultados de la búsqueda para que los usuarios puedan buscar no solo por categoría o campo, sino también por subcategoría o subcampo.

El siguiente ejemplo muestra cómo obtener información de facetas para el atributo personalizado «Ciudad».

```
response=kendra.query(  
    QueryText = query,  
    IndexId = index,  
    Facets = [  
        {  
            "DocumentAttributeKey" : "City"  
        }  
    ]  
)
```

Puede utilizar facetas anidadas para restringir aún más la búsqueda. Por ejemplo, el atributo o faceta del documento «Ciudad» incluye un valor denominado «Seattle». Además, el atributo o faceta del documento «CityRegion» incluye los valores «Norte» y «Sur» para los documentos asignados a «Seattle». Puede mostrar las facetas anidadas con sus recuentos en los resultados de la búsqueda para que los documentos se puedan buscar no solo por ciudad sino también por región dentro de una ciudad.

Tenga en cuenta que las facetas anidadas pueden afectar a la latencia de las consultas. Una regla general es que cuantas más facetas anidadas utilice, mayor será el impacto potencial en la latencia. Otros factores que afectan a la latencia son el tamaño promedio de los documentos indexados, el tamaño del índice, las consultas altamente complejas y la carga total de Amazon Kendra índice.

El siguiente ejemplo muestra cómo obtener información sobre las facetas de «CityRegion» atributo personalizado, como una faceta anidada dentro de «Ciudad».

```
response=kendra.query(  
    QueryText = query,  
    IndexId = index,  
    Facets = [  
        {  
            "DocumentAttributeKey" : "City",  
            "Facets": [  
                {  
                    "DocumentAttributeKey" : "CityRegion"  
                }  
            ]  
        }  
    ]  
)
```

La información de facetas, como el recuento de documentos, se devuelve en el `FacetResults` matriz de respuestas. El contenido se utiliza para mostrar sugerencias de búsqueda por facetas en la aplicación. Por ejemplo, si el atributo «Ciudad» del documento contiene la ciudad a la que podría aplicarse una búsqueda, utilice esa información para mostrar una lista de búsquedas por ciudades. Los usuarios pueden elegir una ciudad para filtrar los resultados de búsqueda. Para realizar la búsqueda por facetas, llame a [Consultar](#). Utilice la API y utilice el atributo de documento elegido para filtrar los resultados.

Puede mostrar hasta 10 valores de faceta por faceta para una consulta y solo una faceta anidada dentro de una faceta. Si desea aumentar estos límites, póngase en contacto con [Soporte](#). Si desea limitar el número de valores de faceta por faceta a menos de 10, puede especificarlo en `FacetObject`.

El siguiente ejemplo de respuesta JSON muestra las facetas incluidas en el atributo de documento «Ciudad». La respuesta incluye el recuento de documentos para el valor de la faceta.

```
{  
    'FacetResults': [  
        {  
            'DocumentAttributeKey': 'City',  
            'DocumentAttributeValueCountPairs': [  
                {  
                    'Value': 'Seattle',  
                    'Count': 100  
                },  
                {  
                    'Value': 'Los Angeles',  
                    'Count': 50  
                },  
                {  
                    'Value': 'Chicago',  
                    'Count': 30  
                },  
                {  
                    'Value': 'Houston',  
                    'Count': 20  
                },  
                {  
                    'Value': 'Phoenix',  
                    'Count': 15  
                },  
                {  
                    'Value': 'San Antonio',  
                    'Count': 10  
                },  
                {  
                    'Value': 'San Diego',  
                    'Count': 8  
                },  
                {  
                    'Value': 'Dallas',  
                    'Count': 7  
                },  
                {  
                    'Value': 'Austin',  
                    'Count': 5  
                },  
                {  
                    'Value': 'Philadelphia',  
                    'Count': 4  
                }  
            ]  
        }  
    ]  
}
```

```
{  
    'Count': 3,  
    'DocumentAttributeValue': {  
        'StringValue': 'Dubai'  
    }  
},  
{  
    'Count': 3,  
    'DocumentAttributeValue': {  
        'StringValue': 'Seattle'  
    }  
},  
{  
    'Count': 1,  
    'DocumentAttributeValue': {  
        'StringValue': 'Paris'  
    }  
}  
]  
}
```

También puedes mostrar la información de facetas de una faceta anidada, como una región de una ciudad, para filtrar aún más los resultados de la búsqueda.

El siguiente ejemplo de respuesta JSON muestra las facetas incluidas en el ámbito de «CityRegion» atributo de documento», como una faceta anidada dentro de «Ciudad». La respuesta incluye el recuento de documentos para los valores de las facetas anidadas.

```
{  
    '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'
            }
          }
        ]
      }
    ]
  }
]
```

Al utilizar un campo de lista de cadenas para crear facetas, los resultados de facetas devueltos se basan en el contenido de la lista de cadenas. Por ejemplo, si tiene un campo de lista de cadenas que contiene dos elementos, uno con la lista «dachshund», «salchicha» y otro con el valor «husky», obtendrá `FacetResults` con tres facetas.

Para obtener más información, consulte [Respuestas a consultas y tipos de respuestas \(p. 538\)](#).

Uso de los atributos del documento para filtrar los resultados de la búsqueda

De forma predeterminada, Query devuelve todos los resultados de la búsqueda. Para filtrar las respuestas, puede realizar operaciones lógicas en los atributos del documento. Por ejemplo, si solo quieres documentos para una ciudad específica, puedes filtrar los atributos de documento personalizados «Ciudad» y «Estado». Usas [AttributeFilter](#) para crear una operación booleana en los filtros que proporcione.

La mayoría de los atributos se pueden usar para filtrar las respuestas de todos [tipos de respuesta](#). Sin embargo, el `_excerpt_page_number` el atributo solo se aplica a ANSWER tipos de respuesta al filtrar las respuestas.

El siguiente ejemplo muestra cómo realizar una operación AND lógica filtrando por una ciudad específica, Seattle, y estado, Washington.

```
response=kendra.query(  
    QueryText = query,  
    IndexId = index,  
    AttributeFilter = {'AndAllFilters':  
        [  
            {"EqualsTo": {"Key": "City", "Value": {"StringValue": "Seattle"}},  
            {"EqualsTo": {"Key": "State", "Value": {"StringValue": "Washington"}}}  
        ]  
    }  
)
```

El siguiente ejemplo muestra cómo realizar una operación OR lógica para cuando alguno de las claves `Fileformat`, `Author`, o `SourceURI` las coinciden con los valores especificados.

```
response=kendra.query(  
    QueryText = query,  
    IndexId = index,  
    AttributeFilter = {'OrAllFilters':  
        [  
            {"EqualsTo": {"Key": "Fileformat", "Value": {"StringValue": "AUTO_DETECT"}},  
            {"EqualsTo": {"Key": "Author", "Value": {"StringValue": "Ana Carolina"}},  
            {"EqualsTo": {"Key": "SourceURI", "Value": {"StringValue": "https://aws.amazonaws.com/234234242342"}},  
        ]  
    }  
)
```

Para `StringList` campos, utilice el `ContainsAny` o `ContainsAll` filtros de atributos para devolver documentos con la cadena especificada. El siguiente ejemplo muestra cómo devolver todos los documentos que tengan los valores «Seattle» o «Portland» en su `Locations` atributo personalizado.

```
response=kendra.query(  
    QueryText = query,  
    IndexId = index,  
    AttributeFilter = {  
        "ContainsAny": { "Key": "Locations", "Value": { "StringListValue":  
            [ "Seattle", "Portland" ] } }  
    }  
)
```

Filtrar los atributos de cada documento en los resultados de búsqueda

Amazon Kendra devuelve los atributos del documento para cada documento de los resultados de la búsqueda. Puede filtrar ciertos atributos del documento que deseé incluir en la respuesta como parte de los resultados de la búsqueda. De forma predeterminada, todos los atributos del documento asignados a un documento se devuelven en la respuesta.

En el siguiente ejemplo, solo el `_source_uri` y `author` los atributos del documento se incluyen en la respuesta de un documento.

```
response=kendra.query(  
    QueryText = query,  
    IndexId = index,
```

```
    RequestedDocumentAttributes = ["_source_uri", "_author"]  
)
```

Filtrar según el contexto del usuario

Puede filtrar los resultados de búsqueda de un usuario en función del acceso del usuario o de su grupo a los documentos. Puede utilizar un identificador de usuario, un identificador de usuario o un atributo de usuario para filtrar documentos. Amazon Kendra también puede asignar usuarios a sus grupos. Puede optar por utilizar AWS IAM Identity Center (successor to AWS Single Sign-On) como almacenamiento o fuente de identidad.

El filtrado de contexto de usuario es un tipo de búsqueda personalizada con la ventaja de controlar el acceso a los documentos. Por ejemplo, no todos los equipos que buscan información en el portal de la empresa deben acceder a documentos de alto secreto de la empresa, ni estos documentos son relevantes para todos los usuarios. Solo los usuarios o grupos de equipos específicos a los que se les dé acceso a documentos de alto secreto deberían ver estos documentos en los resultados de búsqueda.

Cuando un documento se indexa en Amazon Kendra, se incluye la lista de control de acceso (ACL) correspondiente para la mayoría de los documentos. La ACL especifica qué nombres de usuario y nombres de grupos tienen permitido o denegado el acceso al documento. Los documentos sin una ACL son documentos públicos.

Amazon Kendra extrae automáticamente la información de usuario o grupo asociada a cada documento en la mayoría de las fuentes de datos. Por ejemplo, un documento de Quip puede incluir una lista «compartida» de determinados usuarios o grupos que tienen acceso al documento. Si utilizas un bucket de S3 como fuente de datos, proporcionas un [Archivo JSON](#) para su ACL e incluya la ruta S3 a este archivo como parte de la configuración de la fuente de datos. Si agrega documentos directamente a un índice, especifica la ACL en el [Directo](#) o [objeto](#) como parte del objeto de documento en el [BatchPutDocumentAPI](#).

Puede utilizar el [CreateAccessControlConfigurationAPI](#) para reconfigurar el control de acceso a nivel de documento existente sin volver a indexar todos los documentos. Por ejemplo, su índice contiene documentos empresariales de alto secreto a los que solo ciertos empleados o usuarios deben acceder. Uno de estos usuarios deja la empresa o pasa a un equipo al que se le debería impedir el acceso a documentos de alto secreto. El usuario sigue teniendo acceso a los documentos de alto secreto porque tenía acceso cuando sus documentos estaban indexados anteriormente. Puede crear una configuración de control de acceso específica para el usuario con acceso denegado. Más adelante, puede actualizar la configuración del control de acceso para permitir el acceso en caso de que el usuario regrese a la empresa y vuelva a unirse al equipo «ultrasecreto». Puede volver a configurar el control de acceso a sus documentos a medida que cambien las circunstancias.

Para aplicar su configuración de control de acceso a ciertos documentos, llame al [BatchPutDocumentAPI](#) con `AccessControlConfigurationId` incluido en el [Documento](#) objeto. Si utiliza un bucket de S3 como fuente de datos, actualiza el `.metadata.json` con el `AccessControlConfigurationId` y sincroniza tu fuente de datos. Amazon Kendra actualmente solo admite la configuración de control de acceso para fuentes de datos de S3 y documentos indexados mediante el [BatchPutDocumentAPI](#).

Filtrar por token de usuario

Al consultar un índice, puede utilizar un identificador de usuario para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Cuando emites una consulta, Amazon Kendra extrae y valida el token, extrae y comprueba la información del usuario y del grupo y ejecuta la consulta. Se devuelven todos los documentos a los que el usuario tiene acceso, incluidos los documentos públicos. Para obtener más información, consulte [Control de acceso de usuarios basado en tokens](#).

Usted proporciona el token de usuario en el [UserContext](#) objeto y pasar esto al [ConsultaAPI](#).

A continuación se muestra cómo incluir un token de usuario.

```
response = kendra.query(  
    QueryText = query,  
    IndexId = index,  
    UserToken = {  
        Token = "token"  
    })
```

Puede asignar usuarios a grupos. Cuando se utiliza el filtrado de contexto de usuario, no es necesario incluir todos los grupos a los que pertenece un usuario cuando se ejecuta la consulta. Con el [PutPrincipalMapping](#) API, puede asignar usuarios a sus grupos. Si no desea utilizar el [PutPrincipalMapping](#) API, debes proporcionar el nombre de usuario y todos los grupos a los que pertenece el usuario cuando emitas una consulta. También puede obtener los niveles de acceso de los grupos y usuarios de su fuente de identidad de IAM Identity Center mediante el [UserGroupResolutionConfiguration](#) objeto.

Filtrar por ID de usuario y grupo

Al consultar un índice, puede utilizar el identificador de usuario y el grupo para filtrar los resultados de la búsqueda en función del acceso del usuario o de su grupo a los documentos. Cuando emites una consulta, Amazon Kendra comprueba la información del usuario y del grupo y ejecuta la consulta. Se devuelven todos los documentos relevantes para la consulta a los que el usuario tiene acceso, incluidos los documentos públicos.

También puede filtrar los resultados de la búsqueda por fuentes de datos a las que tienen acceso los usuarios y los grupos. Especificar una fuente de datos es útil si un grupo está vinculado a varias fuentes de datos, pero solo desea que el grupo acceda a los documentos de una fuente de datos determinada. Por ejemplo, los grupos «Investigación», «Ingeniería» y «Ventas y marketing» están todos vinculados a los documentos de la empresa almacenados en las fuentes de datos Confluence y Salesforce. Sin embargo, el equipo de «Ventas y marketing» solo necesita acceder a los documentos relacionados con el cliente almacenados en Salesforce. De este modo, cuando los usuarios de ventas y marketing buscan documentos relacionados con los clientes, pueden ver los documentos de Salesforce en sus resultados. Los usuarios que no trabajan en ventas y marketing no ven los documentos de Salesforce en los resultados de búsqueda.

Usted proporciona información sobre los usuarios, los grupos y las fuentes de datos en el [UserContext](#) objeto y pasar esto al [Consulta](#) API. El identificador de usuario y la lista de grupos y fuentes de datos deben coincidir con el nombre que especifique en [Directora](#) objeto para identificar el usuario, los grupos y las fuentes de datos. Con el [Principal](#) objeto, puede añadir un usuario, grupo o fuente de datos a una lista de permitidos o denegados para acceder a un documento.

Debe proporcionar uno de los siguientes datos:

- Información de usuarios y grupos e información de fuentes de datos (opcional).
- Solo la información del usuario si asigna a sus usuarios a grupos y fuentes de datos mediante el [PutPrincipalMapping](#) API. También puede obtener los niveles de acceso de los grupos y usuarios de su fuente de identidad de IAM Identity Center mediante el [UserGroupResolutionConfiguration](#) objeto.

Si esta información no está incluida en la consulta, Amazon Kendra devuelve todos los documentos. Si proporciona esta información, solo se devolverán los documentos con identificadores de usuario, grupos y fuentes de datos coincidentes.

A continuación se muestra cómo incluir el identificador de usuario, los grupos y las fuentes de datos.

```
response = kendra.query(  
    QueryText = query,
```

```
IndexId = index,
UserId = {
    UserId = "user1"
},
Groups = {
    Groups = ["Sales and Marketing"]
},
DataSourceGroups = {
    DataSourceGroups = [{"DataSourceId" : "SalesforceCustomerDocsGroup", "GroupId": "Sales and Marketing"}]
})
```

Filtrar por atributo de usuario

Al consultar un índice, puede utilizar atributos integrados _user_id y _group_id para filtrar los resultados de la búsqueda en función del acceso del usuario y su grupo a los documentos. Puede configurar hasta 100 identificadores de grupo. Cuando emites una consulta, Amazon Kendra comprueba la información del usuario y del grupo y ejecuta la consulta. Se devuelven todos los documentos relevantes para la consulta a los que el usuario tiene acceso, incluidos los documentos públicos.

Usted proporciona los atributos de usuario y grupo en [AttributeFilter](#) objeto y pasar esto al [Consulta API](#).

El siguiente ejemplo muestra una solicitud que filtra la respuesta a la consulta en función del identificador de usuario y de los grupos «HR» e «IT», a los que pertenece el usuario. La consulta devolverá cualquier documento que tenga al usuario o a los grupos «HR» o «IT» en la lista de permitidos. Si el usuario o alguno de los grupos se encuentra en la lista de denegación de un documento, el documento no se devuelve.

```
response = kendra.query(
    QueryText = query,
    IndexId = index,
    AttributeFilter = {
        "OrAllFilters": [
            {
                "EqualsTo": {
                    "Key": "_user_id",
                    "Value": {
                        "StringValue": "user1"
                    }
                }
            },
            {
                "EqualsTo": {
                    "Key": "_group_ids",
                    "Value": {
                        "StringListValue": ["HR", "IT"]
                    }
                }
            }
        ]
    }
)
```

También puede especificar a qué fuente de datos puede acceder un grupo en la `Principal` objeto.

Note

El filtrado del contexto de usuario no es un control de autenticación o autorización para tu contenido. No realiza la autenticación del usuario ni de los grupos enviados al `Query API`. Dependiendo de su aplicación garantizar que la información de usuario y grupo enviada a la `Query API` está autenticada y autorizada.

Hay una implementación del filtrado de contexto de usuario para cada fuente de datos. En la siguiente sección se describe cada implementación.

Temas

- [Filtrado de contexto de usuario para documentos añadidos directamente a un índice \(p. 528\)](#)
- [Filtrado de contexto de usuario para preguntas frecuentes \(p. 528\)](#)
- [Filtrado de contexto de usuario para fuentes de datos \(p. 528\)](#)

Filtrado de contexto de usuario para documentos añadidos directamente a un índice

Al añadir documentos directamente a un índice mediante el [BatchPutDocument API](#), Amazon Kendra obtiene la información de usuarios y grupos de `AccessControlList` campo del documento. Usted proporciona una lista de control de acceso (ACL) para sus documentos y la ACL se incorpora a sus documentos.

Especifica la ACL en [Directora](#) objeto como parte del [Documento](#) objeto en el [BatchPutDocument API](#). Usted proporciona la siguiente información:

- El acceso que debe tener el usuario o el grupo. Puedes decir `ALLOW` o `DENY`.
- El tipo de entidad. Puedes decir `USER` o `GROUP`.
- Nombre del usuario o grupo.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para preguntas frecuentes

Cuando tú [añadir preguntas frecuentes](#) a un índice, Amazon Kendra obtiene la información de usuarios y grupos de `AccessControlList` objeto/campo del archivo JSON de preguntas frecuentes. También puedes usar un archivo CSV de preguntas frecuentes con campos o atributos personalizados para el control de acceso.

Usted proporciona la siguiente información:

- El acceso que debe tener el usuario o el grupo. Puedes decir `ALLOW` o `DENY`.
- El tipo de entidad. Puedes decir `USER` o `GROUP`.
- Nombre del usuario o grupo.

Para obtener más información, consulte [Archivos de preguntas frecuentes](#).

Filtrado de contexto de usuario para fuentes de datos

Amazon Kendra también rastrea la información de la lista de control de acceso (ACL) de usuarios y grupos desde los conectores de fuentes de datos compatibles. Esto resulta útil para filtrar el contexto del usuario, donde los resultados de la búsqueda se filtran en función del acceso del usuario o de su grupo a los documentos.

Temas

- [Filtrado de contexto de usuario para fuentes de datos de Adobe Experience Manager \(p. 529\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Alfresco \(p. 529\)](#)

- [Filtrado de contexto de usuario para Amazon S3fuentes de datos \(p. 530\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de bases de datos \(p. 530\)](#)
- [Filtrado de contexto de usuario para Amazon FSxfuentes de datos \(p. 530\)](#)
- [Filtrado de contexto de usuario para Amazon WorkDocsfuentes de datos \(p. 531\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Box \(p. 531\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Confluence \(p. 531\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Dropbox \(p. 532\)](#)
- [Filtrado de contexto de usuario para GitHubfuentes de datos \(p. 532\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Google Drive \(p. 532\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Gmail \(p. 533\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Jira \(p. 533\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Microsoft Exchange \(p. 534\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Microsoft Teams \(p. 534\)](#)
- [Filtrado de contexto de usuario para MicrosoftOneDrivefuentes de datos \(p. 534\)](#)
- [Filtrado de contexto de usuario para MicrosoftOneDrivefuentes de datos v2.0 \(p. 535\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Microsoft Yammer \(p. 535\)](#)
- [Filtrado de contexto de usuario para MicrosoftSharePointfuentes de datos \(p. 536\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Quip \(p. 536\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Salesforce \(p. 537\)](#)
- [Filtrado de contexto de usuario para ServiceNowfuentes de datos \(p. 537\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Slack \(p. 538\)](#)
- [Filtrado de contexto de usuario para fuentes de datos de Zendesk \(p. 538\)](#)

Filtrado de contexto de usuario para fuentes de datos de Adobe Experience Manager

Al utilizar una fuente de datos de Adobe Experience Manager,Amazon Kendraobtiene la información del usuario y del grupo de la instancia de Adobe Experience Manager.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`—Los ID de grupo existen en el contenido de Adobe Experience Manager donde hay permisos de acceso establecidos. Se mapean a partir de los nombres de los grupos de Adobe Experience Manager.
- `_user_id`—Los identificadores de usuario existen en el contenido de Adobe Experience Manager donde hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores en Adobe Experience Manager.

Puede añadir hasta 200 entradas en el `AccessControlList`campo.

Filtrado de contexto de usuario para fuentes de datos de Alfresco

Cuando utiliza una fuente de datos de Alfresco,Amazon Kendraobtiene la información del usuario y del grupo de la instancia de Alfresco.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`—Los identificadores de grupo existen en Alfresco en los archivos en los que hay permisos de acceso establecidos. Se mapean a partir de los nombres del sistema de los grupos (no de los nombres para mostrar) en Alfresco.

- `_user_id`—Los identificadores de usuario existen en Alfresco en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores en Alfresco.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para Amazon S3 fuentes de datos

Añades el filtrado de contexto de usuario a un documento en un Amazon S3 fuente de datos mediante un archivo de metadatos asociado al documento. Añades la información al `AccessControlList` campo del documento JSON. Para obtener más información sobre cómo añadir metadatos a los documentos indexados desde un Amazon S3 fuente de datos, consulta [Metadatos de documentos S3](#).

Usted proporciona tres datos:

- El acceso que debe tener la entidad. Puedes decir `ALLOW` o `DENY`.
- El tipo de entidad. Puedes decir `USER` o `GROUP`.
- El nombre de la entidad.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para fuentes de datos de bases de datos

Cuando utiliza una fuente de datos de base de datos, como Amazon Aurora PostgreSQL, Amazon Kendra obtiene información de usuarios y grupos de una columna de la tabla fuente. Especifica esta columna en `AclConfiguration` objeto como parte del `DatabaseConfiguration` objeto en el [CreateDataSource API](#).

Una fuente de datos de base de datos tiene las siguientes limitaciones:

- Solo puede especificar una lista de permitidos para una fuente de datos de base de datos. No puedes especificar una lista de denegaciones.
- Solo puede especificar grupos. No puede especificar usuarios individuales para la lista de permitidos.
- La columna de la base de datos debe ser una cadena que contenga una lista de grupos delimitada por punto y coma.

Filtrado de contexto de usuario para Amazon FSx fuentes de datos

Cuando usas un Amazon FSx fuente de datos, Amazon Kendra obtiene información de usuarios y grupos del servicio de directorios del Amazon FSx instancia.

El Amazon FSx Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`—Los identificadores de grupo existen en Amazon FSx en los archivos en los que haya permisos de acceso establecidos. Se mapean a partir de los nombres de los grupos de sistemas en el servicio de directorios de Amazon FSx.
- `_user_id`—Los identificadores de usuario existen en Amazon FSx en los archivos en los que haya permisos de acceso establecidos. Se mapean a partir de los nombres de usuario del sistema en el servicio de directorio de Amazon FSx.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para Amazon WorkDocs fuentes de datos

Cuando usas una fuente de datos de Amazon WorkDocs, Amazon Kendra obtiene la información de usuarios y grupos de la instancia de Amazon WorkDocs.

En Amazon WorkDocs, los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`—Los identificadores de grupo existen en Amazon WorkDocs en los archivos en los que haya permisos de acceso establecidos. Están mapeados a partir de los nombres de los grupos en Amazon WorkDocs.
- `_user_id`—Los identificadores de usuario existen en Amazon WorkDocs en los archivos en los que haya permisos de acceso establecidos. Se mapean a partir de los nombres de usuario en Amazon WorkDocs.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para fuentes de datos de Box

Al utilizar una fuente de datos de Box, Amazon Kendra obtiene información de usuarios y grupos de la instancia de Box.

El grupo Box y los ID de usuario se asignan de la siguiente manera:

- `_group_ids`—Los identificadores de grupo aparecen en Box en los archivos en los que hay permisos de acceso establecidos. Están mapeados a partir de los nombres de los grupos de Box.
- `_user_id`—Los identificadores de usuario aparecen en Box en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores de usuario en Box.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para fuentes de datos de Confluence

Cuando usas una fuente de datos de Confluence, Amazon Kendra obtiene información de usuarios y grupos de la instancia de Confluence.

Puede configurar el acceso de usuarios y grupos a los espacios mediante la página de permisos de espacio. Para las páginas y los blogs, utiliza la página de restricciones. Para obtener más información sobre los permisos de espacio, consulte [Descripción general de los permisos de espacio](#) en el sitio web de soporte de Confluence. Para obtener más información sobre las restricciones de páginas y blogs, consulte [Restricciones de página](#) en el sitio web de soporte de Confluence.

El grupo de Confluence y los nombres de usuario se representan de la siguiente manera:

- `_group_ids`—Los nombres de los grupos aparecen en los espacios, las páginas y los blogs donde hay restricciones. Están mapeados a partir del nombre del grupo en Confluence. Los nombres de los grupos siempre aparecen en minúsculas.
- `_user_id`—Los nombres de usuario aparecen en el espacio, la página o el blog donde hay restricciones. Se mapean según el tipo de instancia de Confluence que utilices.

Para el conector Confluence v1.0

- Servidor:_user_ides el nombre de usuario. El nombre de usuario siempre está en minúsculas.
- Nube:_user_ides el ID de cuenta del usuario.

Para el conector Confluence v2.0

- Servidor:_user_ides el nombre de usuario. El nombre de usuario siempre está en minúsculas.
- Nube:_user_ides el ID de correo electrónico del usuario.

Puede añadir hasta 200 entradas en elAccessControlListcampo.

Filtrado de contexto de usuario para fuentes de datos de Dropbox

Cuando usas una fuente de datos de Dropbox, Amazon Kendra obtiene la información del usuario y del grupo de la instancia de Dropbox.

Los ID de grupo y usuario se asignan de la siguiente manera:

- _group_ids—Los ID de grupo existen en Dropbox en los archivos en los que hay permisos de acceso establecidos. Se mapean a partir de los nombres de los grupos de Dropbox.
- _user_id—Los identificadores de usuario existen en Dropbox en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores en Dropbox.

Puede añadir hasta 200 entradas en elAccessControlListcampo.

Filtrado de contexto de usuario para GitHubfuentes de datos

Cuando usas unGitHubfuente de datos, Amazon Kendra obtiene la información del usuario deGitHubinstancia.

ElGitHubLos identificadores de usuario se asignan de la siguiente manera:

- _user_id—Los identificadores de usuario existen enGitHuben los archivos en los que haya permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores enGitHub.

Puede añadir hasta 200 entradas en elAccessControlListcampo.

Filtrado de contexto de usuario para fuentes de datos de Google Drive

Una fuente de datos de Google Workspace Drive muestra información de usuarios y grupos de usuarios y grupos de Google Drive. La pertenencia a grupos y dominios está asignada a_group_idscampo de índice. El nombre de usuario de Google Drive está asignado a_user_idcampo.

Al proporcionar una o más direcciones de correo electrónico de usuario enQueryAPI, solo se devuelven los documentos que se han compartido con esas direcciones de correo electrónico. Los siguientesAttributeFilterel parámetro solo devuelve los documentos compartidos con "martha@example.com».

```
"AttributeFilter": {  
    "EqualsTo":{  
        "Key": "_user_id",  
        "Value": {
```

```
        "StringValue": "martha@example.com"
    }
}
```

Si proporciona una o más direcciones de correo electrónico de grupo en la consulta, solo se devolverán los documentos compartidos con los grupos. Los siguientes `AttributeFilter` el parámetro solo devuelve los documentos compartidos con el grupo «hr@example.com».

```
"AttributeFilter": {
    "EqualsTo":{
        "Key": "_group_ids",
        "Value": {
            "StringListValue": ["hr@example.com"]
        }
    }
}
```

Si proporciona el dominio en la consulta, se devolverán todos los documentos compartidos con el dominio. Los siguientes `AttributeFilter` El parámetro devuelve los documentos compartidos con el dominio «example.com».

```
"AttributeFilter": {
    "EqualsTo":{
        "Key": "_group_ids",
        "Value": {
            "StringListValue": ["example.com"]
        }
    }
}
```

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para fuentes de datos de Gmail

Cuando utilizas una fuente de datos de Gmail, Amazon Kendra obtiene la información del usuario de la instancia de Gmail.

Los ID de usuario se asignan de la siguiente manera:

- `_user_id`— Los identificadores de usuario existen en Gmail en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores en Gmail.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para fuentes de datos de Jira

Cuando usas una fuente de datos de Jira, Amazon Kendra obtiene información de usuarios y grupos de la instancia de Jira.

Los ID de usuario de Jira se asignan de la siguiente manera:

- `_user_id`—Los identificadores de usuario existen en Jira en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores de usuario en Jira.

Puede añadir hasta 200 entradas en elAccessControlListcampo.

Filtrado de contexto de usuario para fuentes de datos de Microsoft Exchange

Amazon Kendra recupera la información del usuario de Microsoft Exchange cuando indexa los documentos del sitio. La información del usuario se toma del sitio host de Microsoft Exchange subyacente.

Al utilizar un usuario de Exchange para filtrar el contexto de usuario, Amazon Kendra obtiene la información del usuario de la instancia de Microsoft Exchange. Los ID de Exchange se asignan de la siguiente manera:

- `_tenant_id`— Su ID de inquilino de Microsoft es un identificador único a nivel mundial que es necesario para configurar cada instancia de conector. Su ID de inquilino es diferente del nombre o dominio de su organización y se encuentra en la sección de propiedades del panel de control de su cuenta Microsoft.

Puede añadir hasta 200 entradas en elAccessControlListcampo.

Filtrado de contexto de usuario para fuentes de datos de Microsoft Teams

Amazon Kendra recupera la información del usuario de Microsoft Teams cuando indexa los documentos del sitio. La información del usuario se extrae del sitio host subyacente de Microsoft Teams.

Cuando utilizas un usuario de Teams para filtrar el contexto de usuario, Amazon Kendra obtiene la información del usuario de la instancia de Microsoft Teams. Los ID de inquilinos de Teams se asignan de la siguiente manera:

- `_tenant_id`— Su ID de inquilino de Microsoft es un identificador único a nivel mundial que es necesario para configurar cada instancia de conector. Su ID de inquilino es diferente del nombre o dominio de su organización y se encuentra en la sección de propiedades del panel de control de su cuenta Microsoft.

Puede añadir hasta 200 entradas en elAccessControlListcampo.

Filtrado de contexto de usuario para MicrosoftOneDrivefuentes de datos

Amazon Kendra recupera información de usuarios y grupos de MicrosoftOneDrive cuando indexa los documentos del sitio. La información del usuario y del grupo se toma del Microsoft subyacente SharePointsitio que alojaOneDrive.

Cuando usas unOneDriveusuario o grupo para filtrar el contexto de usuario, calcule el ID de la siguiente manera:

1. Obtenga el nombre del sitio. Por ejemplo, <https://host.onmicrosoft.com/sites/siteName>.
2. Tome el hash MD5 del nombre del sitio. Por ejemplo, `430a6b90503eeef95c89295c8999c7981`.
3. Cree el correo electrónico del usuario o el ID del grupo concatenando el hash MD5 con una barra vertical (|) y el ID. Por ejemplo, si el nombre de un grupo es «propietarios del sitio», el identificador del grupo sería:

`"430a6b90503eeef95c89295c8999c7981|site owners"`

Para el nombre de usuario `"someone@host.onmicrosoft.com"`, el ID de usuario sería el siguiente:

"430a6b90503eeef95c89295c8999c7981|someone@host.onmicrosoft.com"

Enviar el ID de usuario o grupo a Amazon Kendra como el_user_id o group_ids atributo cuando llamas al [ConsultaAPI](#). Por ejemplo, el AWS CLI el comando que usa un grupo para filtrar la respuesta de la consulta tiene este aspecto:

```
aws kendra query \  
    --index-id index ID \  
    --query-text "query text" \  
    --attribute-filter '{ \  
        "EqualsTo":{ \  
            "Key": "_group_ids", \  
            "Value": {"StringValue": "430a6b90503eeef95c89295c8999c7981|site owners"} }'
```

Puede añadir hasta 200 entradas en el AccessControlList campo.

Filtrado de contexto de usuario para Microsoft OneDrive fuentes de datos v2.0

Un Microsoft OneDrive La fuente de datos v2.0 devuelve información de secciones y páginas de OneDrive entidades de la lista de control de acceso (ACL). Amazon Kendra utiliza el OneDrive dominio arrendatario para conectarse al OneDrive instancia y puede filtrar según el nombre de la sección, el tipo de página, el nombre de archivo, el tipo de archivo y el contenido del archivo.

Para los objetos estándar, el_user_id y group_id se utilizan de la siguiente manera:

- _user_id— Su Microsoft OneDrive la ID de correo electrónico del usuario está asignada a user_id campo.
- _group_id— Su Microsoft OneDrive el correo electrónico del grupo está asignado al group_id campo.

Puede añadir hasta 200 entradas en el AccessControlList campo.

Filtrado de contexto de usuario para fuentes de datos de Microsoft Yammer

Amazon Kendra recupera la información de usuarios y grupos de Microsoft Yammer cuando indexa los documentos del sitio. La información del usuario y del grupo se obtiene del sitio host de Microsoft Yammer subyacente.

Cuando usas un usuario de Yammer para filtrar el contexto de usuario, Amazon Kendra obtiene la información del usuario de la instancia de Microsoft Yammer. Los ID de usuario de Microsoft Yammer se asignan de la siguiente manera:

- _email_id— Su ID de correo electrónico de Microsoft es un identificador necesario para configurar cada instancia de conector. Puedes encontrar tu ID de correo electrónico en la sección de propiedades del panel de control de tu cuenta Microsoft.
- _group_id— Los ID de grupo existen en las instancias de Microsoft Yammer donde hay permisos de acceso establecidos. Se mapean a partir de los nombres de los grupos en Microsoft Yammer.

Puede añadir hasta 200 entradas en el AccessControlList campo.

Filtrado de contexto de usuario para Microsoft SharePoint fuentes de datos

Amazon Kendra recupera información de usuarios y grupos de Microsoft SharePoint cuando indexa los documentos del sitio. Para filtrar sus documentos, proporcione la información del usuario y del grupo cuando llame al [Query API](#).

Para filtrar mediante un nombre de usuario, utilice la dirección de correo electrónico del usuario. Por ejemplo, johnstiles@example.com.

Cuando usas un SharePoint grupo para filtrar el contexto de usuario, calcule el ID del grupo de la siguiente manera:

Para grupos locales

1. Obtenga el nombre del sitio. Por ejemplo, <https://host.onmicrosoft.com/sites/siteName> ..
2. Tome el hash SHA256 del nombre del sitio. Por ejemplo, 430a6b90503eef95c89295c8999c7981.
3. Cree el ID del grupo concatenando el hash SHA256 con una barra vertical (|) y el nombre del grupo. Por ejemplo, si el nombre del grupo es «propietarios del sitio», el identificador del grupo sería:

```
"430a6b90503eef95c89295c8999c7981|site owners"
```

Enviar el ID del grupo a Amazon Kendra como el `_group_ids` atributo cuando llamas al [API de consulta](#). Por ejemplo, el AWS CLI el comando se ve así:

```
aws kendra query \  
    --index-id index ID  
    --query-text "query text"  
    --attribute-filter '{  
        "EqualsTo":{  
            "Key": "_group_ids",  
            "Value": {"StringValue": "430a6b90503eef95c89295c8999c7981|site  
owners"}  
        }}'
```

Para grupos de AD

1. Use el ID del grupo AD para configurar el contexto de usuario.

Enviar el ID del grupo a Amazon Kendra como el `_group_ids` atributo cuando llamas al [Consulta API](#). Por ejemplo, el AWS CLI el comando se ve así:

```
aws kendra query \  
    --index-id index ID  
    --query-text "query text"  
    --attribute-filter '{  
        "EqualsTo":{  
            "Key": "_group_ids",  
            "Value": {"StringValue": "AD group"}  
        }}'
```

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para fuentes de datos de Quip

Cuando utiliza una fuente de datos de Quip, Amazon Kendra obtiene la información del usuario de la instancia de Quip.

Los ID de usuario de Quip se mapean de la siguiente manera:

- `_user_id`—Los identificadores de usuario existen en Quip en los archivos en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores en Quip.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para fuentes de datos de Salesforce

Una fuente de datos de Salesforce devuelve información de usuarios y grupos de las entidades de la lista de control de acceso (ACL) de Salesforce. Puede aplicar el filtrado de contexto de usuario a los objetos estándar de Salesforce y a las fuentes de chat. El filtrado del contexto de usuario no está disponible para los artículos de conocimiento de Salesforce.

Para los objetos estándar, el `_user_id` y `_group_ids` se utilizan de la siguiente manera:

- `_user_id`—El nombre de usuario del usuario de Salesforce.
- `_group_ids`
 - Nombre de `SalesforceProfile`
 - Nombre de `SalesforceGroup`
 - Nombre de `SalesforceUserRole`
 - Nombre de `SalesforcePermissionSet`

Para los feeds de chat, el `_user_id` y `_group_ids` se utilizan de la siguiente manera:

- `_user_id`—El nombre de usuario del usuario de Salesforce. Solo está disponible si el artículo está publicado en el feed del usuario.
- `_group_ids`—Los ID de grupo se utilizan de la siguiente manera. Solo está disponible si el elemento del feed se publica en un chat o en un grupo de colaboración.
 - El nombre del grupo de chat o colaboración.
 - Si el grupo es público, `PUBLIC:ALL`.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para ServiceNow fuentes de datos

Filtrado de contexto de usuario para ServiceNow solo es compatible con `TemplateConfigurationAPI` y `ServiceNowConector v2.0`. `ServiceNowConfigurationAPI` y `ServiceNowEl conector v1.0` no admite el filtrado de contexto de usuario.

Cuando usas un ServiceNow fuente de datos, Amazon Kendra obtiene la información del usuario y del grupo de ServiceNow instancia.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`—Los identificadores de grupo existen en ServiceNow en los archivos en los que haya permisos de acceso establecidos. Están mapeados a partir de los nombres de los roles `desys_ids` en ServiceNow.
- `_user_id`—Los identificadores de usuario existen en ServiceNow en los archivos en los que haya permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores en ServiceNow.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para fuentes de datos de Slack

Cuando utilizas una fuente de datos de Slack, Amazon Kendra obtiene la información del usuario de la instancia de Slack.

Los seudónimos de Slack se asignan de la siguiente manera:

- `_user_id`—Los seudónimos existen en Slack en los mensajes y canales en los que hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores en Slack.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Filtrado de contexto de usuario para fuentes de datos de Zendesk

Cuando usa una fuente de datos de Zendesk, Amazon Kendra obtiene la información del usuario y del grupo de la instancia de Zendesk.

Los ID de grupo y usuario se asignan de la siguiente manera:

- `_group_ids`—Los ID de grupo existen en los tickets y artículos de Zendesk donde hay permisos de acceso establecidos. Se mapean a partir de los nombres de los grupos en Zendesk.
- `_user_id`—Los ID de grupo existen en los tickets y artículos de Zendesk donde hay permisos de acceso establecidos. Se asignan a partir de los correos electrónicos de los usuarios como identificadores en Zendesk.

Puede añadir hasta 200 entradas en el `AccessControlList` campo.

Respuestas a consultas y tipos de respuestas

Amazon Kendra admite diferentes tipos de respuestas y respuestas a consultas.

Respuestas a consultas

Una llamada al [Consulta](#) La API devuelve información sobre los resultados de una búsqueda.

Los resultados se encuentran en una matriz de [QueryResultItem](#) objetos (`ResultItems`).

Cada [QueryResultItem](#) incluye un resumen del resultado. Se incluyen los atributos del documento asociados al resultado de la consulta.

Información resumida

La información resumida varía según el tipo de resultado. En cada caso, incluye el texto del documento que coincide con el término de búsqueda. También incluye información de resultado que puede utilizar para resaltar el texto de búsqueda en la salida de la aplicación. Por ejemplo, si el término de búsqueda es ¿Cuál es la altura de la Space Needle?, la información resumida incluye la ubicación del texto de las palabras `salt`, `tur`, `aguja` y `espacial`. Para obtener información sobre los tipos de respuesta, consulte [Respuestas a consultas y tipos de respuestas \(p. 538\)](#).

Atributos del documento

Cada resultado contiene los atributos del documento que coincide con una consulta. Algunos de los atributos están predefinidos, como `DocumentId`, `DocumentTitle`, y `DocumentUri`. Otros son atributos personalizados que usted define. Puede utilizar los atributos del documento para filtrar la respuesta de la `QueryAPI`. Por ejemplo, es posible que desee que solo los documentos estén escritos por un autor

específico o por una versión específica de un documento. Para obtener más información, consulte [Filtrado y búsqueda de facetas \(p. 520\)](#). Los atributos del documento se especifican al añadir documentos a un índice. Para obtener más información, consulte [Campos o atributos personalizados](#).

El siguiente es un ejemplo de código JSON para el resultado de una consulta. Tenga en cuenta los atributos del documento en `DocumentAttributes` y `AdditionalAttributes`.

```
{  
    "QueryId": "query-id",  
    "ResultItems": [  
        {  
            "Id": "result-id",  
            "Type": "ANSWER",  
            "AdditionalAttributes": [  
                {  
                    "Key": "AnswerText",  
                    "ValueType": "TEXT_WITH_HIGHLIGHTS_VALUE",  
                    "Value": {  
                        "TextWithHighlightsValue": {  
                            "Text": "text",  
                            "Highlights": [  
                                {  
                                    "BeginOffset": 55,  
                                    "EndOffset": 90,  
                                    "TopAnswer": false  
                                }  
                            ]  
                        }  
                    }  
                }  
            ],  
            "DocumentId": "document-id",  
            "DocumentTitle": {  
                "Text": "title"  
            },  
            "DocumentExcerpt": {  
                "Text": "text",  
                "Highlights": [  
                    {  
                        "BeginOffset": 0,  
                        "EndOffset": 300,  
                        "TopAnswer": false  
                    }  
                ]  
            },  
            "DocumentURI": "uri",  
            "DocumentAttributes": [],  
            "ScoreAttributes": "score",  
            "FeedbackToken": "token"  
        },  
        {  
            "Id": "result-id",  
            "Type": "ANSWER",  
            "Format": "TABLE",  
            "DocumentId": "document-id",  
            "DocumentTitle": {  
                "Text": "title"  
            },  
            "TableExcerpt": {  
                "Rows": [  
                    "Cells": [{  
                        "Header": true,  
                        "Highlighted": false,  
                        "TopAnswer": false,  
                        "Value": "value"  
                    }  
                ]  
            }  
        }  
    ]  
}
```

```
        }, {
            "Header": true,
            "Highlighted": false,
            "TopAnswer": false,
            "Value": "value"
        }, {
            "Header": true,
            "Highlighted": false,
            "TopAnswer": false,
            "Value": "value"
        }, {
            "Header": true,
            "Highlighted": false,
            "TopAnswer": false,
            "Value": "value"
        }
    ],
    "Cells": [
        {
            "Header": false,
            "Highlighted": false,
            "TopAnswer": false,
            "Value": "value"
        },
        {
            "Header": false,
            "Highlighted": false,
            "TopAnswer": false,
            "Value": "value"
        },
        {
            "Header": false,
            "Highlighted": true,
            "TopAnswer": true,
            "Value": "value"
        },
        {
            "Header": false,
            "Highlighted": false,
            "TopAnswer": false,
            "Value": "value"
        }
    ],
    "TotalNumberofRows": number
},
"DocumentURI": "uri",
"ScoreAttributes": "score",
"FeedbackToken": "token"
},
{
    "Id": "result-id",
    "Type": "DOCUMENT",
    "AdditionalAttributes": [],
    "DocumentId": "document-id",
    "DocumentTitle": {
        "Text": "title",
        "Highlights": []
    },
    "DocumentExcerpt": {
        "Text": "text",
        "Highlights": [
            {
                "BeginOffset": 74,
                "EndOffset": 77,
                "TopAnswer": false
            }
        ]
    },
    "DocumentURI": "uri",
    "DocumentAttributes": [
```

```
{  
    "Key": "_source_uri",  
    "Value": {  
        "StringValue": "uri"  
    }  
},  
],  
"ScoreAttributes": "score",  
"FeedbackToken": "token",  
}  
],  
"FacetResults": [],  
"TotalNumberOfResults": number  
}
```

Tipos de respuesta

Amazon Kendra devuelve tres tipos de respuesta a la consulta.

- Respuesta (incluye la respuesta de la tabla)
- Document
- Pregunta y respuesta

El tipo de respuesta se devuelve en `type` campo de respuesta del [QueryResultItem](#) objeto.

Respuesta

Amazon Kendra detectó una o más respuestas a las preguntas en la respuesta. Un dato es la respuesta a una pregunta sobre quién, qué, cuándo o dónde, como ¿Dónde está el centro de servicio más cercano a mí? Amazon Kendra devuelve el texto del índice que mejor coincide con la consulta. El texto está en `AnswerText` campo y contiene información destacada para el término de búsqueda dentro del texto de la respuesta. `AnswerText` incluye el extracto completo del documento con el texto resaltado, mientras que `DocumentExcerpt` incluye el extracto del documento truncado (290 caracteres) con el texto resaltado.

Amazon Kendra solo devuelve una respuesta por documento, y esa es la respuesta con la mayor confianza. Para devolver varias respuestas de un documento, debe dividir el documento en varios documentos.

```
{  
    'AnswerText': {  
        'TextWithHighlights': [  
            {  
                'BeginOffset': 271,  
                'EndOffset': 279,  
                'TopAnswer': False  
            },  
            {  
                'BeginOffset': 481,  
                'EndOffset': 489,  
                'TopAnswer': False  
            },  
            {  
                'BeginOffset': 547,  
                'EndOffset': 555,  
                'TopAnswer': False  
            },  
            {  
                'BeginOffset': 764,
```

```
        'EndOffset': 772,
        'TopAnswer': False
    },
],
'Text': 'Asynchronous operations can\n' 'also process
\\n''documents that are in PDF''format. Using PDF format files allows you to process ''multi-
page\\n''documents.\\n''For information about how Amazon Textract represents
\\n''documents as Block objects,
    'seeDocumentsAndBlockObjects.\\n''\\n''\\n''For information about document ''limits,
    seeLimits in Amazon Textract.
\\n''\\n''\\n''\\n''The Amazon Textract synchronous ''operations can process documents stored in an Amazon
\\n''S3 Bucket or you can pass ''base64 encoded image bytes.\\n''For more information,
see ''Calling Amazon Textract Synchronous Operations. ''Asynchronous operations require input documents
\\n''to be supplied in an Amazon ''S3 Bucket.'
},
'DocumentExcerpt': {
    'Highlights': [
        {
            'BeginOffset': 0,
            'EndOffset': 300,
            'TopAnswer': False
        }
    ],
    'Text': 'Asynchronous operations can\n' 'also process
\\n''documents that are in PDF''format. Using PDF format files allows you to process ''multi-page
\\n''documents.\\n''For information about how Amazon ''Text extract represents\\n'''
},
'Type': 'ANSWER'
}
```

Document

Amazon Kendra devuelve los documentos clasificados para aquellos que coinciden con el término de búsqueda. La clasificación se basa en la confianza de que Amazon Kendra tiene en la precisión del resultado de la búsqueda. La información sobre el documento correspondiente se devuelve en el [QueryResultItem](#). Incluye el título del documento. El extracto incluye información destacada para el texto de búsqueda y la sección de texto coincidente del documento. El URI para los documentos coincidentes se encuentra en `SourceURI` atributo de documento. El siguiente ejemplo de JSON muestra el resumen del documento de un documento coincidente.

```
{
    'DocumentTitle': {
        'Highlights': [
            {
                'BeginOffset': 7,
                'EndOffset': 15,
                'TopAnswer': False
            },
            {
                'BeginOffset': 97,
                'EndOffset': 105,
                'TopAnswer': False
            }
        ],
        'Text': 'Amazon Textract API Permissions: Actions,
\\n''Permissions,
and Resources Reference - ''Amazon Textract'
    },
    'DocumentExcerpt': {
        'Highlights': [
            {
                'BeginOffset': 68,
```

```
        'EndOffset': 76,
        'TopAnswer': False
    },
    {
        'BeginOffset': 121,
        'EndOffset': 129,
        'TopAnswer': False
    }
],
'Text': '....LoggingandMonitoring\tMonitoring
\n''\tCloudWatchMetricsforAmazonTextract
\n''\tLoggingAmazonTextractAPICallswithAWSCloudTrail\n'''tAPIReference\tActions
\tAnalyzeDocument\n'''tDetectDocumentText\n'''tGetDocumentAnalysis...'
},
'Type': 'DOCUMENT'
}
```

Pregunta y respuesta

Se devuelve una respuesta de pregunta y respuesta cuando Amazon Kendra relaciona una pregunta con una de las preguntas más frecuentes del índice. La respuesta incluye la pregunta y la respuesta coincidentes en [QueryResultItem](#) campo. También incluye información de resaltado de los términos de consulta detectados en la cadena de consulta. El siguiente JSON muestra una respuesta a una pregunta y una respuesta. Tenga en cuenta que la respuesta incluye el texto de la pregunta.

```
{
    'AnswerText': {
        'TextWithHighlights': [
            ],
        'Text': '605feet'
    },
    'DocumentExcerpt': {
        'Highlights': [
            {
                'BeginOffset': 0,
                'EndOffset': 8,
                'TopAnswer': False
            }
        ],
        'Text': '605feet'
    },
    'Type': 'QUESTION_ANSWER',
    'QuestionText': {
        'Highlights': [
            {
                'BeginOffset': 12,
                'EndOffset': 18,
                'TopAnswer': False
            },
            {
                'BeginOffset': 26,
                'EndOffset': 31,
                'TopAnswer': False
            },
            {
                'BeginOffset': 32,
                'EndOffset': 38,
                'TopAnswer': False
            }
        ],
        'Text': 'whatistheheightoftheSpaceNeedle?'
    }
}
```

}

Para obtener información sobre cómo añadir texto de preguntas y respuestas a un índice, consulte [Creación de preguntas frecuentes](#).

Ajustar y ordenar las respuestas

Puede modificar el efecto de un campo o atributo en la relevancia de la búsqueda mediante el ajuste de relevancia. También puede ordenar los resultados de la búsqueda por un atributo o campo determinado.

Ajustar las respuestas

Puede modificar el efecto de un campo o atributo en la relevancia de la búsqueda mediante el ajuste de relevancia. Para probar rápidamente el ajuste de relevancia, utilice [Consulta API](#) para introducir las configuraciones de ajuste en la consulta. A continuación, puede ver los diferentes resultados de búsqueda que obtiene de las diferentes configuraciones. La consola no admite el ajuste de relevancia a nivel de consulta. También puede ajustar los campos o atributos que sean del tipo `StringList` solo a nivel de índice. Para obtener más información, consulte [Ajustar la relevancia de la búsqueda](#).

De forma predeterminada, las respuestas a las consultas se ordenan según la puntuación de relevancia que Amazon Kendra determina para cada resultado de la respuesta.

Puede ajustar los resultados para cualquier atributo o campo integrado o personalizado de los siguientes tipos:

- Valor de fecha
- Valor largo
- Valor de cadena

No puede ordenar los atributos del siguiente tipo:

- Valores de la lista de cadenas

Clasifique y ajuste los resultados de los documentos (AWSSDK)

Configure el `Searchable` parámetro to `true` para mejorar la configuración de los metadatos del documento.

Para ajustar un atributo en una consulta, defina `DocumentRelevanceOverrideConfigurations` parámetro del `Query` Utilice la API y especifique el nombre del atributo que deseé ajustar.

El siguiente ejemplo de JSON muestra un `DocumentRelevanceOverrideConfigurations` objeto que anula el ajuste del atributo denominado «departamento» en el índice.

```
"DocumentRelevanceOverrideConfigurations" : [  
    {"Name": "department",  
     "Relevance": {  
         "Importance": 1,  
         "ValueImportanceMap": {  
             "IT": 3,  
             "HR": 7  
         }  
     }  
}
```

]

Clasificación de respuestas

Amazon Kendra utiliza el atributo o campo de clasificación como parte de los criterios de los documentos devueltos por la consulta. Por ejemplo, es posible que los resultados devueltos por una consulta ordenada por «`_created_at`» no contengan los mismos resultados que una consulta ordenada por «`_version`».

De forma predeterminada, las respuestas a las consultas se ordenan según la puntuación de relevancia que Amazon Kendra determina para cada resultado de la respuesta. Para cambiar el orden de clasificación, haga que un atributo de documento sea ordenable y, a continuación, configure Amazon Kendra para usar ese atributo para ordenar las respuestas.

Puede ordenar los resultados en cualquier atributo/campo integrado o personalizado de los siguientes tipos:

- Valor de fecha
- Valor largo
- Valor de cadena

No puede ordenar los atributos del siguiente tipo:

- Valores de la lista de cadenas

Solo puede ordenar por un atributo o campo del documento en cada consulta. Las consultas devuelven 100 resultados. Si hay menos de 100 documentos con el atributo de clasificación establecido, los documentos sin un valor para el atributo de clasificación se devuelven al final de los resultados, ordenados por relevancia para la consulta.

Para ordenar los resultados del documento (AWSSDK)

1. Para utilizar el [UpdateIndex API](#) para hacer que un atributo se pueda ordenar, defina `Sortable` parámetro para `true`. El siguiente ejemplo de JSON usa `DocumentMetadataConfigurationUpdates` para añadir un atributo llamado «Departamento» al índice y hacerlo ordenable.

```
"DocumentMetadataConfigurationUpdates": [  
    {  
        "Name": "Department",  
        "Type": "STRING_VALUE",  
        "Search": {  
            "Sortable": "true"  
        }  
    }  
]
```

2. Para usar un atributo ordenable en una consulta, defina la `SortingConfiguration` parámetro del [Consulta API](#). Especifique el nombre del atributo que se va a ordenar y si desea ordenar la respuesta en orden ascendente o descendente.

El siguiente ejemplo de JSON muestra el `SortingConfiguration` parámetro que se utiliza para ordenar los resultados de una consulta por el atributo «Departamento» en orden ascendente.

```
"SortingConfiguration": {  
    "DocumentAttributeKey": "Department",  
    "SortOrder": "ASC"
```

}

Para ordenar los resultados del documento (consola)

1. Para que un atributo se pueda ordenar en la consola, elija `Clasifiable` en la definición del atributo. Puede hacer que un atributo se pueda ordenar al crearlo o modificarlo más adelante.
2. Para ordenar la respuesta de una consulta en la consola, elija el atributo para ordenar la respuesta en el menú `Clasificar`. En la lista solo aparecen los atributos que se marcaron como ordenables durante la configuración de la fuente de datos.

Ajustar la relevancia de la búsqueda

Amazon Kendra las consultas producen resultados de búsqueda clasificados según su relevancia. Todos los campos o atributos que se pueden buscar en el índice contribuyen a esta clasificación.

Puede modificar el efecto de un campo o atributo en la relevancia de la búsqueda mediante el ajuste de relevancia. El ajuste de la relevancia de la búsqueda se puede realizar manualmente a nivel de índice, donde se establecen las configuraciones de ajuste para el índice, o a nivel de consulta anulando las configuraciones establecidas en el nivel de índice.

Cuando se utiliza el ajuste de relevancia, el resultado mejora la respuesta cuando la consulta incluye términos que coinciden con el campo o el atributo. También especifica la cantidad de impulso que recibe el documento cuando hay una coincidencia. El ajuste de relevancia no hace Amazon Kendra que se incluya un documento en la respuesta a la consulta, sino que es solo uno de los factores que se Amazon Kendra utilizan para determinar la relevancia de un documento.

Puedes aumentar los campos o atributos específicos del índice para asignar más importancia a respuestas específicas. Por ejemplo, cuando alguien busca «¿Cuándo es Re:Invent?» podrías aumentar la relevancia de la actualización del documento en el campo «`_last_update_at`». O bien, en un índice de informes de investigación, puedes destacar una fuente de datos específica en el campo «`fuente`».

También puedes aumentar los documentos en función de los votos o el recuento de visitas, algo habitual en los foros y otras bases de conocimiento de soporte. Puede combinar las ampliaciones, por ejemplo, para aumentar los documentos que se ven más y también los más recientes.

Para establecer la cantidad de impulso que recibe un documento, utilice el `Importance` parámetro. Cuanto más alto sea `Importance`, más aumentará la relevancia del documento el campo o el atributo. Al ajustar el índice o ajustar el nivel de consulta, aumente el valor del `Importance` parámetro en pequeños incrementos hasta obtener el efecto deseado. Para determinar si está mejorando los resultados de la búsqueda, realice la búsqueda y compare los resultados con las consultas anteriores.

Puede especificar atributos de fecha, número o cadena para ajustar un índice o un ajuste a nivel de consulta. Puede ajustar los campos o atributos que son de este tipo `StringList` solo a nivel de índice. Cada campo o atributo tiene criterios específicos para mejorar un resultado.

- Campos o atributos de fecha: existen tres criterios específicos para los campos de fecha `Duration`, `Freshness` y `RankOrder`.
 - `Duration` establece el período de tiempo al que se aplica el impulso. Por ejemplo, si establece el período de tiempo en 86400 segundos (es decir, un día), el impulso comenzará a disminuir después de un día. Cuanto mayor sea la importancia, más rápido disminuirá el efecto de refuerzo.
 - `Freshness` determina qué tan reciente es un documento cuando se aplica a un campo o atributo. Si realiza una solicitud en el campo correspondiente `Freshness` a la fecha de creación o la fecha de última actualización, un documento creado más recientemente o por última vez se considera «más reciente» que un documento anterior. Por ejemplo, si el documento 1 se creó el 14 de noviembre y el documento 2 el 5 de noviembre, el documento 1 es «más reciente» que el documento 2. Y si el documento 1 se actualizó por última vez el 14 de noviembre y el documento 2 se actualizó por última vez el 20 de noviembre, el documento 2 es «más reciente» que el documento 1. Cuanto más fresco esté el documento, más se aplicará este impulso. Solo puede tener un `Freshness` campo en el índice.
 - `RankOrder` aplica el impulso en orden ascendente o descendente. Si lo especifica `ASCENDING`, las fechas posteriores tienen prioridad. Si lo especifica `DESCENDING`, las fechas anteriores tienen prioridad.
- Campos numéricos o atributos: para los campos numéricos o atributos, puede especificar el orden de clasificación que se Amazon Kendra debe utilizar para determinar la relevancia del campo o atributo. Si

Si lo especifica `ASCENDING`, se dará prioridad a los números más altos. Si lo especifica `DESCENDING`, los números más bajos tienen prioridad.

- Campos o atributos de cadena: para los campos o atributos de cadena, puede crear categorías de un campo para dar a cada categoría un impulso diferente. Por ejemplo, si aumentas un campo o atributo denominado «Departamento», puedes dar un impulso diferente a los documentos de «Recursos Humanos» que a los documentos de «Legal». Puede potenciar un campo o atributo de este tipo `String`. Puede aumentar `StringList` los campos solo a nivel de índice.

Ajuste de relevancia a nivel de índice

Para ajustar la relevancia de un campo o atributo a nivel de índice, utilice la [consola](#) para configurar los detalles del índice o la [UpdateIndex API](#).

En el ejemplo siguiente se establece el campo «`_last_updated_at`» como el campo de un documento. `Freshness`

```
"DocumentMetadataConfigurationUpdates" : [
    {
        "Name": "_last_updated_at",
        "Type": "DATE_VALUE",
        "Relevance": {
            "Freshness": TRUE,
            "Importance": 2
        }
    }
]
```

El siguiente ejemplo aplica una importancia diferente a las distintas categorías del campo «departamento».

```
"DocumentMetadataConfigurationUpdates" : [
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 3,
                "Legal": 1
            }
        }
    }
]
```

Ajuste de relevancia a nivel de consulta

Puede ajustar la relevancia de un campo o atributo en el nivel de consulta mediante la API de [consultas](#).

La consola no admite el ajuste de relevancia a nivel de consulta.

El ajuste a nivel de consulta puede acelerar el proceso de prueba del ajuste de relevancia, ya que no es necesario actualizar manualmente las configuraciones de ajuste del índice para cada prueba. Puede ajustar la relevancia de un documento introduciendo las configuraciones de ajuste en la consulta. A continuación, puede ver los diferentes resultados que obtiene de las diferentes configuraciones. Una configuración que se pasa a la consulta anula la configuración que se establece en el nivel de índice.

El siguiente ejemplo anula la importancia que se aplica al campo «departamento» y a cada categoría de departamento establecida en el nivel de índice, como se muestra en el ejemplo anterior. Cuando un usuario introduce su consulta de búsqueda, el campo «departamento» tiene un nivel de importancia razonable y el departamento legal tiene más importancia que el departamento de recursos humanos.

```
"DocumentRelevanceOverrideConfigurations" : [
    {
        "Name": "department",
        "Type": "STRING_VALUE",
        "Relevance": {
            "Importance": 2,
            "ValueImportanceMap": {
                "HR": 2,
                "Legal": 8
            }
        }
    }
]
```

Obtener información con el análisis de búsqueda

Puedes usar Amazon Kendra Search Analytics para obtener información sobre cómo tu aplicación de búsqueda ayuda a los usuarios a encontrar información de forma correcta o no.

Amazon Kendra Los análisis proporcionan una visión general de la forma en que los usuarios interactúan con la aplicación de búsqueda y de la eficacia de la configuración de la aplicación de búsqueda. Puede ver los datos de las métricas mediante la [GetSnapshots API](#) o seleccionando Analytics en el panel de navegación de la consola.

Puede renderizar los datos generados GetSnapshots por su propio panel personalizado. También puede utilizar el panel de métricas que se proporciona en la consola, que incluye gráficos visuales. Con un panel visual, puede buscar tendencias o patrones en el comportamiento de los usuarios a lo largo del tiempo o detectar problemas en la configuración de la aplicación de búsqueda. Por ejemplo, un gráfico lineal que muestre un número constante de consultas por día y un aumento constante podría indicar un aumento de la adopción y el uso. Por otro lado, una caída abrupta podría indicar que hay un problema que debe investigarse.

Puede utilizar las métricas para establecer conexiones entre diferentes puntos de datos a fin de resolver problemas relacionados con la forma en que los usuarios solicitan información o descubren oportunidades de negocio. Por ejemplo, el documento «¿Cómo funciona la IA?» es el documento en el que se hace más clic en los resultados de búsqueda y la consulta más buscada es «¿Cómo funciona el aprendizaje automático?». Esto le informa sobre los términos preferidos y el idioma que utilizan sus usuarios. Puede integrar estos términos en sus documentos o utilizar sinónimos personalizados para estos términos para que los usuarios puedan buscar más en sus documentos.

Métricas de búsqueda

Hay 10 métricas para analizar el rendimiento de la aplicación de búsqueda o la información que buscan los usuarios. Para recuperar los datos de métricas, especifique el nombre de cadena de los datos métricos que desea recuperar cuando llame `GetSnapshots`.

También debe proporcionar un intervalo de tiempo o una ventana de tiempo para ver los datos de las métricas. Puede ver los datos en las siguientes ventanas de tiempo:

- `THIS_WEEK`: la semana actual, que comienza el domingo y termina el día anterior a la fecha actual.
- `ONE_WEEK_AGO`: La semana anterior, comenzando el domingo y finalizando el sábado siguiente.
- `TWO_WEEKS_AGO`: La semana anterior a la semana anterior, comienza el domingo y termina el sábado siguiente.
- `THIS_MONTH`: el mes actual, que comienza el primer día del mes y termina el día anterior a la fecha actual.
- `ONE_MONTH_AGO`: el mes anterior, que comienza el primer día del mes y termina el último día del mes.
- `TWO_MONTHS_AGO`: el mes anterior al mes anterior, que comienza el primer día del mes y termina el último día del mes.

En la consola, las ventanas horarias admitidas son Esta semana, La semana anterior, Este mes, el mes anterior.

Porcentaje de clics

La proporción de consultas que conducen a hacer clic en un documento en los resultados de la búsqueda. Esto le ayuda a entender si la configuración de la aplicación de búsqueda ayuda a los usuarios a encontrar información relevante para sus consultas. En el caso de las consultas que devuelven respuestas instantáneas, es posible que los usuarios no tengan que hacer clic en un documento para obtener más información. Para obtener más información, consulte [the section called “Tasa de respuesta instantánea” \(p. 551\)](#). Debe llamar [SubmitFeedback](#) para asegurarse de que se recopilan los comentarios sobre los clics.

Para recuperar datos sobre la tasa de clics mediante la GetSnapshots API, especifique el metricType as AGG_QUERY_DOC_METRICS También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación.

Tasa de clics cero

La proporción de consultas que conducen a cero clics en los resultados de búsqueda. Esto te ayuda a entender las lagunas en tu contenido, lo que proporciona resultados de búsqueda irrelevantes. En el caso de las consultas que devuelven respuestas instantáneas, es posible que los usuarios no tengan que hacer clic en un documento para obtener más información. Para obtener más información, consulte [the section called “Tasa de respuesta instantánea” \(p. 551\)](#). Además, la configuración de búsqueda, como el ajuste de las configuraciones, podría influir en la forma en que se muestran los documentos en los resultados de la búsqueda.

Para recuperar datos con una tasa de clics cero mediante la GetSnapshots API, especifique el metricType asAGG_QUERY_DOC_METRICS. También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación.

Tasa cero de resultados de búsqueda

La proporción de consultas que conducen a cero resultados de búsqueda. Esto te ayuda a entender las lagunas en tu contenido que no ofrecen resultados de búsqueda relevantes.

Para recuperar datos con una tasa cero de resultados de búsqueda mediante la GetSnapshots API, especifique el metricType asAGG_QUERY_DOC_METRICS. También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación.

Tasa de respuesta instantánea

La proporción de consultas con una respuesta instantánea o preguntas frecuentes devueltas. Esto le ayuda a comprender el papel de las respuestas instantáneas a la hora de proporcionar información.

Para recuperar datos sobre la tasa de respuesta instantánea mediante la GetSnapshots API, especifique el metricType asAGG_QUERY_DOC_METRICS. También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación.

Consultas principales

Las 100 consultas más buscadas por tus usuarios. Esto le ayuda a entender qué consultas son populares y el tipo de información que más interesa a sus usuarios.

Las métricas incluyen el número de veces que se busca en la consulta, la proporción de clics en un documento, la proporción de clics sin clics en un documento, la profundidad media de clics en los

resultados de la búsqueda de la consulta, la proporción de respuestas instantáneas de la consulta y la confianza promedio de los 10 primeros resultados de búsqueda de una consulta.

Para recuperar los datos de las consultas principales mediante la GetSnapshots API, especifique el metricType asQUERIES_BY_COUNT. También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación de la consola y, a continuación, seleccionando Consultas principales en las listas de consultas.

Consultas principales con cero clics

Las 100 consultas principales que conducen a cero clics en los resultados de búsqueda. Esto te ayuda a entender si hay lagunas en tu contenido, si faltan documentos relevantes para algunas consultas o si la configuración de la aplicación de búsqueda arroja resultados de búsqueda irrelevantes. En el caso de las consultas que devuelven respuestas instantáneas, es posible que los usuarios no tengan que hacer clic en un documento para obtener más información. Para obtener más información, consulte [the section called “Tasa de respuesta instantánea” \(p. 551\)](#).

Las métricas incluyen el número de veces que la consulta genera cero clics, la proporción de cero clics en la consulta, la proporción de respuestas instantáneas de la consulta y la confianza promedio de los 10 primeros resultados de búsqueda de una consulta.

Para recuperar los datos de las consultas principales con cero clics mediante la GetSnapshots API, especifique el metricType asQUERIES_BY_ZERO_CLICK_RATE. También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación de la consola y, a continuación, seleccionando las consultas más frecuentes en las listas de consultas.

Consultas principales con cero resultados de búsqueda

Las 100 consultas principales que conducen a cero resultados de búsqueda. Esto te ayuda a comprender las lagunas en su contenido, ya que no hay documentos relevantes para algunas consultas. O bien, los usuarios pueden realizar consultas con términos especializados que posiblemente no generen resultados de búsqueda, lo que le pedirá que cree [sinónimos personalizados](#) para solucionar este problema.

Las métricas incluyen el número de veces que la consulta lleva a cero resultados de búsqueda, la proporción de cero resultados de búsqueda para la consulta y la proporción de veces que se busca la consulta en comparación con todas las consultas.

Para recuperar los datos de las consultas principales con cero resultados de búsqueda mediante la GetSnapshots API, especifique el metricType asQUERIES_BY_ZERO_RESULT_RATE. También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación de la consola y, a continuación, seleccionando Consultas con cero resultados en las listas de consultas.

Documentos en los que más se ha hecho clic

Los 100 documentos con más clics en los resultados de búsqueda. Esto te ayuda a entender qué documentos o resultados de búsqueda son más relevantes para sus usuarios cuando solicitan información.

Las métricas incluyen el número de veces que se hace clic en el documento, el número de «me gusta» que recibe un documento de los usuarios (pulgares hacia arriba) y el número de no me gusta que recibe un documento de los usuarios (pulgares hacia abajo).

Para recuperar los datos de los documentos en los que se hizo clic en la parte superior de los documentos mediante la GetSnapshots API, especifique el metricType asDOCS_BY_CLICK_COUNT. También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación de la consola y, a continuación, seleccionando los documentos más consultados en Listas de consultas.

Total de consultas

El número total de consultas buscadas por los usuarios. Esto le ayuda a comprender qué tan comprometidos están sus usuarios con su aplicación de búsqueda.

Para recuperar datos sobre el total de consultas mediante la GetSnapshots API, especifique el metricType asAGG_QUERY_DOC_METRICS. También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación.

Documentos totales

El número total de documentos del índice. Esto le ayuda a comparar el tamaño del índice con el número total de consultas para comprobar si hay un número adecuado de documentos para el volumen de consultas.

Para recuperar datos sobre el total de documentos mediante la GetSnapshots API, especifique el metricType asAGG_QUERY_DOC_METRICS. También puedes ver esta métrica en la consola seleccionando Analytics en el panel de navegación.

Ejemplo de recuperación de datos métricos

El código siguiente es un ejemplo de cómo recuperar datos de las consultas principales del mes anterior.

Console

Para recuperar las consultas principales del mes anterior

1. En el panel de navegación de la izquierda, en Índices, selecciona tu índice y, a continuación, selecciona Analytics.
2. En la página de análisis, selecciona el botón Esta semana para cambiar el intervalo de tiempo para recuperar los datos al mes anterior.
3. En la página de análisis, en Listas de consultas, selecciona Consultas principales.

CLI

Para recuperar las consultas principales del mes anterior

```
aws kendra get-snapshots \
--index-id index-id \
--interval "ONE_MONTH_AGO" \
--metric-type "QUERIES_BY_COUNT"
```

Python

Para recuperar las consultas principales del mes anterior

```
import boto3

kendra = boto3.client("kendra")

index_id = "index-id"
interval = "ONE_MONTH_AGO"
metric_type = "QUERIES_BY_COUNT"

snapshots_response = kendra.get_snapshots(
```

```
        IndexId = index_id,
        Interval = interval,
        MetricType = metric_type
    )

    print("Top queries data: " + snapshots_response["snapshotsData"])
```

Java

Para recuperar las consultas principales del mes anterior

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.GetSnapshotsRequest;
import software.amazon.awssdk.services.kendra.model.GetSnapshotsResponse;

public class TopQueriesExample {
    public static void main(String[] args) {
        KendraClient kendra = KendraClient.builder().build();

        String indexId = "indexID";
        String interval = "ONE_MONTH_AGO";
        String metricType = "QUERIES_BY_COUNT";

        GetSnapshotsRequest getSnapshotsRequest = GetSnapshotsRequest
            .builder()
            .indexId(indexId)
            .interval(interval)
            .metricType(metricType)
            .build();

        GetSnapshotsResponse getSnapshotsResponse =
            kendra.getSnapshots(GetSnapshotsRequest);

        System.out.println(String.format("Top queries data: ",
            getSnapshotsResponse.snapshotsData()))}
```

Desde métricas hasta información procesable

La información procesable es información significativa que se extrae de datos sin procesar y se utiliza para guiar sus acciones o decisiones. Para extraer el significado de las métricas y utilizarlas para obtener información procesable, es importante no solo analizar las métricas de forma aislada, sino también establecer conexiones entre las métricas.

Por ejemplo, la consulta principal con cero clics es «¿Qué regiones están disponibles actualmente?». Sin embargo, también tiene una tasa de respuesta instantánea del 100 por ciento. Esto sugiere que los usuarios reciban la respuesta a esta pregunta sin necesidad de hacer clic en un resultado de la búsqueda o en un documento que proporcione información sobre las regiones disponibles. Si analizara únicamente cero clics, no obtendría la información completa y, posiblemente, sacaría conclusiones erróneas sobre el éxito de la configuración de su aplicación de búsqueda a la hora de gestionar esta consulta.

Otro ejemplo de una visión práctica es descubrir una oportunidad de negocio. Las empresas suelen buscar oportunidades para aumentar sus clientes analizando las métricas de búsqueda. El documento en el que más clics recibe es «Regiones disponibles». Además, la mayoría de las consultas más buscadas están relacionadas con preguntas sobre la disponibilidad de productos en la región de Oceanía, con tasas de respuesta instantáneas del 100 por ciento y una alta tasa de clics para obtener más información sobre las regiones disponibles como parte de la respuesta. Esto sugiere que hay interés y demanda por tu producto o servicio en esta región.

Visualización y generación de informes de análisis de búsqueda

Hay cinco métricas que incluyen datos de tendencias para que puedas visualizar y buscar tendencias o patrones a lo largo del tiempo. Si utilizas la consola, se proporcionan gráficos de los datos de tendencias. Si utilizas las API, puedes recuperar los datos de tendencias para crear tus propios gráficos o visualizaciones. La mayoría de los gráficos de la consola muestran los puntos de datos diarios durante la ventana de tiempo elegida.

La consola proporciona un panel con las métricas en el que puede seleccionar un gráfico y una lista de los principales que le interese ver. Para exportar las métricas que se muestran en el panel de control en formato CSV, selecciona Exportar en la página principal de Analytics. Puede incluir estos informes en sus documentos o presentaciones empresariales.

Puede visualizar las siguientes métricas:

Gráfico de consultas totales

Un gráfico lineal del número de consultas emitidas por día. El gráfico le ayuda a visualizar los patrones de interacción diaria de los usuarios. Algunos ejemplos incluyen un aumento o una disminución constantes de la participación de los usuarios, o una caída drástica a 0 consultas debido a un fallo de la aplicación de búsqueda o a problemas con tu sitio web.

Si usa la API, puede recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede utilizar los datos para crear sus propios gráficos o utilizar los gráficos que se proporcionan en la consola.

Gráfico de tasa de clics

Un gráfico lineal de las proporciones de clics por día. El gráfico le ayuda a visualizar los patrones de la tasa de clics diaria. Algunos ejemplos incluyen un aumento o una disminución constantes de la tasa de clics, o una disminución de las respuestas instantáneas que podría influir en un aumento de los clics.

Si usa la API, puede recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede utilizar los datos para crear sus propios gráficos o utilizar los gráficos que se proporcionan en la consola.

Gráfico de tasa de clics cero

Un gráfico lineal de la proporción de clics cero por día. El gráfico le ayuda a visualizar patrones con una tasa diaria de clics cero. Algunos ejemplos incluyen un aumento o una disminución constantes de la tasa de clics cero, o un aumento de las respuestas instantáneas que podría influir en un aumento de clics cero.

Si usa la API, puede recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede utilizar los datos para crear sus propios gráficos o utilizar los gráficos que se proporcionan en la consola.

Gráfico de tasas de cero resultados de búsqueda

Un gráfico lineal de la proporción de resultados de búsqueda cero por día. El gráfico te ayuda a visualizar patrones con una tasa diaria cero de resultados de búsqueda. Algunos ejemplos incluyen un aumento o una disminución constantes de la tasa de resultados de búsqueda cero, o una disminución brusca del número de documentos del índice, lo que podría influir en un aumento de resultados de búsqueda cero.

Si usa la API, puede recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede utilizar los datos para crear sus propios gráficos o utilizar los gráficos que se proporcionan en la consola.

Gráfico de tasa de respuesta instantánea

Un gráfico lineal de la proporción de consultas a las que se ha devuelto una respuesta instantánea o preguntas frecuentes. El gráfico le ayuda a visualizar patrones en la tasa de respuesta instantánea diaria. Algunos ejemplos incluyen un aumento o una disminución constantes de las consultas del tipo pregunta-respuesta, o una disminución de los clics que podría influir en un aumento de las respuestas instantáneas.

Si usa la API, puede recuperar estos datos especificando `TREND_QUERY_DOC_METRICS`. Puede utilizar los datos para crear sus propios gráficos o utilizar los gráficos que se proporcionan en la consola.

Enviar comentarios para un aprendizaje incremental

Amazon Kendra utiliza el aprendizaje incremental para mejorar los resultados de búsqueda. Al utilizar los comentarios de las consultas, el aprendizaje incremental mejora los algoritmos de clasificación y optimiza los resultados de búsqueda para una mayor precisión.

Por ejemplo, supongamos que sus usuarios buscan la frase «beneficios de atención médica». Si los usuarios eligen sistemáticamente el segundo resultado de la lista, con el Amazon Kendra tiempo ese resultado pasará a ocupar el primer lugar. El aumento disminuye con el tiempo, por lo que si los usuarios dejan de seleccionar un resultado, Amazon Kendra eventualmente lo elimina y muestra otro resultado más popular en su lugar. Esto ayuda a Amazon Kendra priorizar los resultados en función de la relevancia, la antigüedad y el contenido.

El aprendizaje incremental está activado para todos los índices y para todos los tipos de [documentos admitidos](#).

Amazon Kendra empieza a aprender en cuanto proporcionas tus comentarios, aunque pueden pasar más de 24 horas hasta que veas los resultados de los comentarios. Amazon Kendra proporciona tres métodos para enviar comentarios: la AWS consola, una JavaScript biblioteca que puedes incluir en la página de resultados de búsqueda y una API que puedes usar.

Amazon Kendra acepta dos tipos de comentarios de los usuarios:

- **Clics:** información sobre los resultados de la consulta que eligió el usuario. Los comentarios incluyen el identificador del resultado y la marca de tiempo en Unix de la fecha y la hora en que se eligió el resultado de la búsqueda.

Para enviar comentarios sobre los clics, la aplicación debe recopilar la información sobre los clics de las actividades de sus usuarios y, a continuación, enviar esa información a Amazon Kendra. Puede recopilar información sobre los clics con la consola, la JavaScript biblioteca y la Amazon Kendra API.

- **Relevancia:** información sobre la relevancia de un resultado de búsqueda, que suele proporcionar el usuario. La valoración contiene el identificador del resultado y un indicador de relevancia (RELEVANT o NOT_RELEVANT). El usuario determina la información relevante.

Para enviar comentarios sobre la relevancia, la aplicación debe proporcionar un mecanismo de comentarios que permita al usuario elegir la relevancia adecuada para el resultado de una consulta y, a continuación, enviar esa información a Amazon Kendra. Solo puedes recopilar información relevante con la consola y la Amazon Kendra API.

Los comentarios se utilizan mientras el índice está activo. Los comentarios solo afectan al índice al que se envían, no se pueden usar en todos los índices ni para cuentas diferentes.

Debe proporcionar un contexto de usuario adicional cuando consulte su Amazon Kendra índice. Al proporcionar contexto al usuario, Amazon Kendra puede saber si los comentarios los ha proporcionado un solo usuario o varios usuarios y ajustar los resultados de la búsqueda en consecuencia.

Al proporcionar un contexto de usuario, los comentarios de la consulta se asocian al usuario específico proporcionado en el contexto. Si no especificas el contexto del usuario, puedes proporcionar un identificador de visitante que se utilice para agrupar y agregar consultas.

Si no proporcionas el contexto del usuario ni un identificador de visitante, los comentarios son anónimos y se suman a otros comentarios anónimos.

El código siguiente muestra cómo incluir el contexto del usuario como un símbolo o identificador de visitante.

```
response = kendra.query(  
    QueryText = query,  
    IndexId = index,  
    UserToken = {  
        Token = "token"  
    })  
  
OR  
  
response = kendra.query(  
    QueryText = query,  
    IndexId = index,  
    VisitorId = "visitor-id")
```

En el caso de las aplicaciones web, puede utilizar cookies, ubicaciones o usuarios del navegador para generar un identificador de visitante para cada usuario.

En el caso de las consultas principales, el mayor volumen de consultas, proporcionar comentarios basados en los clics proporciona suficiente información para mejorar la precisión general. En el caso de las consultas finales, aquellas que son poco frecuentes, los expertos en la materia deben enviar comentarios relevantes y no relevantes para mejorar la precisión de esas consultas.

Además de la consola, puede utilizar uno de los dos métodos siguientes: una JavaScript biblioteca o la [SubmitFeedbackAPI](#). Solo debes usar un método para recopilar comentarios. Para obtener los mejores resultados, debes enviar tus comentarios en un plazo de 24 horas después de realizar la consulta.

Temas

- [Uso de la Amazon Kendra JavaScript biblioteca para enviar comentarios \(p. 558\)](#)
- [Uso de la Amazon Kendra API para enviar comentarios \(p. 561\)](#)

Uso de la Amazon Kendra JavaScript biblioteca para enviar comentarios

Amazon Kendra proporciona una JavaScript biblioteca que puede utilizar para añadir comentarios sobre los clics a la página de resultados de la búsqueda. Para utilizar la biblioteca, inserte una etiqueta de script en el código de cliente que muestre el resultado de la búsqueda y, a continuación, añada información a cada uno de los enlaces de documentos de la lista de resultados. Cuando un usuario elige un enlace para ver un documento, se envía la información sobre los clics a Amazon Kendra.

La biblioteca funciona con navegadores compatibles con la JavaScript versión ES6/ES2015.

Paso 1: Inserta una etiqueta de script en tu aplicación Amazon Kendra de búsqueda

En el código de cliente que representa los resultados de la Amazon Kendra búsqueda, inserte una `<script>`etiqueta y añada una referencia a la JavaScript biblioteca:

```
<script>
(function(w, d, s, c, g, n) {
    if(!w[n]) {
        w[n] = w[n] || function () {
            (w[n].q = w[n].q || []).push(arguments);
        }
        w[n].st = new Date().getTime();
        w[n].ep = g;
        var e = document.createElement(s),
            j = document.getElementsByTagName(s)[0];
        e.async = 1;
        e.src = c;
        e.type = 'module';
        j.parentNode.insertBefore(e, j);
    }
})(window, document, 'script',
'library download URL',
'feedback endpoint',
'kendraFeedback');
</script>
```

El script descarga la JavaScript biblioteca de forma asíncrona desde una CDN Amazon Kendra alojada e inicializa una variable global llamada kendraFeedback que permite establecer parámetros opcionales.

Sustituya *la URL de descarga de la biblioteca* y el *punto final de comentarios* por un identificador de la siguiente tabla en función de la región en la que se aloja el Amazon Kendra índice.

Región	Descargar URL	Punto final de retroalimentación
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://d2iezfnpncoujy.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://d3gptlxulu4us.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

Región	Descargar URL	Punto final de retroalimentación
ap-northeast-1	https://d3w0ybsa293kb4.cloudfront.net/ksf-v1.js	https://wg7rz0uzjh.execute-api.ap-northeast-1.amazonaws.com/prod/submit
eu-west-2	https://d1tsrujswld1d1.cloudfront.net/ksf-v1.js	https://qi7mct3x7f.execute-api.eu-west-2.amazonaws.com/prod/submit

Por ejemplo, si tu índice se encuentra en EE. UU. Este (norte de Virginia), la *URL de descarga de la biblioteca* es <https://d2zm01pns956f8.cloudfront.net/ksf-v1.js> y el *punto final de los comentarios* es <https://ujxwp5s92h.execute-api.us-east-1.amazonaws.com/prod/submit>.

Hay dos ajustes opcionales que puede realizar para la Amazon Kendra JavaScript biblioteca:

- `disableCookies`— De forma predeterminada, Amazon Kendra establece una cookie que identifica de forma única al usuario. Establezcalo para deshabilitar la cookie.

```
kendraFeedback('disableCookie', 'true | false');
```

`searchDivClassName`— De forma predeterminada, Amazon Kendra monitorea todos los enlaces de la página de resultados de búsqueda para ver si hay clics. Configure esto en un nombre de `<div>` clase para supervisar solo los enlaces de la clase especificada.

```
kendraFeedback('searchDivClassName', 'class name');
```

Paso 2: Añade el token de valoración a los resultados de la búsqueda

En la página de resultados, añade un atributo HTML llamado `data-kendra-token` a la etiqueta de anclaje o a la etiqueta div principal inmediata que contenga un enlace al documento desde la respuesta a la consulta. Por ejemplo:

```
<a href="document location" data-kendra-token="feedback token value"></a>  
OR  
<div data-url="document location" data-kendra-token="feedback token value"></div>
```

La respuesta de una consulta contiene un símbolo en el `feedbackToken` campo. El token identifica de forma única la respuesta si el usuario la elige. Asigne el valor del token al `data-kendra-token` atributo. La Amazon Kendra JavaScript biblioteca busca este token cuando el usuario elige el resultado y lo envía a un Amazon Kendra punto final como comentario.

La Amazon Kendra JavaScript biblioteca solo envía el token de comentario y otros metadatos, como la hora en que se eligió el resultado y un identificador de visitante único.

Paso 3: Probar el script de comentarios

Para asegurarse de que la JavaScript biblioteca está configurada correctamente y de enviar comentarios al punto final correcto, haga lo siguiente. En este ejemplo se utiliza el navegador Chrome.

1. Abra las herramientas para desarrolladores web en el navegador. En Chrome, abre el menú de Chrome en la esquina superior derecha del navegador, selecciona Más herramientas y, a continuación, selecciona Herramientas para desarrolladores.
2. Asegúrese de que no haya errores relacionados con la Amazon Kendra JavaScript biblioteca en la pestaña de la consola.
3. Realice una búsqueda y elija cualquier resultado. En la pestaña Red de las herramientas para desarrolladores. Deberías ver una solicitud enviada al terminal de comentarios, el token del resultado y un estado de 200 OK.

Uso de la Amazon Kendra API para enviar comentarios

Para usar la Amazon Kendra API para enviar comentarios sobre consultas, usa la [SubmitFeedbackAPI](#). Para identificar la consulta, debe proporcionar el IndexID índice al que se aplica la consulta y el QueryId devuelto en la respuesta de la API de [consultas](#).

En el siguiente ejemplo, se muestra cómo enviar comentarios sobre los clics y la relevancia mediante la Amazon Kendra API. Puede enviar varios conjuntos de comentarios a través de las RelevanceFeedbackItems matrices ClickFeedbackItems y. En este ejemplo se envía un solo clic y un único elemento de valoración sobre la relevancia. El envío de comentarios utiliza la hora actual.

Para enviar comentarios para una búsqueda (AWSSDK)

1. Usa el siguiente código y cambia los siguientes valores:
 - a. index_id: cambie al ID del índice al que se aplica la consulta.
 - b. query_id—Cambia a la consulta sobre la que quieras enviar comentarios.
 - c. result_id—Cambio el ID del resultado de la consulta sobre el que desea enviar comentarios. La respuesta a la consulta contiene el ID del resultado.
 - d. relevance_value—Cambio a RELEVANT (el resultado de la consulta es relevante) o NOT_RELEVANT (el resultado de la consulta no es relevante).

Python

```
import boto3
import time

kendra = boto3.client("kendra")

# Provide the index ID
index_id = "index-id"
# Provide the query ID
query_id = "query-id"
# Provide the search result ID
result_id = "result-id"

# Configure the feedback item
feedback_item = {"ClickTime": int(time.time()),
                 "ResultId":result_id}

# Configure the relevance value
relevance_value = "RELEVANT"
relevance_item = {"RelevanceValue": relevance_value,
                  "ResultId": result_id
                 }
```

```
response = kendra.submit_feedback(  
    QueryId = query_id,  
    IndexId = index_id,  
    ClickFeedbackItems = [feedback_item],  
    RelevanceFeedbackItems = [relevance_item]  
)  
  
print("Submitted feedback for query: " + query_id)
```

Java

```
package com.amazonaws.kendra;  
  
import java.time.Instant;  
import software.amazon.awssdk.services.kendra.KendraClient;  
import software.amazon.awssdk.services.kendra.model.ClickFeedback;  
import software.amazon.awssdk.services.kendra.model.RelevanceFeedback;  
import software.amazon.awssdk.services.kendra.model.RelevanceType;  
import software.amazon.awssdk.services.kendra.model.SubmitFeedbackRequest;  
import software.amazon.awssdk.services.kendra.model.SubmitFeedbackResponse;  
  
public class SubmitFeedbackExample {  
    public static void main(String[] args) {  
        KendraClient kendra = KendraClient.builder().build();  
  
        SubmitFeedbackRequest submitFeedbackRequest = SubmitFeedbackRequest  
            .builder()  
            .indexId("anIndexId")  
            .queryId("aQueryId")  
            .clickFeedbackItems(  
                ClickFeedback  
                    .builder()  
                    .clickTime(Instant.now())  
                    .resultId("aResultId")  
                    .build())  
            .relevanceFeedbackItems(  
                RelevanceFeedback  
                    .builder()  
                    .relevanceValue(RelevanceType.RELEVANT)  
                    .resultId("aResultId")  
                    .build())  
            .build();  
  
        SubmitFeedbackResponse response =  
            kendra.submitFeedback(submitFeedbackRequest);  
  
        System.out.println("Feedback is submitted");  
    }  
}
```

2. Ejecute el código. Una vez enviados los comentarios, el código muestra un mensaje.

Añadir sinónimos personalizados a un índice

Para añadir sinónimos personalizados a un índice, debe especificarlos en un archivo de sinónimos. Puede incluir términos especializados o específicos de la empresa al Amazon Kendra usar sinónimos. Los sinónimos genéricos en inglés `leader`, `head`, `como`, están integrados Amazon Kendra y no deben incluirse en un archivo de sinónimos. Amazon Kendra admite sinónimos para todos los tipos de respuestas, que incluyen tipos de DOCUMENT respuesta QUESTION_ANSWER y/o tipos de ANSWER respuesta. Amazon Kendra actualmente no admite la adición de sinónimos marcados como palabras de parada. Esto se incluirá en una versión futura.

Amazon Kendra establece correlaciones entre sinónimos. Por ejemplo, al usar el par de sinónimos `Dynamo`, `Amazon DynamoDB`, Amazon Kendra correlaciona `Dynamo` con `Amazon DynamoDB`. La consulta «¿Qué es la dinamo?» luego devuelve un documento como «¿Qué es `Amazon DynamoDB`?». Con sinónimos, Amazon Kendra puede captar la correlación más fácilmente.

El archivo de sinónimos es un archivo de texto que se almacena en un Amazon S3 bucket. Consulte [Añadir un diccionario de sinónimos a un índice \(p. 566\)](#).

El archivo de sinónimos utiliza el formato de [sinónimos Solr](#). Amazon Kendra tiene un límite en el número de tesauros por índice. Consulte [Cuotas](#).

Los sinónimos pueden resultar útiles en los siguientes escenarios:

- Términos especializados que no son sinónimos tradicionales del idioma inglés, como NLP, Natural Language Processing.
- Sustantivos propios con asociaciones semánticas complejas. Estos son sustantivos que es poco probable que el público en general comprenda, por ejemplo, en el aprendizaje automático. `cost`, `loss`, `model performance`
- Diferentes formas de nombres de productos, por ejemplo, Elastic Compute Cloud, EC2.
- Términos específicos del dominio o de la empresa, como nombres de productos. Por ejemplo, Route53, DNS.

No utilice sinónimos en los siguientes escenarios:

- Sinónimos genéricos en inglés como `leader`, `head`. Estos sinónimos no son específicos de un dominio y el uso de sinónimos en estos escenarios puede tener efectos no deseados.
- Errores tipográficos como `teh => the`
- Variantes morfológicas como los plurales y posesivos de los sustantivos, la forma comparativa y superlativa de los adjetivos y el tiempo pasado, el participio pasado y la forma progresiva de los verbos. Un ejemplo de adjetivos comparativos y superlativos es. `good`, `better`, `best`
- Palabras de terminación en Unigram (palabra única), como WHO. Las palabras terminativas en Unigram no están permitidas en el diccionario de sinónimos y se excluyen de la búsqueda. Por ejemplo, `WHO => World Health Organization` se rechaza. `W.H.O.` Sin embargo, puede usarlo como un término sinónimo y puede usar palabras de parada como parte de un sinónimo de varias palabras. Por ejemplo, `no of` está permitido pero `United States of America` se acepta.

Los sinónimos personalizados facilitan la comprensión Amazon Kendra de la terminología específica de su empresa al ampliar las consultas para incluir los sinónimos específicos de su empresa. Si bien

los sinónimos pueden mejorar la precisión de la búsqueda, es importante entender cómo afectan los sinónimos a la latencia para poder optimizarlos.

Una regla general para los sinónimos es: cuantos más términos de la consulta coincidan y se amplíen con sinónimos, mayor será el impacto potencial en la latencia. Otros factores que afectan a la latencia son el tamaño promedio de los documentos indexados, el tamaño del índice, cualquier filtrado de los resultados de búsqueda y la carga total del índice. Amazon Kendra Las consultas que no coinciden con ningún sinónimo no se ven afectadas.

Una guía general sobre cómo los sinónimos afectan a la latencia:

Caso de uso	Aumento de latencia*
Consultas típicas en lenguaje natural o palabras clave de 3 a 5 palabras cada una	Menos del 15 por ciento
1 término de consulta se expande a 3 sinónimos	
Índice de unos 500 000 documentos (con un promedio de 10,48 KB de texto extraído por documento) o 30 000 pares de preguntas frecuentes/preguntas	

* El rendimiento varía en función del uso específico de los sinónimos y las configuraciones del índice. Es mejor probar el rendimiento de la búsqueda para obtener puntos de referencia más precisos para tu caso de uso específico.

Si el diccionario de sinónimos es grande, tiene una tasa de expansión a largo plazo y el aumento de latencia no está dentro de los límites aceptables, puede probar una de las siguientes opciones o ambas:

- Recorta tu diccionario de sinónimos para reducir la relación de expansión (número de sinónimos por término).
- Recorta la cobertura general de términos (número de líneas en tu diccionario de sinónimos).

Como alternativa, puede aumentar la capacidad de aprovisionamiento (unidades de almacenamiento virtuales) para compensar el aumento de la latencia.

Temas

- [Creación de un archivo de sinónimos \(p. 564\)](#)
- [Añadir un diccionario de sinónimos a un índice \(p. 566\)](#)
- [Actualización de un diccionario de sinónimos \(p. 569\)](#)
- [Eliminar un diccionario de sinónimos \(p. 572\)](#)
- [Aspectos destacados en los resultados de búsqueda \(p. 573\)](#)

Creación de un archivo de sinónimos

Un archivo de Amazon Kendra sinónimos es un archivo codificado en UTF-8 que contiene una lista de sinónimos en el formato de lista de sinónimos de Solr. El archivo de sinónimos debe tener menos de 5 MB.

Hay dos maneras de especificar los mapeos de sinónimos:

- Los sinónimos bidireccionales se especifican como una lista de términos separados por comas. Si el usuario consulta alguno de los términos, todos los términos de la lista se utilizan para buscar documentos, incluidos los términos consultados originalmente.

- Los sinónimos unidireccionales se especifican como términos separados por el símbolo «=>» entre ellos para asignar los términos a sus sinónimos. Si el usuario consulta un término situado a la izquierda del símbolo «=>», se asigna a un término de la derecha para buscar documentos con el sinónimo. No está mapeado al revés, lo que lo hace unidireccional.

Los sinónimos en sí mismos distinguen entre mayúsculas y minúsculas, pero los términos a los que se asignan no distinguen entre mayúsculas y minúsculas. Por ejemplo, ML => Machine Learning significa que si el usuario consulta «ML» o «ml» o utiliza algún otro caso, se asignará a «Aprendizaje automático». Si tuviera que mapear esto al revésMachine Learning => ML, entonces «Aprendizaje automático» o algún otro caso se asignaría a «ML».

El siguiente ejemplo muestra un archivo de sinónimos con sinónimos de la AWS documentación de ejemplo de Amazon Kendra. Cada línea contiene una única regla de sinónimos. Un sinónimo no coincide exactamente con los caracteres especiales. Por ejemplo, si buscasdead-letter-queue, Kendra hace coincidir los documentos con la frase dead letter queue. Se ignoran las líneas en blanco y los comentarios.

```
# Lines starting with pound are comments and blank lines are ignored.

# Synonym relationships can be defined as unidirectional or bidirectional relationships.

# Unidirection relationships are represented by any term sequence
# on the left hand side (LHS) of "=" followed by synonyms on the right hand side (RHS)
CodeStar => AWS CodeStar
# This will map CodeStar to AWS CodeStar, but not vice-versa

# To map terms vice versa
ML => Machine Learning
Machine Learning => ML

# Multiple synonym relationships may be defined in one line as well by comma separation.
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

# Synonym rule count is the total number of lines defining synonym relationships
# Term count is the total number of unique terms for all rules.
# This thesaurus has a synonym rule count of 6 and a term count of 18.
# Comments and blanks lines do not count.
```

Este ejemplo tiene 6 reglas y 18 términos. Cada línea contiene una única regla de sinónimos. Un sinónimo no coincide exactamente con los caracteres especiales. Por ejemplo, si buscasdead-letter-queue, Kendra buscará los documentos que coincidan con la cola de cartas muertas. Se ignoran las líneas en blanco y los comentarios. Se ignoran algunas reglas. Por ejemplo, a => b es una regla, pero a => a se ignora y no

cuenta como regla. Un sinónimo no coincide exactamente con los caracteres especiales. Por ejemplo, si buscadead-letter-queue, Amazon Kendra coincidirá con el documento que contiene dead letter queue (sin guión). Puede tener un máximo de 10 000 reglas de sinónimos por tesauro.

El recuento de términos es el número de términos únicos en el archivo theaurus. En este ejemplo se incluyen los siguientes términos: AWS CodeStar autoscaling groupasg, Auto Scaling groupautoscaling,DNS,Route53,Route 53,dns, 0, 1, 2, 3, 4, 5, y 6.route53route 53betaAlphaGammaDeltadelta Puedes tener hasta 10 sinónimos por término.

Para obtener más información acerca de las cuotas de Amazon Kendra, consulte [Cuotas para Amazon Kendra \(p. 663\)](#).

Añadir un diccionario de sinónimos a un índice

Los siguientes procedimientos muestran cómo añadir un archivo de sinónimos que contenga sinónimos a un índice. Los efectos del archivo de sinónimos actualizado pueden tardar hasta 30 minutos. Para obtener más información sobre el archivo de sinónimos, consulte [Creación de un archivo de sinónimos \(p. 564\)](#).

Console

Para añadir un diccionario de sinónimos

1. En el panel de navegación de la izquierda, debajo del índice en el que quieras añadir una lista de sinónimos (tu diccionario de sinónimos), selecciona Sinónimos.
2. En la página de sinónimos, selecciona Agregar diccionario de sinónimos.
3. En Definir diccionario de sinónimos, asigne un nombre al diccionario de sinónimos y una descripción opcional.
4. En la configuración del diccionario de sinónimos, proporcione la Amazon S3 ruta al archivo de sinónimos. El archivo debe tener un tamaño inferior a 5 MB.
5. Para Función de IAM, seleccione una función o seleccione Crear una nueva función y especifique un nombre de función para crear una nueva función. Amazon Kendra usa este rol para acceder al Amazon S3 recurso en su nombre. El rol de IAM tiene el prefijo "AmazonKendra->".
6. Elija Guardar para guardar la configuración y añadir el diccionario de sinónimos. Una vez que se ingiere el diccionario de sinónimos, se activa y los sinónimos se resaltan en los resultados. Puede tardar hasta 30 minutos en ver los efectos del archivo de sinónimos.

CLI

Para añadir un diccionario de sinónimos a un índice con el AWS CLI, llame a: `create-thesaurus`

```
aws kendra create-thesaurus \
--index-id index-id \
--name "thesaurus-name" \
--description "thesaurus-description" \
--source-s3-path "Bucket=bucket-name,Key=thesaurus/synonyms.txt" \
--role-arn role-arn
```

Llame `list-thesauri` para ver una lista de tesauros:

```
aws kendra list-thesauri \
--index-id index-id
```

Para ver los detalles de un diccionario de sinónimos, llame `describe-thesaurus`:

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

Puede tardar hasta 30 minutos en ver los efectos del archivo de sinónimos.

Python

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

kendra = boto3.client("kendra")

print("Create a thesaurus")

thesaurus_name = "thesaurus-name"
thesaurus_description = "thesaurus-description"
thesaurus_role_arn = "role-arn"

index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "thesaurus-file"
source_s3_path= {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    thesaurus_response = kendra.create_thesaurus(
        Description = thesaurus_description,
        Name = thesaurus_name,
        RoleArn = thesaurus_role_arn,
        IndexId = index_id,
        SourceS3Path = source_s3_path
    )

    pprint.pprint(thesaurus_response)

    thesaurus_id = thesaurus_response["Id"]

    print("Wait for Kendra to create the thesaurus.")

    while True:
        # Get thesaurus description
        thesaurus_description = kendra.describe_thesaurus(
            Id = thesaurus_id,
            IndexId = index_id
        )
        # If status is not CREATING quit
        status = thesaurus_description["Status"]
        print("Creating thesaurus. Status: " + status)
        if status != "CREATING":
            break
        time.sleep(60)

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

print("Program ends.")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.CreateThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.CreateThesaurusResponse;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusResponse;
import software.amazon.awssdk.services.kendra.model.S3Path;
import software.amazon.awssdk.services.kendra.model.ThesaurusStatus;

public class CreateThesaurusExample {

    public static void main(String[] args) throws InterruptedException {
        KendraClient kendra = KendraClient.builder().build();

        String thesaurusName = "thesaurus-name";
        String thesaurusDescription = "thesaurus-description";
        String thesaurusRoleArn = "role-arn";

        String s3BucketName = "bucket-name";
        String s3Key = "thesaurus-file";
        String indexId = "index-id";

        System.out.println(String.format("Creating a thesaurus named %s", thesaurusName));
        CreateThesaurusRequest createThesaurusRequest = CreateThesaurusRequest
            .builder()
            .name(thesaurusName)
            .indexId(indexId)
            .description(thesaurusDescription)
            .roleArn(thesaurusRoleArn)
            .sourceS3Path(S3Path.builder()
                .bucket(s3BucketName)
                .key(s3Key)
                .build())
            .build();
        CreateThesaurusResponse createThesaurusResponse =
            kendra.createThesaurus(createThesaurusRequest);
        System.out.println(String.format("Thesaurus response %s",
            createThesaurusResponse));

        String thesaurusId = createThesaurusResponse.id();

        System.out.println(String.format("Waiting until the thesaurus with ID %s is
            created.", thesaurusId));

        while (true) {
            DescribeThesaurusRequest describeThesaurusRequest =
            DescribeThesaurusRequest.builder()
                .id(thesaurusId)
                .indexId(indexId)
                .build();
            DescribeThesaurusResponse describeThesaurusResponse =
            kendra.describeThesaurus(describeThesaurusRequest);
            ThesaurusStatus status = describeThesaurusResponse.status();
            if (status != ThesaurusStatus.CREATING) {
                break;
            }
            TimeUnit.SECONDS.sleep(60);
        }
        System.out.println("Thesaurus creation is complete.");
    }
}
```

}

Actualización de un diccionario de sinónimos

Puede cambiar la configuración de un diccionario de sinónimos después de crearlo. Puede cambiar detalles como el nombre del diccionario de sinónimos y la información de IAM. También puede cambiar la ubicación de la ruta Amazon S3 del archivo de sinónimos. Si cambia la ruta del archivo de sinónimos, Amazon Kendra sustituye el diccionario de sinónimos existente por el diccionario de sinónimos especificado en la ruta actualizada.

Los efectos del archivo de sinónimos actualizado pueden tardar hasta 30 minutos.

Note

Si hay errores de validación o de sintaxis en el archivo de sintaxis, se conserva el archivo de sintaxis cargado anteriormente.

Los siguientes procedimientos muestran cómo modificar los detalles del diccionario de sinónimos.

Console

Para modificar los detalles del diccionario de sinónimos

1. En el panel de navegación de la izquierda, debajo del índice que desea modificar, elija Sinónimos.
2. En la página de sinónimos, seleccione el diccionario de sinónimos que desee modificar y, a continuación, elija Editar.
3. En la página Actualizar el diccionario de sinónimos, actualice los detalles del diccionario de sinónimos.
4. (Opcional) Elija Cambiar la ruta del archivo de sinónimos y, a continuación, especifique una Amazon S3 ruta al nuevo archivo de sinónimos. El archivo de sinónimos existente se sustituye por el archivo que especifique. Si no cambia la ruta, Amazon Kendra vuelve a cargar el diccionario de sinónimos desde la ruta existente.

Si selecciona Conservar el archivo de sinónimos actual, Amazon Kendra no se vuelve a cargar el archivo de sinónimos.

5. Seleccione Guardar para guardar la configuración.

También puede volver a cargar el diccionario de sinónimos desde la ruta del diccionario de sinónimos existente.

Para volver a cargar un diccionario de sinónimos desde una ruta existente

1. En el panel de navegación de la izquierda, debajo del índice que desea modificar, elija Sinónimos.
2. En la página de sinónimos, seleccione el diccionario de sinónimos que desee volver a cargar y, a continuación, elija Recargar.
3. En la página Recargar el archivo de sinónimos, confirme que desea volver a cargar el archivo de sinónimos.

CLI

Para actualizar un diccionario de sinónimos, llame al `update-thesaurus`:

```
aws kendra update-thesaurus \
--index-id index-id \
```

```
--name "thesaurus-name" \
--description "thesaurus-description" \
--source-s3-path "Bucket=bucket-name,Key=thesaurus/synonyms.txt" \
--role-arn role-arn
```

Python

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

kendra = boto3.client("kendra")

print("Update a thesaurus")

thesaurus_name = "thesaurus-name"
thesaurus_description = "thesaurus-description"
thesaurus_role_arn = "role-arn"

thesaurus_id = "thesaurus-id"
index_id = "index-id"

s3_bucket_name = "bucket-name"
s3_key = "thesaurus-file"
source_s3_path= {
    'Bucket': s3_bucket_name,
    'Key': s3_key
}

try:
    kendra.update_thesaurus(
        Id = thesaurus_id,
        IndexId = index_id,
        Description = thesaurus_description,
        Name = thesaurus_name,
        RoleArn = thesaurus_role_arn,
        SourceS3Path = source_s3_path
    )

    print("Wait for Kendra to update the thesaurus.")

    while True:
        # Get thesaurus description
        thesaurus_description = kendra.describe_thesaurus(
            Id = thesaurus_id,
            IndexId = index_id
        )
        # If status is not UPDATING quit
        status = thesaurus_description["Status"]
        print("Updating thesaurus. Status: " + status)
        if status != "UPDATING":
            break
        time.sleep(60)

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

print("Program ends.")
```

Java

```
package com.amazonaws.kendra;
```

```
import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.UpdateThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusRequest;
import software.amazon.awssdk.services.kendra.model.DescribeThesaurusResponse;
import software.amazon.awssdk.services.kendra.model.S3Path;
import software.amazon.awssdk.services.kendra.model.ThesaurusStatus;

public class UpdateThesaurusExample {

    public static void main(String[] args) throws InterruptedException {
        KendraClient kendra = KendraClient.builder().build();

        String thesaurusName = "thesaurus-name";
        String thesaurusDescription = "thesaurus-description";
        String thesaurusRoleArn = "role-arn";

        String s3BucketName = "bucket-name";
        String s3Key = "thesaurus-file";

        String thesaurusId = "thesaurus-id";
        String indexId = "index-id";

        UpdateThesaurusRequest updateThesaurusRequest = UpdateThesaurusRequest
            .builder()
            .id(thesaurusId)
            .indexId(indexId)
            .name(thesaurusName)
            .description(thesaurusDescription)
            .roleArn(thesaurusRoleArn)
            .sourceS3Path(S3Path.builder()
                .bucket(s3BucketName)
                .key(s3Key)
                .build())
            .build();
        kendra.updateThesaurus(updateThesaurusRequest);

        System.out.println(String.format("Waiting until the thesaurus with ID %s is updated.", thesaurusId));

        // a new source s3 path requires re-consumption by Kendra
        // and so can take as long as a Create Thesaurus operation
        while (true) {
            DescribeThesaurusRequest describeThesaurusRequest =
            DescribeThesaurusRequest.builder()
                .id(thesaurusId)
                .indexId(indexId)
                .build();
            DescribeThesaurusResponse describeThesaurusResponse =
            kendra.describeThesaurus(describeThesaurusRequest);
            ThesaurusStatus status = describeThesaurusResponse.status();
            if (status != ThesaurusStatus.UPDATING) {
                break;
            }

            TimeUnit.SECONDS.sleep(60);
        }

        System.out.println("Thesaurus update is complete.");
    }
}
```

Eliminar un diccionario de sinónimos

Los siguientes procedimientos muestran cómo eliminar un diccionario de sinónimos.

Console

1. En el panel de navegación de la izquierda, debajo del índice que desea modificar, elija Sinónimos.
2. En la página de sinónimos, seleccione el diccionario de sinónimos que deseé eliminar.
3. En la página de detalles del Tesauro, seleccione Eliminar y, a continuación, confirme la eliminación.

CLI

Para eliminar un diccionario de sinónimos de un índice con el AWS CLI, llame a: `delete-thesaurus`

```
aws kendra delete-thesaurus \
--index-id index-id \
--id thesaurus-id
```

Python

```
import boto3
from botocore.exceptions import ClientError

kendra = boto3.client("kendra")

print("Delete a thesaurus")

thesaurus_id = "thesaurus-id"
index_id = "index-id"

try:
    kendra.delete_thesaurus(
        Id = thesaurus_id,
        IndexId = index_id
    )

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

print("Program ends.")
```

Java

```
package com.amazonaws.kendra;

import software.amazon.awssdk.services.kendra.KendraClient;
import software.amazon.awssdk.services.kendra.model.DeleteThesaurusRequest;

public class DeleteThesaurusExample {

    public static void main(String[] args) throws InterruptedException {
        KendraClient kendra = KendraClient.builder().build();

        String thesaurusId = "thesaurus-id";
        String indexId = "index-id";

        DeleteThesaurusRequest updateThesaurusRequest = DeleteThesaurusRequest
```

```
.builder()
.id(thesaurusId)
.indexId(indexId)
.build();
kendra.deleteThesaurus(updateThesaurusRequest);
}
```

Aspectos destacados en los resultados de búsqueda

El resultado de sinónimos está activado de forma predeterminada. La información destacada se incluye en los resultados de las consultas del Amazon Kendra SDK y la CLI. Si interactúa Amazon Kendra con el SDK o la CLI, usted determina cómo mostrar los resultados.

Los subrayados sinónimos tendrán el tipo THESAURUS_SYNONYM de resultado. Para obtener más información sobre los resultados, consulte el objeto [Highlight](#).

Tutorial: Creación de una solución de búsqueda inteligente enriquecida con metadatos con Amazon Kendra

Este tutorial le muestra cómo crear una solución de búsqueda inteligente, basada en lenguaje natural y enriquecida con metadatos para sus datos empresariales mediante [Amazon Kendra](#), [Amazon Comprehend](#), [Amazon Simple Storage Service \(S3\)](#) y [AWS CloudShell](#).

Amazon Kendra es un servicio de búsqueda inteligente que puede crear un índice de búsqueda para sus repositorios de datos no estructurados en lenguaje natural. Para facilitar a sus clientes la búsqueda y el filtrado de las respuestas relevantes, puede utilizar Amazon Comprehend para extraer los metadatos de sus datos e incorporarlos a su índice de búsqueda de Amazon Kendra.

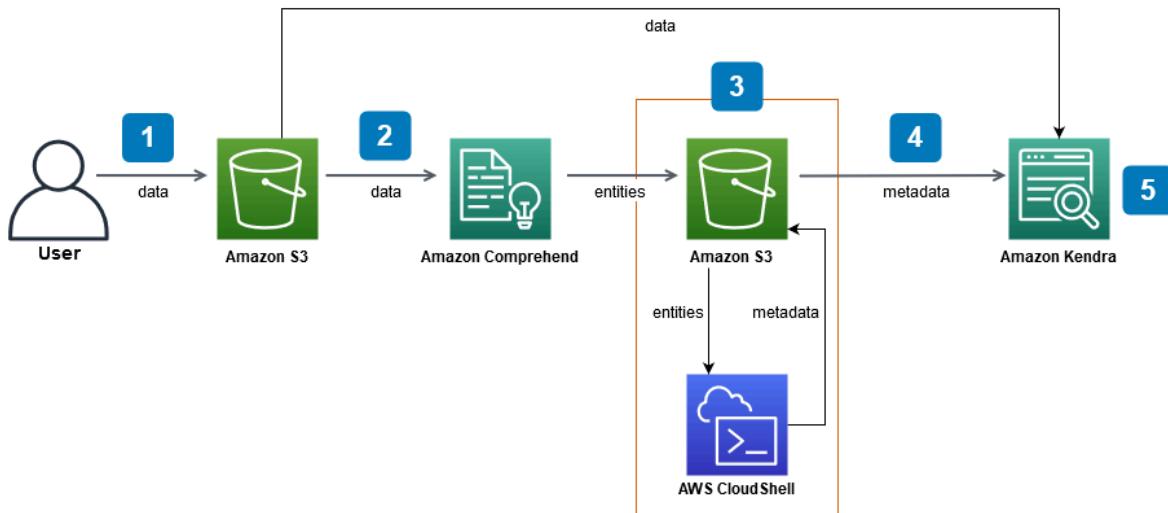
Amazon Comprehend es un servicio de procesamiento de lenguaje natural (PNL) que puede identificar entidades. Las entidades son referencias a personas, lugares, ubicaciones, organizaciones y objetos de sus datos.

Este tutorial utiliza un conjunto de datos de muestra de artículos de noticias para extraer entidades, convertirlas en metadatos e incorporarlas al índice de Amazon Kendra para realizar búsquedas en ellas. Los metadatos añadidos permiten filtrar los resultados de la búsqueda mediante cualquier subconjunto de estas entidades y mejoran la precisión de la búsqueda. Al seguir este tutorial, aprenderá a crear una solución de búsqueda para los datos de su empresa sin ningún conocimiento especializado en aprendizaje automático.

En este tutorial se muestra cómo crear una solución de búsqueda mediante los siguientes pasos:

1. Almacenamiento de un conjunto de datos de muestra de artículos de noticias en Amazon S3.
2. Uso de Amazon Comprehend para extraer entidades de sus datos.
3. Ejecutar un script de Python 3 para convertir las entidades al formato de metadatos de índice de Amazon Kendra y almacenar estos metadatos en S3.
4. Crear un índice de búsqueda de Amazon Kendra e incorporar los datos y los metadatos.
5. Consultar el índice de búsqueda.

El siguiente diagrama muestra el flujo de trabajo:



Tiempo estimado para completar este tutorial: 1 hora

Coste estimado: algunas de las acciones de este tutorial conllevan cargos en tu AWS cuenta. Para obtener más información sobre el costo de cada servicio, consulte las páginas de precios de [Amazon S3](#), [Amazon Comprehend](#) y [Amazon Kendra](#). [AWS CloudShell](#)

Temas

- [Requisitos previos \(p. 575\)](#)
- [Paso 1: Añadir documentos a Amazon S3 \(p. 576\)](#)
- [Paso 2: Ejecutar un trabajo de análisis de entidades en Amazon Comprehend \(p. 583\)](#)
- [Paso 3: Formatear la salida del análisis de entidades como metadatos de Amazon Kendra \(p. 589\)](#)
- [Paso 4: Crear un índice de Amazon Kendra e incorporar los metadatos \(p. 598\)](#)
- [Paso 5: Consultar el índice de Amazon Kendra \(p. 615\)](#)
- [Paso 6: Limpieza \(p. 622\)](#)

Requisitos previos

Para completar este tutorial, necesitará los siguientes recursos:

- Una cuenta de AWS. Si no tiene una AWS cuenta, siga los pasos que se indican en [Configuración de Amazon Kendra](#) para configurar su AWS cuenta.
- Un equipo de desarrollo con Windows, macOS o Linux para acceder a la consola AWS de administración. Para obtener más información, consulte [Configuración de la consola AWS de administración](#).
- Un usuario [AWS Identity and Access Management](#)(IAM). Para obtener información sobre cómo configurar un usuario y un grupo de IAM para su cuenta, consulte la sección [Introducción](#) de la Guía del usuario de IAM.

Si utiliza elAWS Command Line Interface, también debe adjuntar la siguiente política a su usuario de IAM para concederle los permisos básicos necesarios para completar este tutorial.

Para obtener más información, consulte [Crear políticas de IAM](#) y [Aregar y quitar permisos de identidad de IAM](#).

- La [lista de servicios AWS regionales](#). Para reducir la latencia, debe elegir la AWS región más cercana a su ubicación geográfica que sea compatible con Amazon Comprehend y Amazon Kendra.
- (Opcional) Un [AWS Key Management Service](#). Si bien este tutorial no utiliza el cifrado, es posible que desee utilizar las prácticas recomendadas de cifrado para su caso de uso específico.
- (Opcional) Una [nube privada virtual de Amazon](#). Si bien este tutorial no utiliza una VPC, es posible que desee utilizar las prácticas recomendadas de VPC para garantizar la seguridad de los datos para su caso de uso específico.

Paso 1: Añadir documentos a Amazon S3

Antes de ejecutar un trabajo de análisis de entidades de Amazon Comprehend en su conjunto de datos, debe crear un bucket de Amazon S3 para alojar los datos, los metadatos y el resultado del análisis de entidades de Amazon Comprehend.

Temas

- [Descargar el conjunto de datos de muestra \(p. 576\)](#)
- [Creación de un bucket de Amazon S3 \(p. 577\)](#)
- [Crear carpetas de datos y metadatos en su bucket de S3 \(p. 579\)](#)
- [Carga de los datos de entrada \(p. 581\)](#)

Descargar el conjunto de datos de muestra

Antes de que Amazon Comprehend pueda ejecutar un trabajo de análisis de entidades en sus datos, debe descargar y extraer el conjunto de datos y subirlo a un bucket de S3.

Para descargar y extraer el conjunto de datos (consola)

1. Descarga la carpeta [tutorial-dataset.zip](#) en tu dispositivo.
2. Extraiga la `tutorial-dataset` carpeta para acceder a data ella.

Para descargar y extraer el conjunto de datos (Terminal)

1. Para descargar el `tutorial-dataset`, ejecute el siguiente comando en una ventana de terminal:

Linux

```
curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/samples/tutorial-dataset.zip
```

Donde:

- *path/* es la ruta del archivo local a la ubicación en la que desea guardar la carpeta zip.

macOS

```
curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/samples/tutorial-dataset.zip
```

Donde:

- *path/* es la ruta del archivo local a la ubicación en la que desea guardar la carpeta zip.

Windows

```
curl -o path/tutorial-dataset.zip https://docs.aws.amazon.com/kendra/latest/dg/samples/tutorial-dataset.zip
```

Donde:

- *path/* es la ruta del archivo local a la ubicación en la que desea guardar la carpeta zip.
2. Para extraer los datos de la carpeta zip, ejecute el siguiente comando en la ventana del terminal:

Linux

```
unzip path/tutorial-dataset.zip -d path/
```

Donde:

- *path/* es la ruta del archivo local a la carpeta zip guardada.

macOS

```
unzip path/tutorial-dataset.zip -d path/
```

Donde:

- *path/* es la ruta del archivo local a la carpeta zip guardada.

Windows

```
tar -xf path/tutorial-dataset.zip -C path/
```

Donde:

- *path/* es la ruta del archivo local a la carpeta zip guardada.

Al final de este paso, deberías tener los archivos extraídos en una carpeta descomprimida llamada *tutorial-dataset*. Esta carpeta contiene un *README* archivo con una atribución de código abierto de Apache 2.0 y una carpeta llamada *data* que contiene el conjunto de datos de este tutorial. El conjunto de datos consta de 100 archivos con *.story* extensiones.

Creación de un bucket de Amazon S3

Tras descargar y extraer la carpeta de datos de muestra, la almacena en un bucket de Amazon S3.

Important

El nombre de un bucket de Amazon S3 debe ser único en todosAWS.

Para crear un bucket S3 (consola)

1. Inicie sesión en la AWS Management Console y abra la consola de Amazon S3 en <https://console.aws.amazon.com/s3/>.

2. En Depósitos, selecciona Crear depósito.
3. En Nombre del bucket, escriba un nombre único.
4. En Región, elija la AWS región en la que desea crear el depósito.

Note

Debe elegir una región que sea compatible con Amazon Comprehend y Amazon Kendra. No puede cambiar la región de un depósito después de haberlo creado.

5. Mantenga la configuración predeterminada de Bloquear acceso público para este bucket, el control de versiones del bucket y las etiquetas.
6. Para el cifrado predeterminado, selecciona Desactivar.
7. Mantenga la configuración predeterminada para la configuración avanzada.
8. Revisa la configuración del depósito y, a continuación, selecciona Crear depósito.

Para crear un bucket de S3 (AWS CLI)

1. Para crear un bucket de S3, utilice el comando [create-bucket](#) en: AWS CLI

Linux

```
aws s3api create-bucket \  
    --bucket DOC-EXAMPLE-BUCKET \  
    --region aws-region \  
    --create-bucket-configuration LocationConstraint=aws-region
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket,
- *aws-region* es la región en la que desea crear su bucket.

macOS

```
aws s3api create-bucket \  
    --bucket DOC-EXAMPLE-BUCKET \  
    --region aws-region \  
    --create-bucket-configuration LocationConstraint=aws-region
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket,
- *aws-region* es la región en la que desea crear su bucket.

Windows

```
aws s3api create-bucket ^  
    --bucket DOC-EXAMPLE-BUCKET ^  
    --region aws-region ^  
    --create-bucket-configuration LocationConstraint=aws-region
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket,
- *aws-region* es la región en la que desea crear su bucket.

Note

Debe elegir una región que sea compatible con Amazon Comprehend y Amazon Kendra. No puede cambiar la región de un depósito después de haberlo creado.

2. Para asegurarse de que el bucket se creó correctamente, utilice el comando [list](#):

Linux

```
aws s3 ls
```

macOS

```
aws s3 ls
```

Windows

```
aws s3 ls
```

Crear carpetas de datos y metadatos en su bucket de S3

Tras crear el bucket de S3, crea carpetas de datos y metadatos en su interior.

Para crear carpetas en tu bucket de S3 (consola)

1. Abra la consola de Amazon S3 en <https://console.aws.amazon.com/s3>.
2. En Depósitos, haz clic en el nombre de tu depósito de la lista de depósitos.
3. En la pestaña Objetos, selecciona Crear carpeta.
4. Para el nombre de la nueva carpeta, introduzca **data**.
5. Para la configuración de cifrado, seleccione Desactivar.
6. Elija Create folder.
7. Repita los pasos 3 a 6 para crear otra carpeta para almacenar los metadatos de Amazon Kendra y asigne un nombre a la carpeta creada en el paso 4**metadata**.

Para crear carpetas en tu bucket de S3 (AWS CLI)

1. Para crear la data carpeta en su bucket de S3, utilice el comando [put-object](#) en: AWS CLI

Linux

```
aws s3api put-object \
    --bucket DOC-EXAMPLE-BUCKET \
    --key data/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket.

macOS

```
aws s3api put-object \  
    --bucket DOC-EXAMPLE-BUCKET \  
    --key data/
```

Donde:

- *DOC-EXAMPLE-BUCKET es el nombre de tu bucket.*

Windows

```
aws s3api put-object ^  
    --bucket DOC-EXAMPLE-BUCKET ^  
    --key data/
```

Donde:

- *DOC-EXAMPLE-BUCKET es el nombre de tu bucket.*

2. Para crear la metadata carpeta en su bucket de S3, utilice el comando [put-object](#) en: AWS CLI

Linux

```
aws s3api put-object \  
    --bucket DOC-EXAMPLE-BUCKET \  
    --key metadata/
```

Donde:

- *DOC-EXAMPLE-BUCKET es el nombre de tu bucket.*

macOS

```
aws s3api put-object \  
    --bucket DOC-EXAMPLE-BUCKET \  
    --key metadata/
```

Donde:

- *DOC-EXAMPLE-BUCKET es el nombre de tu bucket.*

Windows

```
aws s3api put-object ^  
    --bucket DOC-EXAMPLE-BUCKET ^  
    --key metadata/
```

Donde:

- *DOC-EXAMPLE-BUCKET es el nombre de tu bucket.*

3. Para asegurarte de que las carpetas se crearon correctamente, comprueba el contenido de tu bucket con el comando [list](#):

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket.

Carga de los datos de entrada

Tras crear las carpetas de datos y metadatos, cargue el conjunto de datos de muestra en la data carpeta.

Para cargar el conjunto de datos de muestra en la carpeta de datos (consola)

1. Abra la consola de Amazon S3 en <https://console.aws.amazon.com/s3>.
2. En Depósitos, haga clic en el nombre de su depósito de la lista de depósitos y, a continuación, haga clic en data
3. Selecciona Cargar y, a continuación, selecciona Añadir archivos.
4. En el cuadro de diálogo, navegue hasta la data carpeta que se encuentra dentro de la tutorial-dataset carpeta del dispositivo local, seleccione todos los archivos y, a continuación, elija Abrir.
5. Conserve la configuración predeterminada para Destino, Permisos y Propiedades.
6. Seleccione Upload (Cargar).

Para cargar el conjunto de datos de muestra en la carpeta de datos () AWS CLI

1. Para cargar los datos de muestra en la data carpeta, utilice el comando [copy](#) deAWS CLI:

Linux

```
aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
```

Donde:

- *path*/ es la ruta del archivo a la tutorial-dataset carpeta de su dispositivo,

- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket.

macOS

```
aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
```

Donde:

- *path/* es la ruta del archivo a la tutorial-dataset carpeta de su dispositivo,
- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket.

Windows

```
aws s3 cp path/tutorial-dataset/data s3://DOC-EXAMPLE-BUCKET/data/ --recursive
```

Donde:

- *path/* es la ruta del archivo a la tutorial-dataset carpeta de su dispositivo,
- *DOC-EXAMPLE-BUCKET* es el nombre de tu bucket.

2. Para asegurarse de que los archivos del conjunto de datos se hayan cargado correctamente en la data carpeta, utilice el comando [list](#) de: AWS CLI

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/data/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

Al final de este paso, tendrá un bucket de S3 con su conjunto de datos almacenado dentro de la data carpeta y una metadata carpeta vacía, que almacenará sus metadatos de Amazon Kendra.

Paso 2: Ejecutar un trabajo de análisis de entidades en Amazon Comprehend

Tras almacenar el conjunto de datos de muestra en su bucket de S3, ejecuta un trabajo de análisis de entidades de Amazon Comprehend para extraer las entidades de sus documentos. Estas entidades formarán atributos personalizados de Amazon Kendra y le ayudarán a filtrar los resultados de búsqueda de su índice. Para obtener más información, consulte [Detectar entidades](#).

Temas

- [Ejecución de un trabajo de análisis de entidades de Amazon Comprehend \(p. 583\)](#)

Ejecución de un trabajo de análisis de entidades de Amazon Comprehend

Para extraer entidades de su conjunto de datos, ejecute un trabajo de análisis de entidades de Amazon Comprehend.

Si utiliza la AWS CLI en este paso, primero debe crear y adjuntar una función y una política de AWS IAM para Amazon Comprehend y, a continuación, ejecutar un trabajo de análisis de entidades. Para ejecutar un trabajo de análisis de entidades en los datos de muestra, Amazon Comprehend necesita:

- un rol AWS Identity and Access Management (IAM) que lo reconoce como una entidad de confianza
- una política de AWS IAM adjunta a la función de IAM que le otorga permisos para acceder a su bucket de S3

Para obtener más información, consulte [Cómo funciona Amazon Comprehend con las políticas de IAM y basadas en identidades para Amazon Comprehend](#).

Para ejecutar un trabajo de análisis de entidades de Amazon Comprehend (consola)

1. Abra la consola de Amazon Comprehend en <https://console.aws.amazon.com/comprehend/>.

Important

Asegúrese de estar en la misma región en la que creó su bucket de Amazon S3. Si se encuentra en otra región, elija la AWS región en la que creó su bucket de S3 en el selector de regiones de la barra de navegación superior.

2. Elija Lanzar Amazon Comprehend.
3. En el panel de navegación de la izquierda, seleccione Trabajos de análisis.
4. Seleccione Create job (Crear trabajo).
5. En la sección Configuración del trabajo, haga lo siguiente:
 - a. En Name (Nombre), ingrese **data-entities-analysis**.
 - b. En Tipo de análisis, elija Entidades.
 - c. En Idioma, selecciona inglés.
 - d. Mantenga el cifrado de tareas desactivado.
6. En la sección Datos de entrada, haga lo siguiente:
 - a. En Fuente de datos, seleccione Mis documentos.

- b. Para la ubicación de S3, elija Explorar S3.
 - c. En Elegir recursos, haz clic en el nombre de tu depósito de la lista de depósitos.
 - d. Para Objetos, seleccione el botón de opción data y elija Elegir.
 - e. En Formato de entrada, elija Un documento por archivo.
7. En la sección Datos de salida, haga lo siguiente:
- a. Para la ubicación de S3, elija Examinar en S3 y, a continuación, seleccione la casilla de opciones de su depósito en la lista de depósitos y elija Elegir.
 - b. Mantenga el cifrado desactivado.
8. En la sección Permisos de acceso, haga lo siguiente:
- a. Para el rol de IAM, elija Crear un rol de IAM.
 - b. Para obtener permisos de acceso, elija cubos S3 de entrada y salida.
 - c. En el sufijo Nombre, introduzca**comprehend-role**. Esta función proporciona acceso a su bucket de Amazon S3.
9. Conserva la configuración de VPC predeterminada.
10. Seleccione Create job (Crear trabajo).

Para ejecutar un trabajo de análisis de entidades de Amazon Comprehend () AWS CLI

1. Para crear y adjuntar un rol de IAM para Amazon Comprehend que lo reconozca como una entidad de confianza, haga lo siguiente:
 - a. Guarde la siguiente política de confianza como un archivo JSON llamado **comprehend-trust-policy.json** en un editor de texto de su dispositivo local.

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

- b. Para crear un rol de IAM llamado **comprehend-role** y adjuntarle **comprehend-trust-policy.json** el archivo guardado, utilice el comando [create-role](#):

Linux

```
aws iam create-role \  
    --role-name comprehend-role \  
    --assume-role-policy-document file://path/comprehend-trust-  
    policy.json
```

Donde:

- *path*/ es la ruta del archivo a su **comprehend-trust-policy.json** dispositivo local.

macOS

```
aws iam create-role \
    --role-name comprehend-role \
    --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Donde:

- *path* es la ruta del archivo a su comprehend-trust-policy.json dispositivo local.

Windows

```
aws iam create-role ^
    --role-name comprehend-role ^
    --assume-role-policy-document file://path/comprehend-trust-
policy.json
```

Donde:

- *path* es la ruta del archivo a su comprehend-trust-policy.json dispositivo local.

- Copie el nombre del recurso de Amazon (ARN) en su editor de texto y guárdelo localmente como comprehend-role-arn.

Note

El ARN tiene un formato similar a *arn:aws:iam: :123456789012:role/comprehend-role*. Necesita el ARN como el que guardó comprehend-role-arn para ejecutar el trabajo de análisis de Amazon Comprehend.

- Para crear y adjuntar una política de IAM a su rol de IAM que le otorgue permisos para acceder a su bucket de S3, haga lo siguiente:

- Guarde la siguiente política de confianza como un archivo JSON llamado comprehend-S3-access-policy.json en un editor de texto de su dispositivo local.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::DOC-EXAMPLE-BUCKET/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3>ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::DOC-EXAMPLE-BUCKET"
            ],
            "Effect": "Allow"
        }
    ]
}
```

```
        "Action": [
            "s3:PutObject"
        ],
        "Resource": [
            "arn:aws:s3:::DOC-EXAMPLE-BUCKET/*"
        ],
        "Effect": "Allow"
    }
}
```

- b. Para crear una política de IAM llamada comprehend-S3-access-policy para acceder a su bucket de S3, utilice el comando [create-policy](#):

Linux

```
aws iam create-policy \
    --policy-name comprehend-S3-access-policy \
    --policy-document file://path/comprehend-S3-access-policy.json
```

Donde:

- *path/* es la ruta del archivo a su comprehend-S3-access-policy.json dispositivo local.

macOS

```
aws iam create-policy \
    --policy-name comprehend-S3-access-policy \
    --policy-document file://path/comprehend-S3-access-policy.json
```

Donde:

- *path/* es la ruta del archivo a su comprehend-S3-access-policy.json dispositivo local.

Windows

```
aws iam create-policy ^
    --policy-name comprehend-S3-access-policy ^
    --policy-document file://path/comprehend-S3-access-policy.json
```

Donde:

- *path/* es la ruta del archivo a su comprehend-S3-access-policy.json dispositivo local.

- c. Copie el nombre del recurso de Amazon (ARN) en su editor de texto y guárdelo localmente como comprehend-S3-access-arn.

Note

El ARN tiene un formato similar a *arn:aws:iam: :123456789012:role/comprehend-S3-access-policy*. Necesita el ARN como el que guardó comprehend-S3-access-arn para adjuntarlo comprehend-S3-access-policy a su rol de IAM.

- d. Para adjuntar el comprehend-S3-access-policy a su rol de IAM, utilice el [attach-role-policy](#) comando:

Linux

```
aws iam attach-role-policy \  
    --policy-arn policy-arn \  
    --role-name comprehend-role
```

Donde:

- *policy-arn* es el ARN como el que ha guardado. comprehend-S3-access-arn

macOS

```
aws iam attach-role-policy \  
    --policy-arn policy-arn \  
    --role-name comprehend-role
```

Donde:

- *policy-arn* es el ARN como el que ha guardado. comprehend-S3-access-arn

Windows

```
aws iam attach-role-policy ^  
    --policy-arn policy-arn ^  
    --role-name comprehend-role
```

Donde:

- *policy-arn* es el ARN como el que ha guardado. comprehend-S3-access-arn

3. Para ejecutar un trabajo de análisis de entidades de Amazon Comprehend, utilice el [start-entities-detection-job](#) comando:

Linux

```
aws comprehend start-entities-detection-job \  
    --input-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/ \  
    data/,InputFormat=ONE_DOC_PER_FILE \  
    --output-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/ \  
    --data-access-role-arn role-arn \  
    --job-name data-entities-analysis \  
    --language-code en \  
    --region aws-region
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de su bucket de S3,
- *role-arn* es el ARN como el que has guardado, comprehend-role-arn
- *aws-region* es su AWS región.

macOS

```
aws comprehend start-entities-detection-job \  
    
```

```
--input-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/\ndata,InputFormat=ONE_DOC_PER_FILE \\n\n--output-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/ \\n\n--data-access-role-arn role-arn \\n\n--job-name data-entities-analysis \\n\n--language-code en \\n\n--region aws-region
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de su bucket de S3,
- *role-arn* es el ARN como el que has guardado, comprehend-role-arn
- *aws-region* es su AWS región.

Windows

```
aws comprehend start-entities-detection-job ^\n    --input-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/\ndata,InputFormat=ONE_DOC_PER_FILE ^\n    --output-data-config S3Uri=s3://DOC-EXAMPLE-BUCKET/ ^\n    --data-access-role-arn role-arn ^\n    --job-name data-entities-analysis ^\n    --language-code en ^\n    --region aws-region
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre de su bucket de S3,
- *role-arn* es el ARN como el que has guardado, comprehend-role-arn
- *aws-region* es su AWS región.

4. Copia el análisis de entidades JobId y guárdalo en un editor de texto como comprehend-job-id. Le JobId ayuda a realizar un seguimiento del estado de su trabajo de análisis de entidades.
5. Para realizar un seguimiento del progreso de su trabajo de análisis de entidades, utilice el [describe-entities-detection-job](#) comando:

Linux

```
aws comprehend describe-entities-detection-job \\n\n    --job-id entities-job-id \\n\n    --region aws-region
```

Donde:

- *entities-job-id* es tu salvacióncomprehend-job-id,
- *aws-region* es su AWS región.

macOS

```
aws comprehend describe-entities-detection-job \\n\n    --job-id entities-job-id \\n\n    --region aws-region
```

Donde:

- *entities-job-id* es tu salvacióncomprehend-job-id,

- *aws-region* es su AWS región.

Windows

```
aws comprehend describe-entities-detection-job ^
--job-id entities-job-id ^
--region aws-region
```

Donde:

- *entities-job-id*es tu salvacióncomprehend-job-id,
- *aws-region* es su AWS región.

El cambio a. Puede tardar varios minutos en JobStatus cambiarse aCOMPLETED.

Al final de este paso, Amazon Comprehend almacena los resultados del análisis de entidades como un `output.tar.gz` archivo comprimido dentro de una `output` carpeta generada automáticamente en su bucket de S3. Asegúrese de que el estado de su trabajo de análisis esté completo antes de continuar con el siguiente paso.

Paso 3: Formatear la salida del análisis de entidades como metadatos de Amazon Kendra

Para convertir las entidades extraídas por Amazon Comprehend al formato de metadatos requerido por un índice de Amazon Kendra, ejecute un script de Python 3. Los resultados de la conversión se almacenan en la metadata carpeta de su bucket de Amazon S3.

Para obtener más información sobre el formato y la estructura de los metadatos de Amazon Kendra, consulte [Metadatos de documentos de S3](#).

Temas

- [Descargar y extraer la salida de Amazon Comprehend \(p. 589\)](#)
- [Cargar la salida al bucket de S3 \(p. 592\)](#)
- [Conversión de la salida al formato de metadatos de Amazon Kendra \(p. 593\)](#)
- [Limpiar su depósito de Amazon S3 \(p. 596\)](#)

Descargar y extraer la salida de Amazon Comprehend

Para formatear el resultado del análisis de entidades de Amazon Comprehend, primero debe descargar el archivo de análisis de entidades de Amazon Comprehend y extraer el `output.tar.gz` archivo de análisis de entidades.

Para descargar y extraer el archivo de salida (consola)

1. En el panel de navegación de la consola Amazon Comprehend, vaya a Trabajos de análisis.
2. Elija su trabajo de análisis de entidades*data-entities-analysis*.
3. En Salida, elija el enlace que aparece junto a la ubicación de los datos de salida. Esto lo redirige al `output.tar.gz` archivo de su bucket de S3.
4. En la pestaña Descripción general, selecciona Descargar.

Tip

Los resultados de todos los trabajos de análisis de Amazon Comprehend tienen el mismo nombre. Cambiar el nombre de tu archivo te ayudará a rastrearlo más fácilmente.

5. Descomprime y extrae el archivo Amazon Comprehend descargado en tu dispositivo.

Para descargar y extraer el archivo de salida (AWS CLI)

1. Para acceder al nombre de la carpeta generada automáticamente por Amazon Comprehend en su bucket de S3 que contiene los resultados del trabajo de análisis de entidades, utilice el [describe-entities-detection-job](#) comando:

Linux

```
aws comprehend describe-entities-detection-job \
    --job-id entities-job-id \
    --region aws-region
```

Donde:

- *entities-job-id*es de lo que te comprehend-job-id salvaste[the section called “Paso 2: Detectar entidades” \(p. 583\)](#),
- *aws-region* es su AWS región.

macOS

```
aws comprehend describe-entities-detection-job \
    --job-id entities-job-id \
    --region aws-region
```

Donde:

- *entities-job-id*es de lo que te comprehend-job-id salvaste[the section called “Paso 2: Detectar entidades” \(p. 583\)](#),
- *aws-region* es su AWS región.

Windows

```
aws comprehend describe-entities-detection-job ^
    --job-id entities-job-id ^
    --region aws-region
```

Donde:

- *entities-job-id*es de lo que te comprehend-job-id salvaste[the section called “Paso 2: Detectar entidades” \(p. 583\)](#),
- *aws-region* es su AWS región.

2. Desde el OutputDataConfig objeto de la descripción del trabajo de su entidad, copie y guarde el S3Uri valor como comprehend-S3uri en un editor de texto.

Note

El S3Uri valor tiene un formato similar a `s3://DOC-EXAMPLE-BUCKET /... /output/output.tar.gz`.

3. Para descargar el archivo de salida de entidades, utilice el comando [copy](#):

Linux

```
aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
```

Donde:

- `s3://DOC-EXAMPLE-BUCKET /... /output/output.tar.gz` es el S3Uri valor como el que guardócomprehend-S3uri,
- `path/` es el directorio local donde desea guardar la salida.

macOS

```
aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
```

Donde:

- `s3://DOC-EXAMPLE-BUCKET /... /output/output.tar.gz` es el S3Uri valor como el que guardócomprehend-S3uri,
- `path/` es el directorio local donde desea guardar la salida.

Windows

```
aws s3 cp s3://DOC-EXAMPLE-BUCKET/.../output/output.tar.gz path/output.tar.gz
```

Donde:

- `s3://DOC-EXAMPLE-BUCKET /... /output/output.tar.gz` es el S3Uri valor como el que guardócomprehend-S3uri,
- `path/` es el directorio local donde desea guardar la salida.

4. Para extraer la salida de las entidades, ejecute el siguiente comando en una ventana de terminal:

Linux

```
tar -xf path/output.tar.gz -C path/
```

Donde:

- `path/` es la ruta del `output.tar.gz` archivo descargado en su dispositivo local.

macOS

```
tar -xf path/output.tar.gz -C path/
```

Donde:

- *path/* es la ruta del output.tar.gz archivo descargado en su dispositivo local.

Windows

```
tar -xf path/output.tar.gz -C path/
```

Donde:

- *path/* es la ruta del output.tar.gz archivo descargado en su dispositivo local.

Al final de este paso, deberías tener un archivo en tu dispositivo llamado output con una lista de entidades identificadas en Amazon Comprehend.

Cargar la salida al bucket de S3

Tras descargar y extraer el archivo de análisis de entidades de Amazon Comprehend, cargue el output archivo extraído a su bucket de Amazon S3.

Para cargar el archivo de salida extraído de Amazon Comprehend (consola)

1. Abra la consola de Amazon S3 en <https://console.aws.amazon.com/s3>.
2. En Buckets, haz clic en el nombre de tu bucket y, a continuación, selecciona Cargar.
3. En Archivos y carpetas, selecciona Agregar archivos.
4. En el cuadro de diálogo, navegue hasta output el archivo extraído en el dispositivo, selecciónelo y elija Abrir.
5. Conserve la configuración predeterminada para Destino, Permisos y Propiedades.
6. Seleccione Upload (Cargar).

Para cargar el archivo de salida extraído de Amazon Comprehend () AWS CLI

1. Para subir el output archivo extraído a tu bucket, usa el comando [copy](#):

Linux

```
aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
```

Donde:

- *path/* es la ruta del archivo local al archivo extraído, output
- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

macOS

```
aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
```

Donde:

- *path/* es la ruta del archivo local al archivo extraído, output
- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

Windows

```
aws s3 cp path/output s3://DOC-EXAMPLE-BUCKET/output
```

Donde:

- *path/* es la ruta del archivo local al archivo extraído, output
- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

2. Para asegurarse de que el output archivo se ha cargado correctamente en su bucket de S3, compruebe su contenido mediante el comando [list](#):

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

Conversión de la salida al formato de metadatos de Amazon Kendra

Para convertir la salida de Amazon Comprehend en metadatos de Amazon Kendra, ejecute un script de Python 3. Si utiliza la consola, utilice AWS CloudShell este paso.

Para ejecutar el script de Python 3 (consola)

1. Descarga el archivo [comprimido converter.py.zip](#) en tu dispositivo.
2. Extraiga el archivo Python 3converter.py.
3. Inicie sesión en la [consola AWS de administración](#) y asegúrese de que su AWS región esté configurada en la misma región que su bucket de S3 y su trabajo de análisis de Amazon Comprehend.
4. Elija el AWS CloudShell icono o escriba AWSCloudShell en el cuadro de búsqueda de la barra de navegación superior para iniciar un entorno.

Note

Cuando AWS CloudShell se inicia por primera vez en una nueva ventana del navegador, un panel de bienvenida muestra y enumera las características clave. La consola estará lista para la interacción cuando cierre este panel y aparezca la línea de comandos.

5. Una vez que el terminal esté preparado, seleccione Acciones en el panel de navegación y, a continuación, elija Cargar archivo en el menú.
6. En el cuadro de diálogo que se abre, selecciona Seleccionar archivo y, a continuación, elige el archivo Python 3 descargado converter.py de tu dispositivo. Seleccione Upload (Cargar).
7. En el AWS CloudShell entorno, introduzca el siguiente comando:

```
python3 converter.py
```

8. Cuando la interfaz shell le pida que introduzca el nombre de su bucket de S3, introduzca el nombre de su bucket de S3 y pulse enter.
9. Cuando la interfaz shell le pida que introduzca la ruta completa del archivo de salida de Comprehend, introduzca y pulse enter**output**.
10. Cuando la interfaz shell le pida que introduzca la ruta de archivo completa a su carpeta de metadatos, introduzca **metadata/** y pulse enter.

Important

Para que los metadatos tengan el formato correcto, los valores de entrada de los pasos 8 a 10 deben ser exactos.

Para ejecutar el script de Python 3 (AWS CLI)

1. Para descargar el archivo Python 3converter.py, ejecute el siguiente comando en una ventana de terminal:

Linux

```
curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/samples/converter.py.zip
```

Donde:

- *path/* es la ruta del archivo a la ubicación en la que desea guardar el archivo comprimido.

macOS

```
curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/samples/converter.py.zip
```

Donde:

- *path/* es la ruta del archivo a la ubicación en la que desea guardar el archivo comprimido.

Windows

```
curl -o path/converter.py.zip https://docs.aws.amazon.com/kendra/latest/dg/samples/converter.py.zip
```

Donde:

- *path/* es la ruta del archivo a la ubicación en la que desea guardar el archivo comprimido.
2. Para extraer el archivo Python 3, ejecute el siguiente comando en la ventana del terminal:

Linux

```
unzip path/converter.py.zip -d path/
```

Donde:

- *path/* es la ruta del archivo a tu archivo guardado. converter.py.zip

macOS

```
unzip path/converter.py.zip -d path/
```

Donde:

- *path/* es la ruta del archivo a tu archivo guardado. converter.py.zip

Windows

```
tar -xf path/converter.py.zip -C path/
```

Donde:

- *path/* es la ruta del archivo a tu archivo guardado. converter.py.zip

3. Asegúrese de que Boto3 esté instalado en su dispositivo ejecutando el siguiente comando.

Linux

```
pip3 show boto3
```

macOS

```
pip3 show boto3
```

Windows

```
pip3 show boto3
```

Note

Si no tiene Boto3 instalado, ejecútelo `pip3 install boto3` para instalarlo.

4. Para ejecutar el script de Python 3 para convertir el output archivo, ejecute el siguiente comando.

Linux

```
python path/converter.py
```

Donde:

- *path/* es la ruta del archivo a tu archivo guardado. converter.py.zip

macOS

```
python path/converter.py
```

Donde:

- *path/* es la ruta del archivo a tu archivo guardado. converter.py.zip

Windows

```
python path/converter.py
```

Donde:

- *path/* es la ruta del archivo a tu archivo guardado. converter.py.zip

5. Cuando se AWS CLI le soliciteEnter the name of your S3 bucket, introduzca el nombre de su bucket de S3 y pulse enter.
6. Cuando se AWS CLI le pida que lo hagaEnter the full filepath to your Comprehend output file, introduzca **output** y pulse enter.
7. Cuando se AWS CLI le pida que lo hagaEnter the full filepath to your metadata folder, introduzca **metadata/** y pulse enter.

Important

Para que los metadatos tengan el formato correcto, los valores de entrada de los pasos 5 a 7 deben ser exactos.

Al final de este paso, los metadatos formateados se depositan dentro de la metadata carpeta del bucket de S3.

Limpiar su depósito de Amazon S3

Dado que el índice de Amazon Kendra sincroniza todos los archivos almacenados en un bucket, le recomendamos que limpie su bucket de Amazon S3 para evitar resultados de búsqueda redundantes.

Para limpiar su bucket de Amazon S3 (consola)

1. Abra la consola de Amazon S3 en <https://console.aws.amazon.com/s3>.
2. En Buckets, elija su depósito y, a continuación, seleccione la carpeta de salida del análisis de entidades de Amazon Comprehend, el archivo de análisis de entidades de Amazon Comprehend y el .temp archivo de Amazon Comprehend extraído. **output**
3. En la pestaña Descripción general, selecciona Eliminar.
4. En Eliminar objetos, elija ¿Eliminar objetos permanentemente? e introduzca **permanently delete** en el campo de entrada de texto.
5. Elija Delete objects (Eliminar objetos).

Para limpiar su bucket de Amazon S3 (AWS CLI)

1. Para eliminar todos los archivos y carpetas de su bucket de S3, excepto las carpetas data y, utilice el comando [remove](#) en AWS CLI:

Linux

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude "metadata/*"
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

macOS

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude "metadata/*"
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

Windows

```
aws s3 rm s3://DOC-EXAMPLE-BUCKET/ --recursive --exclude "data/*" --exclude "metadata/*"
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

2. Para asegurarse de que los objetos se eliminaron correctamente del bucket de S3, compruebe su contenido mediante el comando [list](#):

Linux

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

macOS

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

Windows

```
aws s3 ls s3://DOC-EXAMPLE-BUCKET/
```

Donde:

- *DOC-EXAMPLE-BUCKET* es el nombre del bucket de S3.

Al final de este paso, ha convertido el resultado del análisis de entidades de Amazon Comprehend en metadatos de Amazon Kendra. Ya está listo para crear un índice de Amazon Kendra.

Paso 4: Crear un índice de Amazon Kendra e incorporar los metadatos

Para implementar su solución de búsqueda inteligente, debe crear un índice de Amazon Kendra e incorporar sus datos y metadatos de S3 en él.

Antes de añadir metadatos al índice de Amazon Kendra, debe crear campos de índice personalizados que correspondan a los atributos personalizados del documento, que a su vez corresponden a los tipos de entidades de Amazon Comprehend. Amazon Kendra utiliza los campos de índice y los atributos de documentos personalizados que usted crea para buscar y filtrar sus documentos.

Para obtener más información, consulte [Indexación](#) y [Creación de atributos de documentos personalizados](#).

Temas

- [Creación de un índice de Amazon Kendra \(p. 598\)](#)
- [Actualización del rol de IAM para el acceso a Amazon S3 \(p. 604\)](#)
- [Creación de campos de índice de búsqueda personalizados de Amazon Kendra \(p. 606\)](#)
- [Añadir el bucket de Amazon S3 como fuente de datos para el índice \(p. 610\)](#)
- [Sincronización del índice de Amazon Kendra \(p. 613\)](#)

Creación de un índice de Amazon Kendra

Para consultar los documentos fuente, debe crear un índice de Amazon Kendra.

Si utiliza el AWS CLI en este paso, cree y adjunte una función y una política de AWS IAM que permitan a Amazon Kendra acceder a sus CloudWatch registros antes de crear un índice. Para obtener más información, consulte [Requisitos previos](#).

Para crear un índice de Amazon Kendra (consola)

1. Abra la consola de Amazon Kendra en <https://console.aws.amazon.com/kendra/>.

Important

Asegúrese de estar en la misma región en la que creó su trabajo de análisis de entidades de Amazon Comprehend y su bucket de Amazon S3. Si se encuentra en otra región, elija la AWS región en la que creó su bucket de Amazon S3 en el selector de regiones de la barra de navegación superior.

2. Elige Crear un índice.
3. Para ver los detalles del índice en la página Especificar los detalles del índice, haga lo siguiente:
 - a. En Index name (Nombre de índice), ingrese el **kendra-index**.
 - b. Mantenga el campo Descripción en blanco.
 - c. En IAM Role (Rol de IAM), elija Create a New Role (Crear un nuevo rol). Esta función proporciona acceso a su bucket de Amazon S3.
 - d. En Nombre del rol, ingrese **kendra-role**. El rol de IAM tendrá el prefijoAmazonKendra-.
 - e. Conserve la configuración predeterminada para el cifrado y las etiquetas y seleccione Siguiente.
4. Para configurar el control de acceso en la página Configurar el control de acceso de usuario, elija No y, a continuación, elija Siguiente.
5. Para las ediciones de aprovisionamiento, en la página de detalles del aprovisionamiento, elija Edición para desarrolladores y elija Crear.

Para crear un índice de Amazon Kendra () AWS CLI

1. Para crear y adjuntar un rol de IAM para Amazon Kendra que lo reconozca como una entidad de confianza, haga lo siguiente:
 - a. Guarde la siguiente política de confianza como un archivo JSON llamado **kendra-trust-policy.json** en un editor de texto de su dispositivo local.

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

- b. Para crear un rol de IAM llamado **kendra-role** y adjuntarle **kendra-trust-policy.json** el archivo guardado, utilice el comando [create-role](#):

Linux

```
aws iam create-role \  
    --role-name kendra-role \  
    --assume-role-policy-document file://path/kendra-trust-policy.json
```

Donde:

- *path/* es la ruta del archivo a su **kendra-trust-policy.json** dispositivo local.

macOS

```
aws iam create-role \  
    --role-name kendra-role \  
    --assume-role-policy-document file://path/kendra-trust-policy.json
```

Donde:

- *path/* es la ruta del archivo a su **kendra-trust-policy.json** dispositivo local.

Windows

```
aws iam create-role ^
--role-name kendra-role ^
--assume-role-policy-document file://path/kendra-trust-policy.json
```

Donde:

- *path*/ es la ruta del archivo a su kendra-trust-policy.json dispositivo local.
- c. Copie el nombre del recurso de Amazon (ARN) en su editor de texto y guárdelo localmente como kendra-role-arn.

Note

El ARN tiene un formato similar a *arn:aws:iam: :123456789012:role/kendra-role*. Necesita el ARN que guardó kendra-role-arn para ejecutar trabajos de Amazon Kendra.

2. Antes de crear un índice, debe proporcionar permiso kendra-role para escribir en CloudWatch Logs. Para ello, siga los pasos que se describen a continuación:
 - a. Guarde la siguiente política de confianza como un archivo JSON llamado kendra-cloudwatch-policy.json en un editor de texto de su dispositivo local.

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

Sustituya *aws-region* por su AWS región y por su ID de *aws-account-id*cuenta de 12 dígitosAWS.

- b. Para crear una política de IAM para acceder a CloudWatch los registros, utilice el comando [create-policy](#):

Linux

```
aws iam create-policy \  
    --policy-name kendra-cloudwatch-policy \  
    --policy-document file://path/kendra-cloudwatch-policy.json
```

Donde:

- *path*/ es la ruta del archivo a su kendra-cloudwatch-policy.json dispositivo local.

macOS

```
aws iam create-policy \  
    --policy-name kendra-cloudwatch-policy \  
    --policy-document file://path/kendra-cloudwatch-policy.json
```

Donde:

- *path*/ es la ruta del archivo a su kendra-cloudwatch-policy.json dispositivo local.

Windows

```
aws iam create-policy ^  
    --policy-name kendra-cloudwatch-policy ^  
    --policy-document file://path/kendra-cloudwatch-policy.json
```

Donde:

- *path*/ es la ruta del archivo a su kendra-cloudwatch-policy.json dispositivo local.

- c. Copie el nombre del recurso de Amazon (ARN) en su editor de texto y guárdelo localmente comokendra-cloudwatch-arn.

Note

El ARN tiene un formato similar a *arn:aws:iam: :123456789012:role/*. kendra-cloudwatch-policy Necesita el ARN como el que guardó kendra-cloudwatch-arn para adjuntarlo kendra-cloudwatch-policy a su rol de IAM.

- d. Para adjuntar el kendra-cloudwatch-policy a su rol de IAM, utilice el [attach-role-policy](#) comando:

Linux

```
aws iam attach-role-policy \  
    --policy-arn policy-arn \  
    --role-name kendra-role
```

Donde:

- *policy-arn* es tu guardado. kendra-cloudwatch-arn

macOS

```
aws iam attach-role-policy \  
    --policy-arn policy-arn \  
    --role-name kendra-role
```

Donde:

- *policy-arn* es tu guardado. kendra-cloudwatch-arn

Windows

```
aws iam attach-role-policy ^  
    --policy-arn policy-arn ^  
    --role-name kendra-role
```

Donde:

- *policy-arn* es tu guardado. kendra-cloudwatch-arn

3. Para crear un índice, utilice el comando [create-index](#):

Linux

```
aws kendra create-index \  
    --name kendra-index \  
    --edition DEVELOPER_EDITION \  
    --role-arn role-arn \  
    --region aws-region
```

Donde:

- *role-arn* es tu salvación, kendra-role-arn
- *aws-region* es su AWS región.

macOS

```
aws kendra create-index \  
    --name kendra-index \  
    --edition DEVELOPER_EDITION \  
    --role-arn role-arn \  
    --region aws-region
```

Donde:

- *role-arn* es tu salvación, kendra-role-arn
- *aws-region* es su AWS región.

Windows

```
aws kendra create-index ^  
    --name kendra-index ^  
    --edition DEVELOPER_EDITION ^  
    --role-arn role-arn ^
```

```
--region aws-region
```

Donde:

- *role-arn* es tu salvación, `kendra-role-arn`
- *aws-region* es su AWS región.

4. Copia el índice Id y guárdalo en un editor de texto comokendra-index-id. Le Id ayuda a realizar un seguimiento del estado de la creación de su índice.
5. Para realizar un seguimiento del progreso del trabajo de creación de índices, utilice el comando [describe-index](#):

Linux

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

Donde:

- *kendra-index-id*es tu salvación*kendra-index-id*,
- *aws-region* es su AWS región.

macOS

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

Donde:

- *kendra-index-id*es tu salvación*kendra-index-id*,
- *aws-region* es su AWS región.

Windows

```
aws kendra describe-index ^  
  --id kendra-index-id ^  
  --region aws-region
```

Donde:

- *kendra-index-id*es tu salvación*kendra-index-id*,
- *aws-region* es su AWS región.

El proceso de creación del índice tarda, en promedio, 15 minutos, pero puede tardar más. Cuando el estado del índice esté activo, el índice estará listo para usarse. Mientras se crea el índice, puede iniciar el siguiente paso.

Si utiliza el AWS CLI en este paso, cree y adjunte una política de IAM a su rol de Amazon Kendra que otorgue a su índice permisos para acceder a su bucket de S3.

Actualización del rol de IAM para el acceso a Amazon S3

Mientras se crea el índice, debe actualizar su rol de Amazon Kendra IAM para permitir que el índice que creó lea los datos de su bucket de Amazon S3. Para obtener más información, consulte las [funciones de acceso de IAM para Amazon Kendra](#).

Para actualizar su rol de IAM (consola)

1. Abra la consola de IAM en <https://console.aws.amazon.com/iam/>.
2. En el panel de navegación de la izquierda, elija Funciones e introduzca **kendra-role** en el cuadro de búsqueda situado encima del nombre del rol.
3. De las opciones sugeridas, haz clic en **kendra-role**.
4. En Resumen, elija Adjuntar políticas.
5. En Adjuntar permisos, en el cuadro de búsqueda, introduce **S3** y selecciona la casilla situada junto a la **ReadOnlyAccess** política de AmazonS3 entre las opciones sugeridas.
6. Elija Attach policy (Asociar política). En la página de resumen, ahora verá dos políticas adjuntas a la función de IAM.
7. Vuelva a la consola de Amazon Kendra en <https://console.aws.amazon.com/kendra/> y espere a que el estado del índice cambie de Creando a Activo antes de continuar con el siguiente paso.

Para actualizar su rol de IAM () AWS CLI

1. Guarda el siguiente texto en un archivo JSON llamado **kendra-S3-access-policy.json** en un editor de texto de tu dispositivo local.

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Action": [  
                "s3:GetObject"  
            ],  
            "Resource": [  
                "arn:aws:s3:::DOC-EXAMPLE-BUCKET/*"  
            ],  
            "Effect": "Allow"  
        },  
        {  
            "Action": [  
                "s3>ListBucket"  
            ],  
            "Resource": [  
                "arn:aws:s3:::DOC-EXAMPLE-BUCKET"  
            ],  
            "Effect": "Allow"  
        },  
        {  
            "Effect": "Allow",  
            "Action": [  
                "kendra:BatchPutDocument",  
                "kendra:BatchDeleteDocument",  
                "kendra>ListDataSourceSyncJobs"  
            ],  
            "Resource": [  
                "arn:aws:kendra:aws-region:aws-account-id:index/kendra-index-id"  
            ]  
        }  
    ]  
}
```

```
        ]  
    }]  
}
```

Sustituya **DOC-EXAMPLE-BUCKET** por su nombre de bucket de S3, **aws-region por su AWS región, aws-account-id** por su ID de AWS cuenta de 12 dígitos y por su archivo guardado **kendra-index-id** **kendra-index-id**

2. Para crear una política de IAM para acceder a su bucket de S3, utilice el comando [create-policy](#):

Linux

```
aws iam create-policy \  
    --policy-name kendra-S3-access-policy \  
    --policy-document file://path/kendra-S3-access-policy.json
```

Donde:

- **path/** es la ruta del archivo a su **kendra-S3-access-policy.json** dispositivo local.

macOS

```
aws iam create-policy \  
    --policy-name kendra-S3-access-policy \  
    --policy-document file://path/kendra-S3-access-policy.json
```

Donde:

- **path/** es la ruta del archivo a su **kendra-S3-access-policy.json** dispositivo local.

Windows

```
aws iam create-policy ^  
    --policy-name kendra-S3-access-policy ^  
    --policy-document file://path/kendra-S3-access-policy.json
```

Donde:

- **path/** es la ruta del archivo a su **kendra-S3-access-policy.json** dispositivo local.

3. Copie el nombre del recurso de Amazon (ARN) en su editor de texto y guárdelo localmente como **kendra-S3-access-arn**.

Note

El ARN tiene un formato similar a **arn:aws:iam: :123456789012:role/kendra-S3-access-policy**. Necesita el ARN como el que guardó **kendra-S3-access-arn** para adjuntarlo **kendra-S3-access-policy** a su rol de IAM.

4. Para **kendra-S3-access-policy** adjuntar el rol de IAM de Amazon Kendra, utilice el [attach-role-policy](#) comando:

Linux

```
aws iam attach-role-policy \  
    --policy-arn policy-arn \  
    --role-name kendra-role
```

Donde:

- *policy-arn* es tu guardado. kendra-S3-access-arn

macOS

```
aws iam attach-role-policy \
    --policy-arn policy-arn \
    --role-name kendra-role
```

Donde:

- *policy-arn* es tu guardado. kendra-S3-access-arn

Windows

```
aws iam attach-role-policy ^
    --policy-arn policy-arn ^
    --role-name kendra-role
```

Donde:

- *policy-arn* es tu guardado. kendra-S3-access-arn

Creación de campos de índice de búsqueda personalizados de Amazon Kendra

Para preparar a Amazon Kendra para que reconozca sus metadatos como atributos de documentos personalizados, cree campos personalizados correspondientes a los tipos de entidades de Amazon Comprehend. Introduce los siguientes nueve tipos de entidades de Amazon Comprehend como campos personalizados:

- ARTÍCULO_COMERCIAL
- DATE
- EVENT
- LOCATION
- ORGANIZACIÓN
- OTHER
- PERSONA
- CANTIDAD
- TITLE

Important

El índice no reconocerá los tipos de entidades mal escritos.

Para crear campos personalizados para su índice de Amazon Kendra (consola)

1. Abra la consola de Amazon Kendra en <https://console.aws.amazon.com/kendra/>.

2. En la lista de índices, haga clic en `kendra-index`
3. En el panel de navegación de la izquierda, en Administración de datos, elija Definición de faceta.
4. En el menú Campos de índice, selecciona Agregar campo.
5. En el cuadro de diálogo Agregar campo de índice, haga lo siguiente:
 - a. En Nombre de campo, escriba **COMMERCIAL_ITEM**.
 - b. En Tipo de datos, elija Lista de cadenas.
 - c. En Tipos de uso, seleccione Tabla de facetas, Buscable y Visualizable y, a continuación, elija Agregar.
 - d. Repita los pasos a a c para cada tipo de entidad de Amazon Comprehend: COMMERCIAL_ITEM, DATE, EVENT, LOCATION, ORGANIZATION, OTHER, PERSON, QUANTITY, TITLE.

La consola muestra los mensajes de adición de campos correctos. Puede optar por cerrarlos antes de continuar con el siguiente paso.

Para crear campos personalizados para su índice de Amazon Kendra () AWS CLI

1. Guarda el siguiente texto como un archivo JSON llamado `custom-attributes.json` en un editor de texto de tu dispositivo local.

```
[  
  {  
    "Name": "COMMERCIAL_ITEM",  
    "Type": "STRING_LIST_VALUE",  
    "Search": {  
      "Facetable": true,  
      "Searchable": true,  
      "Displayable": true  
    }  
  },  
  {  
    "Name": "DATE",  
    "Type": "STRING_LIST_VALUE",  
    "Search": {  
      "Facetable": true,  
      "Searchable": true,  
      "Displayable": true  
    }  
  },  
  {  
    "Name": "EVENT",  
    "Type": "STRING_LIST_VALUE",  
    "Search": {  
      "Facetable": true,  
      "Searchable": true,  
      "Displayable": true  
    }  
  },  
  {  
    "Name": "LOCATION",  
    "Type": "STRING_LIST_VALUE",  
    "Search": {  
      "Facetable": true,  
      "Searchable": true,  
      "Displayable": true  
    }  
  },  
  {  
    "Name": "ORGANIZATION",  
    "Type": "STRING_LIST_VALUE",  
  }
```

```
"Search": {  
    "Facetable": true,  
    "Searchable": true,  
    "Displayable": true  
},  
{  
    "Name": "OTHER",  
    "Type": "STRING_LIST_VALUE",  
    "Search": {  
        "Facetable": true,  
        "Searchable": true,  
        "Displayable": true  
}  
},  
{  
    "Name": "PERSON",  
    "Type": "STRING_LIST_VALUE",  
    "Search": {  
        "Facetable": true,  
        "Searchable": true,  
        "Displayable": true  
}  
},  
{  
    "Name": "QUANTITY",  
    "Type": "STRING_LIST_VALUE",  
    "Search": {  
        "Facetable": true,  
        "Searchable": true,  
        "Displayable": true  
}  
},  
{  
    "Name": "TITLE",  
    "Type": "STRING_LIST_VALUE",  
    "Search": {  
        "Facetable": true,  
        "Searchable": true,  
        "Displayable": true  
}  
}  
]
```

2. Para crear campos personalizados en el índice, utilice el comando [update-index](#):

Linux

```
aws kendra update-index \  
    --id kendra-index-id \  
    --document-metadata-configuration-updates file://path/custom-  
    attributes.json \  
    --region aws-region
```

Donde:

- *kendra-index-id*es tu salvación*kendra-index-id*,
- *path/* es la ruta del archivo a su *custom-attributes.json* dispositivo local,
- *aws-region* es su AWS región.

macOS

```
aws kendra update-index \
    --id kendra-index-id \
    --document-metadata-configuration-updates file://path/custom-
attributes.json \
    --region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *path/* es la ruta del archivo a su custom-attributes.json dispositivo local,
- *aws-region* es su AWS región.

Windows

```
aws kendra update-index ^
    --id kendra-index-id ^
    --document-metadata-configuration-updates file://path/custom-
attributes.json ^
    --region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *path/* es la ruta del archivo a su custom-attributes.json dispositivo local,
- *aws-region* es su AWS región.

3. Para comprobar que los atributos personalizados se han agregado al índice, utilice el comando [describe-index](#):

Linux

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

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

macOS

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

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

Windows

```
aws kendra describe-index ^
--id kendra-index-id ^
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvación*kendra-index-id*,
- *aws-region* es su AWS región.

Añadir el bucket de Amazon S3 como fuente de datos para el índice

Antes de poder sincronizar el índice, debe conectar su fuente de datos de S3 a él.

Para conectar un bucket de S3 a su índice de Amazon Kendra (consola)

1. Abra la consola de Amazon Kendra en <https://console.aws.amazon.com/kendra/>.
2. En la lista de índices, haga clic en *kendra-index*
3. En el menú de navegación de la izquierda, en Administración de datos, selecciona Fuentes de datos.
4. En la sección Seleccione el tipo de conector de fuente de datos, vaya a Amazon S3 y seleccione Agregar conector.
5. En la página Especificar detalles de la fuente de datos, haga lo siguiente:
 - a. En Nombre y descripción, en Nombre de la fuente de datos, introduzca**S3-data-source**.
 - b. Mantenga la sección Descripción en blanco.
 - c. Conserve la configuración predeterminada para las etiquetas.
 - d. Elija Siguiente.
6. En la página Configurar los ajustes de sincronización, en la sección Alcance de la sincronización, haga lo siguiente:
 - a. En Introducir la ubicación de la fuente de datos, seleccione Examinar S3.
 - b. En Elegir recursos, seleccione su bucket de S3 y, a continuación, elija Elegir.
 - c. En Ubicación de carpeta con prefijos de archivos de metadatos, elija Examinar S3.
 - d. En Elegir recursos, haz clic en el nombre de tu depósito de la lista de depósitos.
 - e. Para Objetos, seleccione la casilla de opción **metadata** y elija Elegir. El campo de ubicación ahora debería decir**metadata/**.
 - f. Mantenga la configuración predeterminada para la ubicación del archivo de configuración de la lista de control de acceso, seleccione la clave de descifrado y la configuración adicional.
7. Para el rol de IAM, en la página Configurar los ajustes de sincronización, seleccione*kendra-role*.
8. En la página Configurar los ajustes de sincronización, en Sincronizar programación de ejecución, en Frecuencia, elija Ejecutar bajo demanda y, a continuación, Siguiente.
9. En la página Revisar y crear, revise sus opciones para los detalles de la fuente de datos y seleccione Agregar fuente de datos.

Para conectar un bucket de S3 a su índice de Amazon Kendra () AWS CLI

- Guarda el siguiente texto como un archivo JSON llamado S3-data-connector.json en un editor de texto de tu dispositivo local.

```
{  
    "S3Configuration":{  
        "BucketName":"DOC-EXAMPLE-BUCKET",  
        "DocumentsMetadataConfiguration":{  
            "S3Prefix":"metadata"  
        }  
    }  
}
```

Sustituya *DOC-EXAMPLE-BUCKET* por el nombre de su bucket de S3.

- Para conectar el bucket de S3 al índice, utilice el [create-data-source](#) comando:

Linux

```
aws kendra create-data-source \  
    --index-id kendra-index-id \  
    --name S3-data-source \  
    --type S3 \  
    --configuration file://path/S3-data-connector.json \  
    --role-arn role-arn \  
    --region aws-region
```

Donde:

- kendra-index-id*es tu salvaciónkendra-index-id,
- path/* es la ruta del archivo a su S3-data-connector.json dispositivo local,
- role-arn* es tu salvación, kendra-role-arn
- aws-region* es su AWS región.

macOS

```
aws kendra create-data-source \  
    --index-id kendra-index-id \  
    --name S3-data-source \  
    --type S3 \  
    --configuration file://path/S3-data-connector.json \  
    --role-arn role-arn \  
    --region aws-region
```

Donde:

- kendra-index-id*es tu salvaciónkendra-index-id,
- path/* es la ruta del archivo a su S3-data-connector.json dispositivo local,
- role-arn* es tu salvación, kendra-role-arn
- aws-region* es su AWS región.

Windows

```
aws kendra create-data-source ^  
    --index-id kendra-index-id ^
```

```
--name S3-data-source ^
--type S3 ^
--configuration file://path/S3-data-connector.json ^
--role-arn role-arn ^
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
 - *path/* es la ruta del archivo a su S3-data-connector.json dispositivo local,
 - *role-arn* es tu salvación, kendra-role-arn
 - *aws-region* es su AWS región.
3. Copia el conector Id y guárdalo en un editor de texto comoS3-connector-id. Le Id ayuda a realizar un seguimiento del estado del proceso de conexión de datos.
 4. Para asegurarse de que la fuente de datos de S3 se ha conectado correctamente, utilice el [describe-data-source](#) comando:

Linux

```
aws kendra describe-data-source \
    --id S3-connector-id \
    --index-id kendra-index-id \
    --region aws-region
```

Donde:

- *S3-Connector-ID* es tu salvación, S3-connector-id
- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

macOS

```
aws kendra describe-data-source \
    --id S3-connector-id \
    --index-id kendra-index-id \
    --region aws-region
```

Donde:

- *S3-Connector-ID* es tu salvación, S3-connector-id
- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

Windows

```
aws kendra describe-data-source ^
    --id S3-connector-id ^
    --index-id kendra-index-id ^
    --region aws-region
```

Donde:

- *S3-Connector-ID* es tu salvación, S3-connector-id

- *kendra-index-id* es tu salvación *kendra-index-id*,
- *aws-region* es su AWS región.

Al final de este paso, la fuente de datos de Amazon S3 se conectará al índice.

Sincronización del índice de Amazon Kendra

Con la fuente de datos de Amazon S3 agregada, ahora puede sincronizar su índice de Amazon Kendra con ella.

Para sincronizar el índice de Amazon Kendra (consola)

1. Abra la consola de Amazon Kendra en <https://console.aws.amazon.com/kendra/>.
2. En la lista de índices, haga clic en *kendra-index*
3. En el menú de navegación de la izquierda, selecciona Fuentes de datos.
4. En Fuentes de datos, seleccione *S3-data-source*.
5. En la barra de navegación superior, selecciona Sincronizar ahora.

Para sincronizar tu índice de Amazon Kendra () AWS CLI

1. Para sincronizar el índice, utilice el comando [start-data-source-sync-job](#):

Linux

```
aws kendra start-data-source-sync-job \
  --id S3-connector-id \
  --index-id kendra-index-id \
  --region aws-region
```

Donde:

- *S3-Connector-ID* es tu salvación, *S3-connector-id*
- *kendra-index-id* es tu salvación *kendra-index-id*,
- *aws-region* es su AWS región.

macOS

```
aws kendra start-data-source-sync-job \
  --id S3-connector-id \
  --index-id kendra-index-id \
  --region aws-region
```

Donde:

- *S3-Connector-ID* es tu salvación, *S3-connector-id*
- *kendra-index-id* es tu salvación *kendra-index-id*,
- *aws-region* es su AWS región.

Windows

```
aws kendra start-data-source-sync-job ^
```

```
--id S3-connector-id ^
--index-id kendra-index-id ^
--region aws-region
```

Donde:

- *S3-Connector-ID* es tu salvación, *S3-connector-id*
- *kendra-index-id* es tu salvación *kendra-index-id*,
- *aws-region* es su AWS región.

2. Para comprobar el estado de la sincronización del índice, utilice el comando [list-data-source-sync-jobs](#):

Linux

```
aws kendra list-data-source-sync-jobs \
--id S3-connector-id \
--index-id kendra-index-id \
--region aws-region
```

Donde:

- *S3-Connector-ID* es tu salvación, *S3-connector-id*
- *kendra-index-id* es tu salvación *kendra-index-id*,
- *aws-region* es su AWS región.

macOS

```
aws kendra list-data-source-sync-jobs \
--id S3-connector-id \
--index-id kendra-index-id \
--region aws-region
```

Donde:

- *S3-Connector-ID* es tu salvación, *S3-connector-id*
- *kendra-index-id* es tu salvación *kendra-index-id*,
- *aws-region* es su AWS región.

Windows

```
aws kendra list-data-source-sync-jobs ^
--id S3-connector-id ^
--index-id kendra-index-id ^
--region aws-region
```

Donde:

- *S3-Connector-ID* es tu salvación, *S3-connector-id*
- *kendra-index-id* es tu salvación *kendra-index-id*,
- *aws-region* es su AWS región.

Al final de este paso, ha creado un índice de Amazon Kendra con capacidad de búsqueda y filtrado para su conjunto de datos.

Paso 5: Consultar el índice de Amazon Kendra

Su índice de Amazon Kendra ya está listo para realizar consultas en lenguaje natural. Cuando busca en su índice, Amazon Kendra utiliza todos los datos y metadatos que ha proporcionado para obtener las respuestas más precisas a la consulta de búsqueda.

Amazon Kendra puede responder a tres tipos de consultas:

- Consultas factoides (preguntas de «quién», «qué», «cuándo» o «dónde»)
- Consultas descriptivas (preguntas de «cómo»)
- Búsquedas de palabras clave (preguntas cuya intención y alcance no están claros)

Temas

- [Consultar su índice de Amazon Kendra \(p. 615\)](#)
- [Filtrar los resultados de búsqueda \(p. 619\)](#)

Consultar su índice de Amazon Kendra

Puede consultar su índice de Amazon Kendra mediante preguntas que correspondan a los tres tipos de consultas que admite Amazon Kendra. Para obtener más información, consulte [Consultas](#).

Las preguntas de ejemplo de esta sección se han elegido en función del conjunto de datos de muestra.

Para consultar su índice de Amazon Kendra (consola)

1. Abra la consola de Amazon Kendra en <https://console.aws.amazon.com/kendra/>.
2. En la lista de índices, haga clic en `kendra-index`
3. En el menú de navegación de la izquierda, selecciona la opción para buscar en tu índice.
4. Para ejecutar un ejemplo de consulta factoid, introduzca **Who is Lewis Hamilton?** en el cuadro de búsqueda y pulse enter.

El primer resultado devuelto es la respuesta sugerida por Amazon Kendra, junto con el archivo de datos que contiene la respuesta. El resto de los resultados forman el conjunto de documentos recomendados.

The screenshot shows the Amazon Kendra console interface. At the top, there is a search bar with the query "Who is Lewis Hamilton?". Below the search bar, it says "Test query with user name or groups" and "1-8 of 8 results". A section titled "Amazon Kendra suggested answers" displays a result for "Formula One driver". The result includes a snippet from CNN: "7d87db6157b9a3142a96dd6f4a13f85b555c4f24 Formula One driver (CNN) -- Formula One driver Lewis Hamilton has become the latest high-profile British sports star to regret a hastily dashed off tweet after lashing out at McLaren teammate Jenson Button on Twitter. Hamilton accused fellow Briton Button of "unfollowing" him -- not subscribing to his tweets -- on the micro-blogging site, before later discovering his colleague had never followed him. The tweets were sent just hours after the conclusion of the Japanese Grand Prix, where Button finished one place above Hamilton in fourth position. The 2008 world champion Hamilton will leave McLaren at the end of the 2012 season to join German team Mercedes in a three-year deal." Below this, there are two more snippets with similar content, each with a link to the full article on AWS S3.

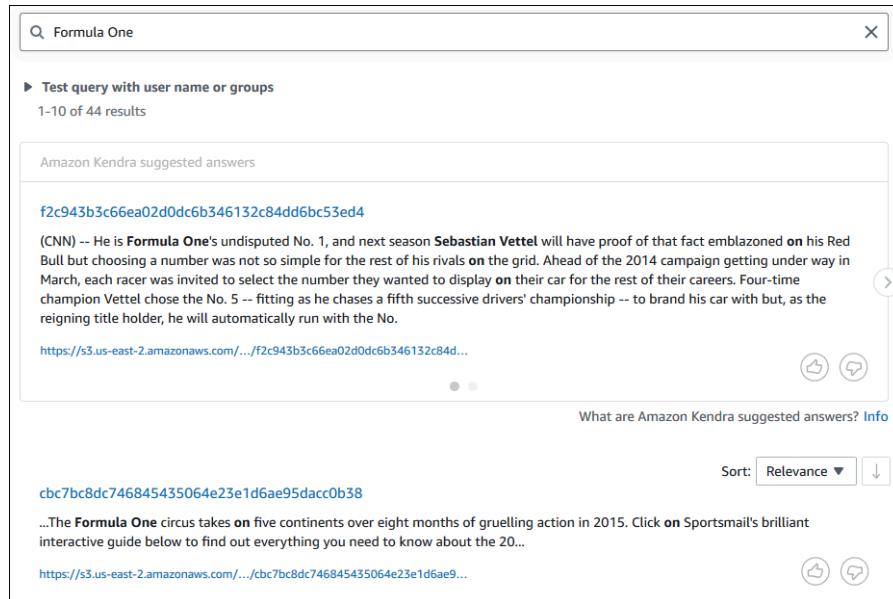
- Para ejecutar una consulta descriptiva, introduzca **How does Formula One work?** en el cuadro de búsqueda y pulse enter.

Verás otro resultado devuelto por la consola de Amazon Kendra, esta vez con la frase correspondiente resaltada.

The screenshot shows the Amazon Kendra console interface. At the top, there is a search bar with the query "How does Formula One work?". Below the search bar, it says "Test query with user name or groups" and "1-10 of 51 results". A section titled "Amazon Kendra suggested answers" displays a result for "Formula One". The result includes a snippet from CNN: "cbc7bc8dc746845435064e23e1d6ae95dacc0b38 ...The Formula One circus takes on five continents over eight months of gruelling action in 2015. Click on Sportsmail's brilliant interactive guide below to find out everything you need to know about the 20...". Below this, there are two more snippets with similar content, each with a link to the full article on AWS S3.

- Para ejecutar una búsqueda por palabra clave, introduzca **Formula One** en el cuadro de búsqueda y pulse enter.

Verá otro resultado devuelto por la consola de Amazon Kendra, seguido de los resultados de todas las demás menciones de la frase en el conjunto de datos.



Para consultar su índice de Amazon Kendra () AWS CLI

1. Para ejecutar una consulta de factoid de ejemplo, utilice el comando [query](#):

Linux

```
aws kendra query \
  --index-id kendra-index-id \
  --query-text "Who is Lewis Hamilton?" \
  --region aws-region
```

Donde:

- *kendra-index-ides* tu salvación*kendra-index-id*,
- *aws-region* es su AWS región.

macOS

```
aws kendra query \
  --index-id kendra-index-id \
  --query-text "Who is Lewis Hamilton?" \
  --region aws-region
```

Donde:

- *kendra-index-ides* tu salvación*kendra-index-id*,
- *aws-region* es su AWS región.

Windows

```
aws kendra query ^
  --index-id kendra-index-id ^
  --query-text "Who is Lewis Hamilton?" ^
```

```
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

AWS CLIMuestra los resultados de la consulta.

2. Para ejecutar un ejemplo de consulta descriptiva, utilice el comando [query](#):

Linux

```
aws kendra query \  
    --index-id kendra-index-id \  
    --query-text "How does Formula One work?" \  
    --region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

macOS

```
aws kendra query \  
    --index-id kendra-index-id \  
    --query-text "How does Formula One work?" \  
    --region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

Windows

```
aws kendra query ^  
    --index-id kendra-index-id ^  
    --query-text "How does Formula One work?" ^  
    --region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

AWS CLIMuestra los resultados de la consulta.

3. Para ejecutar un ejemplo de búsqueda por palabra clave, utilice el comando [query](#):

Linux

```
aws kendra query \  
    --index-id kendra-index-id \  
    --query-text "Formula One"
```

```
--query-text "Formula One" \
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

macOS

```
aws kendra query \
--index-id kendra-index-id \
--query-text "Formula One" \
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

Windows

```
aws kendra query ^
--index-id kendra-index-id ^
--query-text "Formula One" ^
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

AWS CLI Muestra las respuestas devueltas a la consulta.

Filtrar los resultados de búsqueda

Puede filtrar y ordenar los resultados de la búsqueda mediante atributos de documentos personalizados en la consola de Amazon Kendra. Para obtener más información sobre cómo Amazon Kendra procesa las consultas, consulte [Filtrar consultas](#).

Para filtrar los resultados de la búsqueda (Consola)

1. Abra la consola de Amazon Kendra en <https://console.aws.amazon.com/kendra/>.
2. En la lista de índices, haga clic en *kendra-index*
3. En el menú de navegación de la izquierda, selecciona la opción para buscar en tu índice.
4. En el cuadro de búsqueda, introduzca **Soccer matches** como consulta y pulse enter.
5. En el menú de navegación de la izquierda, selecciona Filtrar resultados de búsqueda para ver una lista de facetas que puedes usar para filtrar la búsqueda.
6. Selecciona la casilla «Liga de Campeones» situada debajo del subtítulo EVENTO para ver los resultados de tu búsqueda filtrados únicamente por los resultados que contengan «Liga de Campeones».

Amazon Kendra Guía para desarrolladores

Filtrar los resultados de búsqueda

The screenshot shows the Amazon Kendra search interface. At the top, there is a search bar with the query "Soccer matches". Below the search bar, there is a section titled "Test query with user name or groups" with the note "1-4 of 4 results". A "Filter search results" dropdown is open, showing categories like LOCATION, OTHER, ORGANIZATION, DATE, PERSON, QUANTITY, TITLE, and EVENT. Under the EVENT category, "Champions League" is selected. On the right side, the search results are displayed in a list format. The first result is a suggested answer: "7e5db27742008942b2f9cf6ac41826f86148d1f". It contains a snippet of text about a soccer match and a link to the full article. Below this, there are "Like" and "Comment" icons. The second result is another suggested answer with a snippet and a link, also with "Like" and "Comment" icons. The third result is a link to a page about a footballer's career, with "Like" and "Comment" icons. The fourth result is a link to a page about a footballer's battle with betting, with "Like" and "Comment" icons. At the bottom of the results, there is a "Sort: Relevance" dropdown and a "What are Amazon Kendra suggested answers? Info" link.

Para filtrar los resultados de búsqueda (AWS CLI)

1. Para ver las entidades de un tipo específico (por ejemplo `EVENT`) que están disponibles para una búsqueda, utilice el comando [query](#):

Linux

```
aws kendra query \
--index-id kendra-index-id \
--query-text "Soccer matches" \
--facets '[{"DocumentAttributeKey": "EVENT"}]' \
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvación `kendra-index-id`,
- *aws-region* es su AWS región.

macOS

```
aws kendra query \
--index-id kendra-index-id \
--query-text "Soccer matches" \
--facets '[{"DocumentAttributeKey": "EVENT"}]' \
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvación*kendra-index-id*,
- *aws-region* es su AWS región.

Windows

```
aws kendra query ^
--index-id kendra-index-id ^
--query-text "Soccer matches" ^
--facets '[{"DocumentAttributeKey": "EVENT"}]' ^
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvación*kendra-index-id*,
- *aws-region* es su AWS región.

AWS CLIMuestra los resultados de la búsqueda. Para obtener una lista de tipos de facetasEVENT, navegue hasta la sección «FacetResults» del AWS CLI resultado para ver una lista de facetas filtrables con sus recuentos. Por ejemplo, una de las facetas es la «Liga de Campeones».

Note

En su lugarEVENT, puede elegir cualquiera de los campos de índice que creó [the section called “Creación de un índice de Amazon Kendra” \(p. 598\)](#) para el DocumentAttributeKey valor.

2. Para realizar la misma búsqueda pero filtrar solo por los resultados que contengan «Liga de Campeones», utilice el comando [query](#):

Linux

```
aws kendra query \
--index-id kendra-index-id \
--query-text "Soccer matches" \
--attribute-filter '{"ContainsAny":{"Key": "EVENT", "Value": \
["StringListValue":["Champions League"]]} }' \
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvación*kendra-index-id*,
- *aws-region* es su AWS región.

macOS

```
aws kendra query \
--index-id kendra-index-id \
--query-text "Soccer matches" \
--attribute-filter '{"ContainsAny":{"Key": "EVENT", "Value": \
["StringListValue":["Champions League"]]} }' \
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvación*kendra-index-id*,

- *aws-region* es su AWS región.

Windows

```
aws kendra query ^
--index-id kendra-index-id ^
--query-text "Soccer matches" ^
--attribute-filter '[{"ContainsAny":{"Key": "EVENT", "Value": [
{"StringListValue": ["Champions League"]}}}' ^
--region aws-region
```

Donde:

- *kendra-index-id*es tu salvaciónkendra-index-id,
- *aws-region* es su AWS región.

AWS CLIMuestra los resultados de búsqueda filtrados.

Paso 6: Limpieza

Limpiar tus archivos

Para dejar de incurrir en cargos en tu AWS cuenta después de completar este tutorial, puedes seguir los siguientes pasos:

1. Elimine su bucket de Amazon S3

Para obtener información sobre cómo eliminar un depósito, consulte [Eliminar un depósito](#).

2. Eliminar el índice de Amazon Kendra

Para obtener información sobre cómo eliminar un índice de Amazon Kendra, consulte [Eliminar un índice](#).

3. Eliminar **converter.py**

- Para la consola: vaya a [AWS CloudShell](#) asegúrese de que la región esté configurada como su AWS región. Una vez que se haya cargado el shell de bash, escriba el siguiente comando en el entorno y presione enter.

```
rm converter.py
```

- Para AWS CLI: Ejecute el siguiente comando en una ventana de terminal.

Linux

```
rm file/converter.py
```

Donde:

- *file*/ es la ruta del archivo a su converter.py dispositivo local.

macOS

```
rm file/converter.py
```

Donde:

- *file/* es la ruta del archivo a su converter.py dispositivo local.

Windows

```
rm file/converter.py
```

Donde:

- *file/* es la ruta del archivo a su converter.py dispositivo local.

Más información

Para obtener más información sobre la integración de Amazon Kendra en su flujo de trabajo, consulte las siguientes publicaciones de blog:

- [Etiquetado de metadatos de contenido para una búsqueda mejorada](#)
- [Cree una solución de búsqueda inteligente con enriquecimiento de contenido automatizado](#)

Para obtener más información sobre Amazon Comprehend, consulte la Guía para [desarrolladores de Amazon Comprehend](#).

Monitorización y registro para Amazon Kendra

Temas

- [Supervisión del índice \(consola\) \(p. 624\)](#)
- [Registro de llamadas a la API de Amazon Kendra con AWS CloudTrail registros \(p. 627\)](#)
- [Registro de llamadas a la API de clasificación inteligente de Amazon Kendra con AWS CloudTrail registros \(p. 629\)](#)
- [Monitorización de Amazon Kendra con Amazon CloudWatch \(p. 631\)](#)
- [Supervisión de Amazon Kendra con Amazon Logs CloudWatch \(p. 636\)](#)

Supervisión del índice (consola)

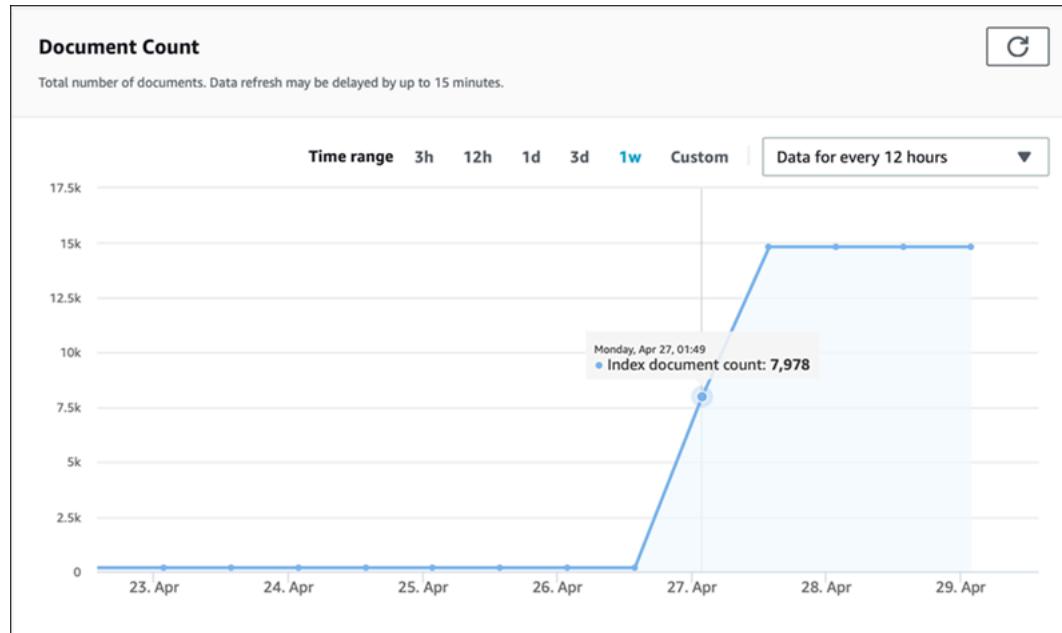
Utilice la consola de Amazon Kendra para supervisar el estado de los índices y las fuentes de datos. Puede utilizar esta información para realizar un seguimiento de los requisitos de tamaño y almacenamiento del índice y para supervisar el progreso y el éxito de la sincronización entre el índice y las fuentes de datos.

Para ver las métricas del índice (consola)

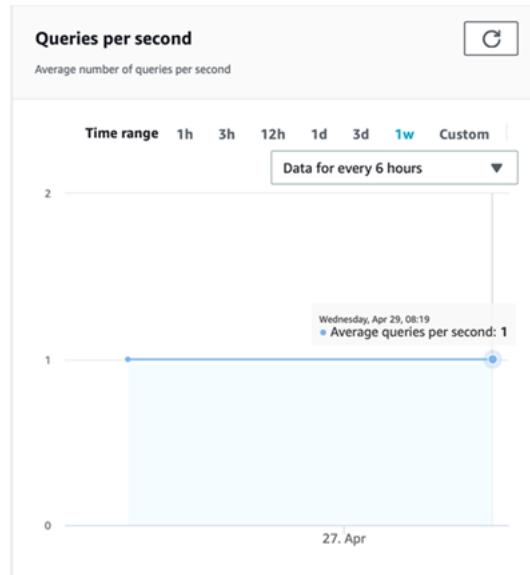
1. Inicie sesión en la consola de Amazon Kendra AWS Management Console y ábrala en <https://console.aws.amazon.com/kendra/home>.
2. En la lista de índices, seleccione el índice que desee ver.
3. Desplázate por la pantalla para ver las métricas del índice.

Puedes ver las siguientes métricas sobre tu índice.

- Recuento de documentos: el número total de documentos indexados. Esto incluye todos los documentos de todas las fuentes de datos. Usa esta métrica para determinar si necesitas comprar más o menos unidades de almacenamiento para tu índice.



- Consultas por segundo: el número de consultas de índice que se solicitan cada segundo. Usa esta métrica para determinar si necesitas comprar más o menos unidades de consulta para tu índice.



Para supervisar el progreso y el éxito de la sincronización entre el índice y una fuente de datos, utilice la consola de Amazon Kendra. Utilice esta información para ayudar a determinar el estado de la fuente de datos.

Para ver las métricas de sincronización (consola)

1. Inicie sesión en la consola de Amazon Kendra AWS Management Console y ábrala en <https://console.aws.amazon.com/kendra/home>.
2. En la lista de índices, elija el índice para el que desea ver las métricas de sincronización.
3. En el menú de la izquierda, selecciona Fuentes de datos.
4. En la lista de fuentes de datos, elija la fuente de datos que desee ver.

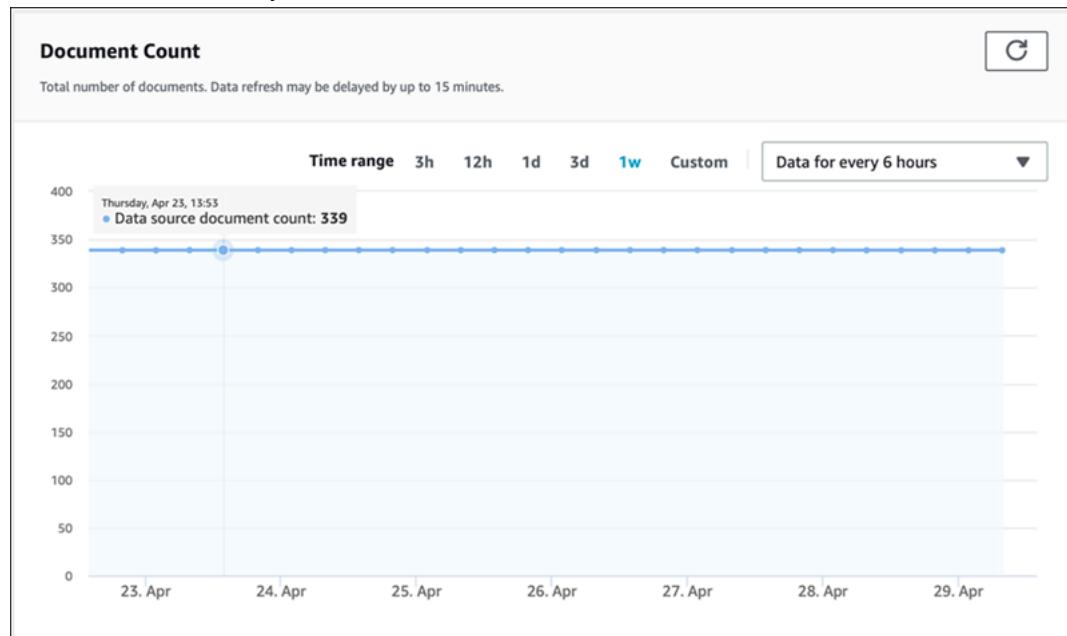
- Desplázate por la pantalla para ver las métricas de ejecución de la sincronización.

Puede ver la siguiente información.

- Historial de ejecución de sincronización: estadísticas sobre la ejecución de la sincronización, incluidas la hora de inicio y finalización, el número de documentos añadidos, eliminados y con errores. Si se produce un error en la ejecución de la sincronización, hay un enlace a CloudWatch los registros con más información. Elija el ícono de configuración en la esquina superior izquierda para cambiar las columnas que se muestran en el historial. Utilice esta información para determinar el estado general de su fuente de datos.

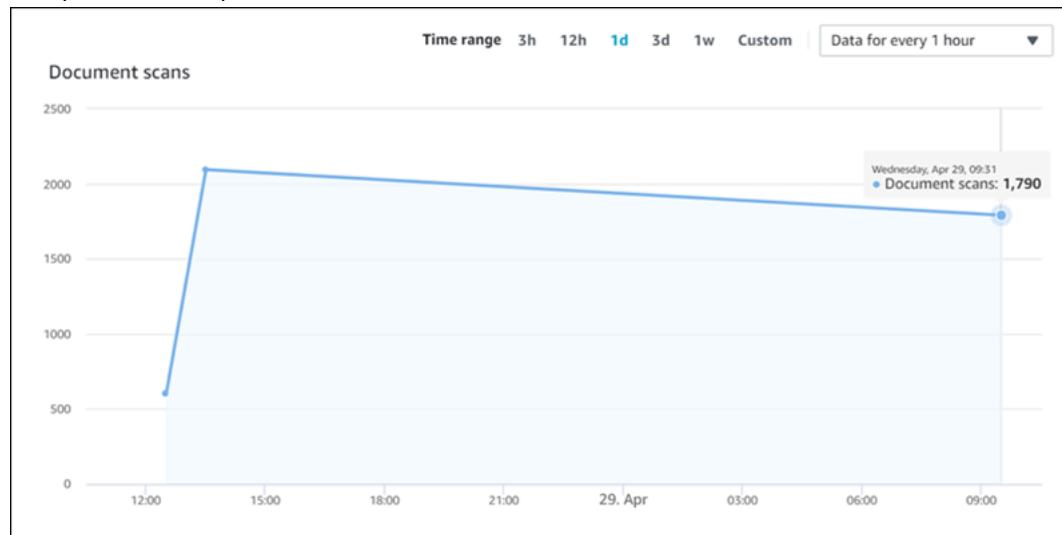
Sync run history (5)						
Status / Summary	Start time	End time	Added / Modified	Deleted	Failed	Details
Syncing - indexing	Apr 29, 2020, 9:53 AM PDT	Apr 29, 2020, 9:54 AM PDT	0	0	0	View in CloudWatch
Succeeded	Apr 28, 2020, 1:35 PM PDT	Apr 28, 2020, 1:37 PM PDT	1484	0	2	Service is operating normally
Succeeded	Apr 28, 2020, 1:32 PM PDT	Apr 28, 2020, 1:32 PM PDT	0	0	0	Service is operating normally
Succeeded	Apr 28, 2020, 1:05 PM PDT	Apr 28, 2020, 1:06 PM PDT	5	0	0	Service is operating normally
Succeeded	Apr 28, 2020, 1:05 PM PDT	Apr 28, 2020, 1:05 PM PDT	298	0	1	Service is operating normally

- Recuento de documentos: el número total de documentos indexados desde esta fuente de datos. Es el total de todos los documentos añadidos a la fuente de datos menos el total de todos los documentos eliminados de la fuente de datos. Utilice esta información para determinar cuántos documentos de esta fuente de datos se incluyen en el índice.

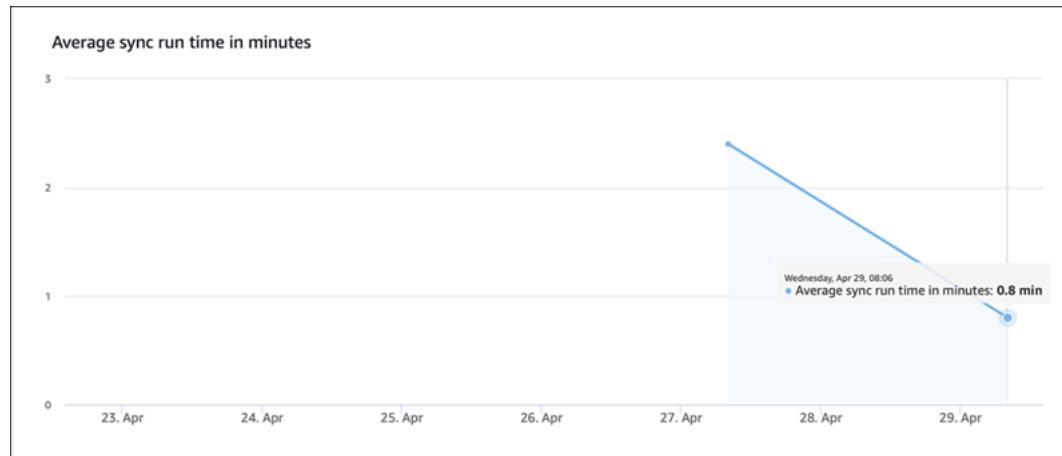


- Escaneos de documentos: el número total de documentos escaneados durante la ejecución de la sincronización. Esto incluye todos los documentos de la fuente de datos, incluidos los agregados,

actualizados, eliminados o sin cambios. Utilice esta información para determinar si Amazon Kendra está escaneando todos los documentos de la fuente de datos. La cantidad de documentos escaneados afecta al importe cobrado por el servicio.



- Tiempo medio de ejecución de la sincronización en minutos: el tiempo promedio que tarda en completarse una ejecución de sincronización. El tiempo que se tarda en sincronizar una fuente de datos afecta al importe que se cobra por el servicio.



Registro de llamadas a la API de Amazon Kendra con AWS CloudTrail registros

Amazon Kendra está integrado con AWS CloudTrail un servicio que proporciona un registro de las acciones realizadas por un usuario, un rol o un AWS servicio en Amazon Kendra. CloudTrail captura todas las llamadas a la API de Amazon Kendra como eventos, incluidas las llamadas desde la consola de Amazon Kendra y las llamadas de código a las API de Amazon Kendra. Si crea una ruta, puede activar la entrega continua de CloudTrail eventos a un bucket de Amazon S3, incluidos los eventos de Amazon Kendra. Si no configura un registro de seguimiento, puede ver los eventos más recientes en la consola de CloudTrail en el Event history (Historial de eventos). Con la información recopilada por CloudTrail, puede determinar la solicitud que se realizó a Amazon Kendra, la dirección IP desde la que se realizó la solicitud, quién hizo la solicitud, cuándo se hizo y detalles adicionales.

Para obtener más información CloudTrail, incluido cómo configurarlo y activarlo, consulte la [Guía del AWS CloudTrail usuario](#).

Información sobre Amazon Kendra en CloudTrail

CloudTrail se activa en su AWS cuenta al crear la cuenta. Cuando se produce una actividad en Amazon Kendra, esa actividad se registra en un CloudTrail evento junto con otros eventos de AWS servicio en el historial de CloudTrail eventos. Puede ver, buscar y descargar los últimos eventos de la cuenta de AWS. Para obtener más información, consulte [Visualización de eventos con el historial de eventos de CloudTrail](#).

Para obtener un registro continuo de los eventos de tu AWS cuenta, incluidos los eventos de Amazon Kendra, crea una ruta. Un registro es una configuración que permite CloudTrail entregar eventos como archivos de registro a un bucket de S3 especificado. De manera predeterminada, cuando se crea un registro de seguimiento en la consola, el registro de seguimiento se aplica a todas las regiones de AWS. El registro de seguimiento registra los eventos de todas las regiones de la partición de AWS y envía los archivos de registro al bucket de S3 especificado. También puede configurar otros servicios de AWS para analizar y actuar en función de los datos de eventos recopilados en los registros de CloudTrail. Para obtener más información, consulte:

- [Introducción a la creación de registros de seguimiento](#)
- [Servicios e integraciones compatibles con CloudTrail](#)
- [Configuración de notificaciones de Amazon SNS para CloudTrail](#)
- [Recibir archivos de registro de CloudTrail de varias regiones](#) y [Recepción de archivos de registro de CloudTrail de varias cuentas](#)

CloudTrail registra todas las acciones de Amazon Kendra, que están documentadas en la [referencia de la API](#). Por ejemplo, las llamadas a las operaciones `CreateIndex`, `CreateDataSource` y `Query` generan entradas en los archivos de registro de CloudTrail.

Cada entrada de registro o evento contiene información sobre quién generó la solicitud. Para obtener más información, consulte el [elemento `userIdentity` de CloudTrail](#).

Ejemplo: entradas del archivo de registro de Amazon Kendra

Un registro es una configuración que permite la entrega de eventos como archivos de registro a un bucket de S3 especificado. Los archivos de registro de CloudTrail contienen una o varias entradas de registro. Un evento representa una solicitud específica realizada desde un origen y contiene información sobre la acción solicitada, la fecha y la hora de la acción, los parámetros de la solicitud, etc. Los archivos de log de CloudTrail no representan un seguimiento de la pila ordenado de las llamadas públicas al API, por lo que no aparecen en ningún orden específico.

Las llamadas a la `Query` operación crean la siguiente entrada.

```
{  
    "eventVersion": "1.05",  
    "userIdentity": {  
        "type": "AssumedRole | FederatedUser | IAMUser | Root | SAMLUser | WebIdentityUser",  
        "principalId": "principal ID",  
        "arn": "ARN",  
        "accountId": "account ID",  
        "accessKeyId": "access key ID",  
        "sessionContext": {  
            "sessionIssuer": {  
                "type": "AWS",  
                "principalId": "principal ID",  
                "arn": "ARN",  
                "accountId": "account ID",  
                "accessKeyId": "access key ID",  
                "sessionName": "session name",  
                "sessionDuration": 3600  
            }  
        }  
    }  
}
```

```
        "type": "Role",
        "principalId": "principal Id",
        "arn": "ARN",
        "accountId": "account ID",
        "userName": "user name"
    },
    "webIdFederationData": {

    },
    "attributes": {
        "mfaAuthenticated": false,
        "creationDate": "timestamp"
    }
},
"eventTime": "timestamp",
"eventSource": "kendra.amazonaws.com",
"eventName": "Query",
"awsRegion": "region",
"sourceIPAddress": "source IP address",
"userAgent": "user agent",
"requestParameters": {
    "indexId": "index ID"
},
"responseElements": null,
"requestID": "request ID",
"eventID": "event ID",
"eventType": "AwsApiCall",
"recipientAccountId": "account ID"
},
```

Registro de llamadas a la API de clasificación inteligente de Amazon Kendra con AWS CloudTrail registros

Amazon Kendra Intelligent Ranking está integrado con AWS CloudTrail un servicio que proporciona un registro de las acciones realizadas por un usuario, rol o AWS servicio en Amazon Kendra Intelligent Ranking. CloudTrail captura todas las llamadas a la API de Amazon Kendra Intelligent Ranking como eventos, incluidas las llamadas de código a las API de Amazon Kendra Intelligent Ranking. Si crea una ruta, puede activar la entrega continua de CloudTrail eventos a un bucket de Amazon S3, incluidos los eventos de Amazon Kendra Intelligent Ranking. Si no configura un registro de seguimiento, puede ver los eventos más recientes en la consola de CloudTrail en el Event history (Historial de eventos). Con la información recopilada por CloudTrail, puede determinar la solicitud que se realizó a Amazon Kendra Intelligent Ranking, la dirección IP desde la que se realizó la solicitud, quién la hizo, cuándo se realizó y detalles adicionales.

Para obtener más información sobre CloudTrail, incluido cómo configurarlo y activarlo, consulte la [Guía del AWS CloudTrail usuario](#).

Información sobre la clasificación inteligente de Amazon Kendra en CloudTrail

CloudTrail se activa en su AWS cuenta al crear la cuenta. Cuando se produce actividad en Amazon Kendra Intelligent Ranking, esa actividad se registra en un CloudTrail evento junto con otros eventos AWS de servicio del historial de CloudTrail eventos. Puede ver, buscar y descargar los últimos eventos de la cuenta

de AWS. Para obtener más información, consulte [Visualización de eventos con el historial de eventos de CloudTrail](#).

Para obtener un registro continuo de los eventos de su AWS cuenta, incluidos los eventos de la Clasificación Inteligente de Amazon Kendra, cree un registro. Un registro es una configuración que permite CloudTrail entregar eventos como archivos de registro a un bucket de S3 especificado. De manera predeterminada, cuando se crea un registro de seguimiento en la consola, el registro de seguimiento se aplica a todas las regiones de AWS. El registro de seguimiento registra los eventos de todas las regiones de la partición de AWS y envía los archivos de registro al bucket de S3 especificado. También puede configurar otros servicios de AWS para analizar y actuar en función de los datos de eventos recopilados en los registros de CloudTrail. Para obtener más información, consulte:

- [Introducción a la creación de registros de seguimiento](#)
- [Servicios e integraciones compatibles con CloudTrail](#)
- [Configuración de notificaciones de Amazon SNS para CloudTrail](#)
- [Recibir archivos de registro de CloudTrail de varias regiones](#) y [Recepción de archivos de registro de CloudTrail de varias cuentas](#)

CloudTrail registra todas las acciones de clasificación inteligente de Amazon Kendra, que se documentan en la [referencia de la API](#). Por ejemplo, las llamadas a `CreateRescoreExecutionPlan` generan entradas en los archivos de CloudTrail registro.

Cada entrada de registro o evento contiene información sobre quién generó la solicitud. Para obtener más información, consulte el [elemento userIdentity de CloudTrail](#).

Ejemplo: entradas del archivo de registro de Amazon Kendra Intelligent Ranking

Un registro es una configuración que permite la entrega de eventos como archivos de registro a un bucket de S3 especificado. Los archivos de registro de CloudTrail contienen una o varias entradas de registro. Un evento representa una solicitud específica realizada desde un origen y contiene información sobre la acción solicitada, la fecha y la hora de la acción, los parámetros de la solicitud, etc. Los archivos de log de CloudTrail no representan un seguimiento de la pila ordenado de las llamadas públicas al API, por lo que no aparecen en ningún orden específico.

Las llamadas a la `CreateRescoreExecutionPlan` operación crean la siguiente entrada.

```
{  
    "eventVersion": "1.08",  
    "userIdentity": {  
        "type": "AssumedRole",  
        "principalId": "principal ID",  
        "arn": "ARN",  
        "accountId": "account ID",  
        "accessKeyId": "access key ID",  
        "sessionContext": {  
            "sessionIssuer": {  
                "type": "Role",  
                "principalId": "principal ID",  
                "arn": "ARN",  
                "accountId": "account ID",  
                "userName": "user name"  
            },  
            "webIdFederationData": {},  
            "attributes": {  
                "creationDate": "yyyy-mm-ddThh:mm:ssZ",  
                "mfaAuthenticated": "false"  
            }  
        }  
    }  
}
```

```
        }
    },
    "eventTime": "yyyy-mm-ddThh:mm:ssZ",
    "eventSource": "kendra-ranking.amazonaws.com",
    "eventName": "CreateRescoreExecutionPlan",
    "awsRegion": "region",
    "sourceIPAddress": "source IP address",
    "userAgent": "user agent",
    "requestParameters": {
        "name": "name",
        "description": "description",
        "clientToken": "client token"
    },
    "responseElements": {
        "id": "rescore execution plan ID",
        "arn": "rescore execution plan ARN"
    },
    "requestID": "request ID",
    "eventID": "event ID",
    "readOnly": false,
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "account ID",
    "eventCategory": "Management",
    "tlsDetails": {
        "tlsVersion": "TLS version",
        "cipherSuite": "cipher suite",
        "clientProvidedHostHeader": "kendra-ranking.[region].api.aws"
    }
}
```

Monitorización de Amazon Kendra con Amazon CloudWatch

Para hacer un seguimiento del estado de tus índices, usa AmazonCloudWatch. ConCloudWatch, puedes obtener métricas para la sincronización de documentos para su índice. También puedes configurar alarmas de CloudWatch para recibir una notificación cuando una o varias métricas superen el umbral que definas. Por ejemplo, puedes supervisar el número de documentos enviados para su indexación o el número de documentos que no se han indexado.

Debes disponer de los CloudWatch permisos adecuados para supervisar Amazon Kendra. CloudWatch para obtener más información, consulte [Autenticación y control de acceso para Amazon CloudWatch](#) en la Guía del CloudWatch usuario de Amazon.

Visualización de las métricas de Amazon Kendra

Consulta las métricas de Amazon Kendra mediante la CloudWatch consola.

Para ver las métricas (consola de CloudWatch)

1. Inicie sesión en la CloudWatch consola AWS Management Console y ábrala en <https://console.aws.amazon.com/cloudwatch/>.
2. Elija Métricas, elija Todas las métricas y, a continuación, elija Kendra.
3. Elija la dimensión, un nombre de métrica y, a continuación, Add to graph (Añadir al gráfico).
4. Elija un valor para el intervalo de fechas. El recuento de las métricas del intervalo de fechas seleccionado se muestra en el gráfico.

Creación de una alarma

Una CloudWatch alarma monitorea una sola métrica durante un período de tiempo específico y realiza una o más acciones: enviar una notificación a una política superior de Amazon Simple Notification Service (Amazon SNS) o a una política de escalado automático. Las acciones o acciones se basan en el valor de la métrica en relación con un umbral determinado durante los períodos de tiempo que especifique. CloudWatch también puede enviarle un mensaje de Amazon SNS cuando la alarma cambie de estado.

Las alarmas de CloudWatch solo invocan acciones cuando el estado cambia y se mantiene durante el período que ha especificado.

Para configurar una alarma

1. Inicie sesión en la CloudWatch consola AWS Management Console y ábrala en <https://console.aws.amazon.com/cloudwatch/>.
2. Elija Alarms (Alarmas) y, a continuación, seleccione Create Alarm (Crear alarma).
3. Elija las métricas de Kendra y, a continuación, elija una métrica.
4. En Time Range, elija un intervalo de tiempo para monitorizar y, a continuación, elija Next.
5. Introduzca un Name (Nombre) y una Description (Descripción).
6. En Whenever, elija \geq y escriba un valor máximo.
7. Si desea que CloudWatch envíe un correo electrónico cuando se alcance el estado de alarma, dentro de la sección Actions, en Whenever this alarm, elija State is ALARM. En Enviar notificación a, elige una lista de correo o elige Nueva lista y crea una nueva lista de correo
8. Obtenga una vista previa de la alarma en la sección Alarm Preview. Si está satisfecho con la alarma, elija Create Alarm.

CloudWatchMétricas para trabajos de sincronización de índices

La siguiente tabla describe las métricas de Amazon Kendra para los trabajos de sincronización de fuentes de datos.

Métrica	Descripción
DocumentsCrawled	<p>La cantidad de documentos que el trabajo de sincronización escaneó o descubrió durante la ejecución.</p> <p>Dimensiones:</p> <ul style="list-style-type: none">• IndexId• DataSourceld <p>Unidad: recuento</p>
DocumentsSubmittedForIndexing	<p>El número de documentos que el trabajo de sincronización envió al índice.</p> <p>Dimensiones:</p> <ul style="list-style-type: none">• IndexId• DataSourceld

Métrica	Descripción
	Unidad: recuento
DocumentsSubmittedForIndexingFailed	<p>El número de documentos que no se pudieron indexar. Compruebe el contenido del CloudWatch registro del trabajo de sincronización para obtener más información.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld
	Unidad: recuento
DocumentsSubmittedForDeletion	<p>El número de documentos que el trabajo de sincronización solicitó que se eliminaran del índice.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld
	Unidad: recuento
DocumentsSubmittedForDeletionFailed	<p>El número de documentos que no se pudieron eliminar. Compruebe el contenido del CloudWatch registro del trabajo de sincronización para obtener más información.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld
	Unidad: recuento

Métricas para las fuentes de datos de Amazon Kendra

La siguiente tabla describe las métricas de Amazon Kendra para los trabajos de sincronización de fuentes de datos. Las métricas marcadas con un asterisco (*) solo se utilizan para las fuentes de datos de Amazon S3.

Métrica	Descripción
DocumentsSkippedNoChange *	<p>La cantidad de documentos examinados y comprobados que no cambiaron, por lo que no se enviaron para su indexación.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld

Métrica	Descripción
	Unidad: recuento
DocumentsSkippedInvalidMetadata *	<p>El número de documentos omitidos porque había un problema con el archivo de metadatos asociado. Compruebe el contenido del CloudWatch registro de la ejecución de sincronización para obtener más información.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld
	Unidad: recuento
DocumentsCrawled	<p>El número de archivos de documentos examinados.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld
	Unidad: recuento
DocumentsSubmittedForDeletion	<p>El número de documentos examinados que se eliminaron de la fuente de datos y se enviaron para su eliminación.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld
	Unidad: recuento
DocumentsSubmittedForDeletionFailed	<p>El número de documentos que no se eliminaron de una fuente de datos.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld
	Unidad: recuento

Métrica	Descripción
DocumentsSubmittedForIndexing	<p>El número de documentos examinados y presentados para su indexación.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld <p>Unidad: recuento</p>
DocumentsSubmittedForIndexingFailed	<p>La cantidad de documentos enviados para su indexación que no se pudieron indexar.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld <p>Unidad: recuento</p>

Métricas para documentos indexados

En la siguiente tabla se describen las métricas de Amazon Kendra para los documentos indexados. Para los documentos que se indexan mediante la [BatchPutDocument](#) operación, solo se admite la IndexId dimensión.

Métrica	Descripción
DocumentsIndexed	<p>El número de documentos indexados.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld <p>Unidad: recuento</p>
DocumentsFailedToIndex	<p>El número de documentos que no se han podido indexar. Consulte el contenido del CloudWatch registro para obtener más información.</p> <p>Dimensiones:</p> <ul style="list-style-type: none"> • IndexId • DataSourceld <p>Unidad: recuento</p>
IndexQueryCount	<p>El número de consultas de índice por minuto.</p> <p>Dimensiones:</p>

Métrica	Descripción
	<ul style="list-style-type: none">• IndexId <p>Unidad: recuento</p>

Supervisión de Amazon Kendra con Amazon Logs CloudWatch

Amazon Kendra utiliza Amazon CloudWatch Logs para proporcionarle información sobre el funcionamiento de sus fuentes de datos. Amazon Kendra registra los detalles del proceso de los documentos a medida que se indexan. Registra los errores de la fuente de datos que se producen mientras se indexan los documentos. Los registros se utilizan para supervisar, almacenar y acceder a los archivos de registro.

CloudWatch Los registros almacenan los eventos de registro en un flujo de registro que forma parte de un grupo de registros. Amazon Kendra utiliza estas funciones de la siguiente manera:

- Grupos de registro: Amazon Kendra almacena todos sus flujos de registro en un solo grupo de registros para cada índice. Amazon Kendra crea el grupo de registros cuando se crea el índice. El identificador del grupo de registros siempre comienza por «aws/kendra/».
- Flujo de registro: Amazon Kendra crea un nuevo flujo de registro de fuentes de datos en el grupo de registros para cada trabajo de sincronización de índices que ejecute. También crea un nuevo flujo de registro de documentos cuando un flujo alcanza aproximadamente 500 entradas.
- Entradas de registro: Amazon Kendra crea una entrada de registro en el flujo de registro a medida que indexa los documentos. Cada entrada proporciona información sobre el procesamiento del documento o sobre cualquier error que se encuentre.

Para obtener más información sobre el uso de CloudWatch registros, consulte [Qué son los registros de Amazon Cloud Watch](#) en la Guía del usuario de Amazon Cloud Watch Logs.

Amazon Kendra crea dos tipos de flujos de registro:

- [Flujos de registro de fuentes de datos \(p. 636\)](#)
- [Flujos de registro de documentos \(p. 637\)](#)

Flujos de registro de fuentes de datos

Los flujos de registro de fuentes de datos publican entradas sobre sus trabajos de sincronización de índices. Cada trabajo de sincronización crea un nuevo flujo de registro que utiliza para publicar entradas. El nombre del flujo de registro es:

```
data source id/YYYY-MM-DD-HH/data source sync job ID
```

Se crea un nuevo flujo de registro para cada trabajo de sincronización ejecutado.

Hay tres tipos de mensajes de registro publicados en un flujo de registro de una fuente de datos:

- Un mensaje de registro de un documento que no se pudo enviar para su indexación. A continuación se muestra un ejemplo de este mensaje para un documento de una fuente de datos de S3:

```
{
```

```
    "DocumentId": "document ID",
    "S3Path": "s3://bucket/prefix/object",
    "Message": "Failed to ingest document via BatchPutDocument.",
    "ErrorCode": "InvalidRequest",
    "ErrorMessage": "No document metadata configuration found for document attribute key
city."
}
```

- Un mensaje de registro de un documento que no se pudo enviar para su eliminación. El siguiente es un ejemplo de este mensaje:

```
{
    "DocumentId": "document ID",
    "Message": "Failed to delete document via BatchDeleteDocument.",
    "ErrorCode": "InvalidRequest",
    "ErrorMessage": "Document can't be deleted because it doesn't exist."
}
```

- Un mensaje de registro cuando se encuentra un archivo de metadatos no válido para un documento en un bucket de Amazon S3. El siguiente es un ejemplo de este mensaje.

```
{
    "Message": "Found invalid metadata
file bucket/prefix/filename.extension.metadata.json."
}
```

- En el caso de SharePoint los conectores de bases de datos, Amazon Kendra solo escribe mensajes en el flujo de registro si no se puede indexar un documento. A continuación se muestra un ejemplo del mensaje de error que registra Amazon Kendra.

```
{
    "DocumentID": "document ID",
    "IndexID": "index ID",
    "SourceURI": "",
    "CrawlStatus": "FAILED",
    "ErrorCode": "403",
    "ErrorMessage": "Access Denied",
    "DataSourceErrorCode": "403"
}
```

Flujos de registro de documentos

Amazon Kendra registra la información sobre el procesamiento de los documentos mientras se indexan. Registra un conjunto de mensajes para documentos almacenados en una fuente de datos de Amazon S3. Registra los errores solo de los documentos almacenados en una fuente de datos de Microsoft SharePoint o de una base de datos.

Si los documentos se agregaron al índice mediante la [BatchPutDocument](#) operación, el flujo de registro se denomina de la siguiente manera:

YYYY-MM-DD-HH/UUID

Si los documentos se agregaron al índice mediante una fuente de datos, el flujo de registro se denomina de la siguiente manera:

dataSourceId/YYYY-MM-DD-HH/UUID

Cada flujo de registro contiene hasta 500 mensajes.

Si se produce un error al indexar un documento, se envía este mensaje a la secuencia de registro:

```
{  
    "DocumentId": "document ID",  
    "IndexName": "index name",  
    "IndexId": "index ID"  
    "SourceURI": "source URI"  
    "IndexingStatus": "DocumentFailedToIndex",  
    "ErrorCode": "400 | 500",  
    "ErrorMessage": "message"  
}
```

Seguridad en Amazon Kendra

La seguridad en la nube de AWS es la mayor prioridad. Como cliente de AWS, se beneficia de una arquitectura de red y un centro de datos que se han diseñado para satisfacer los requisitos de seguridad de las organizaciones más exigentes.

La seguridad es una responsabilidad compartida entre AWS y usted. El [modelo de responsabilidad compartida](#) la describe como seguridad de la nube y seguridad en la nube:

- Seguridad de la nube—AWSes responsable de proteger la infraestructura que funcionaAWSservicios en elAWSNube.AWStambién le proporciona servicios que puede utilizar de forma segura. Los auditores externos prueban y verifican periódicamente la eficacia de nuestra seguridad como parte de los [AWSProgramas de conformidad de](#). Para obtener más información sobre los programas de cumplimiento que se aplican a Amazon Kendra, consulte[AWSServicios incluidos en el ámbito de aplicación por programa de cumplimiento](#).
- Seguridad en la nube—Su responsabilidad está determinada por laAWSservicio que utilizas. También es responsable de otros factores, incluida la confidencialidad de los datos, los requisitos de la empresa y la legislación y los reglamentos aplicables.

Esta documentación le ayuda a entender cómo aplicar el modelo de responsabilidad compartida al utilizar Amazon Kendra. En los siguientes temas se muestra cómo configurar Amazon Kendra para cumplir sus objetivos de seguridad y cumplimiento. También aprenderás a usar otrosAWSservicios que le ayudan a supervisar y proteger sus recursos de Amazon Kendra.

Temas

- [Protección de datos en Amazon Kendra \(p. 639\)](#)
- [Amazon Kendra y puntos finales de VPC de interfaz \(AWS PrivateLink\) \(p. 641\)](#)
- [Administración de identidades y accesos para Amazon Kendra \(p. 642\)](#)
- [Prácticas recomendadas de seguridad \(p. 660\)](#)
- [Registro y monitoreo en Amazon Kendra \(p. 660\)](#)
- [Validación de conformidad para Amazon Kendra \(p. 660\)](#)
- [Resiliencia en Amazon Kendra \(p. 661\)](#)
- [Seguridad de infraestructura en Amazon Kendra \(p. 662\)](#)
- [Configuración y análisis de vulnerabilidades en AWS Identity and Access Management \(p. 662\)](#)

Protección de datos en Amazon Kendra

EIAWS [modelo de responsabilidad compartida](#)se aplica a la protección de datos en Amazon Kendra. Como se describe en este modelo, AWS es responsable de proteger la infraestructura global que ejecuta toda la Nube de AWS. Usted es responsable de mantener el control sobre el contenido alojado en esta infraestructura. Este contenido incluye la configuración de seguridad y las tareas de administración para el que utiliza Servicios de AWS. Para obtener más información sobre la privacidad de los datos, consulte las [Preguntas frecuentes sobre la privacidad de datos](#). Para obtener información sobre la protección de datos en Europa, consulte la publicación de blog [AWSShared Responsibility Model and GDPR](#) en el Blog de seguridad de AWS.

Con fines de protección de datos, recomendamos proteger las credenciales de Cuenta de AWS y configurar cuentas de usuario individuales con AWS IAM Identity Center (successor to AWS Single Sign-On) o AWS Identity and Access Management (IAM). De esta manera, solo se otorgan a cada usuario los permisos necesarios para cumplir con sus obligaciones laborales. También recomendamos proteger sus datos de las siguientes formas:

- Utilice la autenticación multifactor (MFA) con cada cuenta.
- Utilice SSL/TLS para comunicarse con los recursos de AWS. Nosotros exigimos TLS 1.2 y recomendamos TLS 1.3.
- Configure la API y el registro de actividad del usuario con AWS CloudTrail.
- Utilice las soluciones de cifrado de AWS, junto con todos los controles de seguridad predeterminados dentro de los servicios de Servicios de AWS.
- Utilice servicios de seguridad administrados avanzados, como Amazon Macie, que lo ayuden a detectar y proteger los datos confidenciales almacenados en Amazon S3.
- Si necesita módulos criptográficos validados FIPS 140-2 al acceder a AWS a través de una interfaz de línea de comandos o una API, utilice un punto de conexión de FIPS. Para obtener más información sobre los puntos de conexión de FIPS disponibles, consulte [Estándar de procesamiento de la información federal \(FIPS\) 140-2](#).

Se recomienda encarecidamente no ingresar nunca información confidencial o sensible, como, por ejemplo, direcciones de correo electrónico de clientes, en etiquetas o campos de formato libre, tales como el campo Nombre. Esto incluye cuando trabajas con Amazon, Kendra u otros Servicios de AWS mediante la consola, la API, AWS CLI, o AWSSDK. Cualquier dato que ingrese en etiquetas o campos de formato libre utilizados para nombres se pueden emplear para los registros de facturación o diagnóstico. Si proporciona una URL a un servidor externo, recomendamos encarecidamente que no incluya información de credenciales en la URL a fin de validar la solicitud para ese servidor.

Cifrado en reposo

Amazon Kendra cifra sus datos en reposo con la clave de cifrado que elija. Puede elegir una de las siguientes opciones:

- Un AWS-de propiedad AWS KMS. Si no especificas una clave de cifrado, tus datos se cifrarán con esta clave de forma predeterminada.
- Un AWS-clave de KMS gestionada en su cuenta. Amazon Kendra crea, administra y utiliza esta clave en su nombre. El nombre de la clave esaws/kendra.
- Una clave gestionada por el cliente. Puedes proporcionar el ARN de una clave de cifrado que creaste en tu cuenta. Cuando utilice una clave de KMS administrada por el cliente, debe proporcionar a la clave una política de claves que permita a Amazon Kendra utilizarla. Seleccione una clave KMS de cifrado simétrico administrada por el cliente. Amazon Kendra no admite claves KMS asimétricas. Para obtener más información, consulte [Administración de claves \(p. 640\)](#).

Cifrado en tránsito

Amazon Kendra usa el protocolo HTTPS para comunicarse con la aplicación cliente. Utiliza HTTPS y AWS Firms para comunicarse con otros servicios en nombre de su aplicación. Si usas una VPC, puedes usar AWS PrivateLink para establecer una conexión privada entre su VPC y Amazon Kendra.

Administración de claves

Amazon Kendra cifra el contenido del índice mediante uno de los tres tipos de claves. Puede elegir una de las siguientes opciones:

- UnAWS-de propiedadAWSKILÓMETROS. Esta es la opción predeterminada.
- UnAWS-clave KMS gestionada. Esta clave se crea en su cuenta y Amazon Kendra la administra y utiliza en su nombre.
- Una clave de KMS administrada por el cliente. Puede crear la clave al crear un índice o una fuente de datos de Amazon Kendra, o bien puede crear la clave mediante laAWS KMSconsola. Seleccione una clave de KMS de cifrado simétrico administrada por el cliente. Amazon Kendra no admite claves KMS asimétricas. Para obtener más información, consulte [Uso de claves simétricas y asimétricas](#) en la Guía para desarrolladores de AWS Key Management Service.

Amazon Kendra y puntos finales de VPC de interfaz (AWS PrivateLink)

Puede establecer una conexión privada entre su VPC y Amazon Kendra creando un punto final de interfaz VPC. Los puntos finales de la interfaz funcionan con[AWS PrivateLink](#), una tecnología que le permite acceder de forma privada a las API de Amazon Kendra sin una puerta de enlace a Internet, un dispositivo NAT, una conexión VPN oAWSConexión Direct Connect. Las instancias de su VPC no necesitan direcciones IP públicas para comunicarse con las API de Amazon Kendra. El tráfico entre su VPC y Amazon Kendra no sale de la red de Amazon.

Cada punto de enlace de la interfaz está representado por una o más [interfaces de red elásticas](#) en las subredes.

Para obtener más información, consulte[Terminales de interfaz VPC \(AWS PrivateLink\)](#)en elGuía del usuario de Amazon VPC.

Consideraciones sobre los endpoints de Amazon Kendra VPC

Antes de configurar una interfaz VPC endpoint para Amazon Kendra, asegúrese de revisar[Propiedades y limitaciones de los terminales de la interfaz](#)en elGuía del usuario de Amazon VPC.

Amazon Kendra permite realizar llamadas a todas sus acciones de API desde su VPC.

Creación de un endpoint de interfaz VPC para Amazon Kendra

Puede crear un punto de conexión de VPC para el servicio Amazon Kendra mediante la consola de Amazon VPC o laAWS Command Line Interface(AWS CLI). Para obtener más información, consulte [Creación de un punto de conexión de interfaz](#) en la Guía del usuario de Amazon VPC.

Cree un punto de conexión de VPC para Amazon Kendra con el siguiente nombre de servicio:

- com.amazonaws.*region*.kendra

Tras crear un punto de enlace de VPC, puede utilizar el siguiente ejemploAWS CLI comando que usa el `endpoint-url` parámetro para especificar un punto final de interfaz para la API de Amazon Kendra:

```
aws kendra list-indices --endpoint-url https://VPC endpoint
```

donde *Punto final de VPC* es el nombre DNS generado cuando se crea el extremo de la interfaz. Este nombre incluye el ID del punto de enlace de VPC, el nombre del servicio Amazon Kendra y el nombre de la región. Por ejemplo, vpce-1234-abcdef.kendra.us-west-2.vpce.amazonaws.com.

Si activa el DNS privado para el endpoint, puede realizar solicitudes de API a Amazon Kendra utilizando su nombre DNS predeterminado para la región, por ejemplo,kendra.us-east-1.amazonaws.com.

Para obtener más información, consulte [Acceso a un servicio a través de un punto de conexión de interfaz](#) en la Guía del usuario de Amazon VPC.

Creación de una política de puntos de conexión de VPC para Amazon Kendra

Puede adjuntar una política de punto de conexión a su punto de conexión de VPC que controle el acceso a Amazon Kendra. La política especifica la siguiente información:

- La entidad principal que puede realizar acciones.
- Las acciones que se pueden realizar.
- Los recursos en los que se pueden llevar a cabo las acciones.

Para obtener más información, consulte [Control del acceso a los servicios con puntos de enlace de la VPC](#) en la guía del usuario de Amazon VPC.

Ejemplo: política de punto de conexión de VPC para acciones de Amazon Kendra

El siguiente es un ejemplo de una política de puntos de conexión para Amazon Kendra. Cuando se adjunta a un punto de conexión, esta política otorga acceso a las acciones de Amazon Kendra listadas a todos los principales de todos los recursos.

```
{  
    "Statement": [  
        {  
            "Principal": "*",  
            "Effect": "Allow",  
            "Action": [  
                "kendra:Query"  
            ],  
            "Resource": "*"  
        }  
    ]  
}
```

Administración de identidades y accesos para Amazon Kendra

AWS Identity and Access Management (IAM) es un Servicio de AWS que ayuda a los administradores a controlar de forma segura el acceso a los recursos de AWS. Los administradores de IAM controlan quién puede ser autenticado(iniciado sesión) y autorizado(tienen permisos) para usar los recursos de Amazon Kendra. IAM es un Servicio de AWS que se puede utilizar sin cargo adicional.

Temas

- [Público \(p. 643\)](#)

- [Autenticación con identidades \(p. 643\)](#)
- [Administración de acceso mediante políticas \(p. 645\)](#)
- [Cómo funciona Amazon Kendra con IAM \(p. 647\)](#)
- [Ejemplos de políticas basadas en la identidad de Amazon Kendra \(p. 651\)](#)
- [AWSpolíticas gestionadas para Amazon Kendra \(p. 655\)](#)
- [Solución de problemas de identidad y acceso a Amazon Kendra \(p. 658\)](#)

Público

Cómo se usa AWS Identity and Access Management(IAM) varía según el trabajo que realice en Amazon Kendra.

Usuario del servicio— Si utiliza el servicio Amazon Kendra para realizar su trabajo, su administrador le proporcionará las credenciales y los permisos que necesita. A medida que utilice más funciones de Amazon Kendra para realizar su trabajo, es posible que necesite permisos adicionales. Entender cómo se administra el acceso puede ayudarlo a solicitar los permisos correctos a su administrador. Si no puede acceder a una función de Amazon Kendra, consulte [Solución de problemas de identidad y acceso a Amazon Kendra \(p. 658\)](#).

Administrador de servicios— Si está a cargo de los recursos de Amazon Kendra en su empresa, probablemente tenga acceso completo a Amazon Kendra. Su trabajo es determinar a qué funciones y recursos de Amazon Kendra deben acceder los usuarios de su servicio. Luego, debe enviar solicitudes a su administrador de IAM para cambiar los permisos de los usuarios de su servicio. Revise la información de esta página para conocer los conceptos básicos de IAM. Para obtener más información sobre cómo su empresa puede utilizar IAM con Amazon Kendra, consulte [Cómo funciona Amazon Kendra con IAM \(p. 647\)](#).

Administrador de IAM— Si es administrador de IAM, puede que desee obtener detalles sobre cómo puede redactar políticas para administrar el acceso a Amazon Kendra. Para ver ejemplos de políticas basadas en identidades de Amazon Kendra que puede usar en IAM, consulte [Ejemplos de políticas basadas en la identidad de Amazon Kendra \(p. 651\)](#).

Autenticación con identidades

La autenticación es la manera de iniciar sesión en AWS mediante credenciales de identidad. Debe estar autenticado (haber iniciado sesión en AWS) como el Usuario raíz de la cuenta de AWS, como un usuario de IAM o asumiendo un rol de IAM.

Puede iniciar sesión en AWS como una identidad federada mediante las credenciales proporcionadas a través de una fuente de identidad. AWS IAM Identity Center (successor to AWS Single Sign-On) Los usuarios (IAM Identity Center), la autenticación de inicio de sesión único de su empresa y sus credenciales de Google o Facebook son ejemplos de identidades federadas. Al iniciar sesión como identidad federada, su administrador habrá configurado previamente la federación de identidades mediante roles de IAM. Cuando accede a AWS mediante la federación, está asumiendo un rol de forma indirecta.

Según el tipo de usuario que sea, puede iniciar sesión en la AWS Management Console o en el portal de acceso AWS. Para obtener más información sobre el inicio de sesión en AWS, consulte [Cómo iniciar sesión en su Cuenta de AWS](#) en la Guía del usuario de AWS Sign-In.

Si accede a AWS mediante programación, AWS proporciona un kit de desarrollo de software (SDK) y una interfaz de línea de comandos (CLI) para firmar criptográficamente las solicitudes mediante el uso de las credenciales. Si no usa las herramientas de AWS, debe firmar usted mismo las solicitudes. Para obtener más información sobre la firma de solicitudes, consulte [Firma de solicitudes API de AWS](#) en la Guía del usuario de IAM.

Independientemente del método de autenticación que use, es posible que deba proporcionar información de seguridad adicional. Por ejemplo, AWS le recomienda el uso de la autenticación multifactor (MFA) para aumentar la seguridad de su cuenta. Para obtener más información, consulte [Autenticación multifactor](#) en la Guía del usuario de AWS IAM Identity Center (successor to AWS Single Sign-On) y [Uso de la autenticación multifactor \(MFA\) en AWS](#) en la Guía del usuario de IAM.

Usuario raíz de Cuenta de AWS

Cuando se crea una Cuenta de AWS, se comienza con una identidad de inicio de sesión que tiene acceso completo a todos los recursos y Servicios de AWS de la cuenta. Esta identidad recibe el nombre de usuario raíz de la Cuenta de AWS y se accede a ella iniciando sesión con el email y la contraseña que utilizó para crear la cuenta. Recomendamos que no utilice el usuario raíz para las tareas cotidianas. Proteja las credenciales del usuario raíz y utilícelas solo para las tareas que este pueda realizar. Para ver la lista completa de las tareas que requieren que inicie sesión como usuario raíz, consulte [Tasks that require root user credentials](#) (Tareas que requieren credenciales de usuario raíz) en la Guía de referencia de AWS Account Management.

Usuarios y grupos de IAM

Un [usuario de IAM](#) es una identidad de la Cuenta de AWS que dispone de permisos específicos para una sola persona o aplicación. Siempre que sea posible, recomendamos emplear credenciales temporales, en lugar de crear usuarios de IAM que tengan credenciales de larga duración como contraseñas y claves de acceso. No obstante, si tiene casos de uso específicos que requieran credenciales de larga duración con usuarios de IAM, recomendamos rotar las claves de acceso. Para obtener más información, consulte [Rotar las claves de acceso periódicamente para casos de uso que requieran credenciales de larga duración](#) en la Guía del usuario de IAM.

Un [grupo de IAM](#) es una identidad que especifica un conjunto de usuarios de IAM. No puede iniciar sesión como grupo. Puede usar los grupos para especificar permisos para varios usuarios a la vez. Los grupos facilitan la administración de los permisos de grandes conjuntos de usuarios. Por ejemplo, podría tener un grupo cuyo nombre fuese IAMAdmins y conceder permisos a dicho grupo para administrar los recursos de IAM.

Los usuarios son diferentes de los roles. Un usuario se asocia exclusivamente a una persona o aplicación, pero la intención es que cualquier usuario pueda asumir un rol que necesite. Los usuarios tienen credenciales permanentes a largo plazo y los roles proporcionan credenciales temporales. Para obtener más información, consulte [Cuándo crear un usuario de IAM \(en lugar de un rol\)](#) en la Guía del usuario de IAM.

IAM roles

Un [rol de IAM](#) es una identidad de la Cuenta de AWS que dispone de permisos específicos. Es similar a un usuario de IAM, pero no está asociado a una determinada persona. Puede asumir temporalmente un rol de IAM en la AWS Management Console[cambiando de roles](#). Puede asumir un rol llamando a una operación de la AWS CLI o de la API de AWS, o utilizando una URL personalizada. Para obtener más información acerca de los métodos para el uso de roles, consulte [Uso de roles de IAM](#) en la Guía del usuario de IAM.

Los roles de IAM con credenciales temporales son útiles en las siguientes situaciones:

- Acceso de usuario federado: para asignar permisos a una identidad federada, puede crear un rol y definir permisos para este. Cuando se autentica una identidad federada, se asocia la identidad al rol y se le conceden los permisos que están definidos en este. Para obtener información acerca de roles para federación, consulte [Creación de un rol para un proveedor de identidades de terceros](#) en la Guía del usuario de IAM. Si utiliza el Centro de identidades de IAM, debe configurar un conjunto de permisos. El Centro de identidades de IAM correlaciona el conjunto de permisos con un rol en IAM para controlar a qué pueden acceder las identidades después de autenticarse. Para obtener información acerca de los conjuntos de permisos, consulte [Conjuntos de permisos](#) en la Guía del usuario de AWS IAM Identity Center (successor to AWS Single Sign-On).

- Permisos de usuario de IAM temporales: un usuario de IAM puede asumir un rol de IAM para recibir temporalmente permisos distintos que le permitan realizar una tarea concreta.
- Acceso entre cuentas: puede utilizar un rol de IAM para permitir que alguien (una entidad principal de confianza) de otra cuenta acceda a los recursos de la cuenta. Los roles son la forma principal de conceder acceso entre cuentas. No obstante, con algunos Servicios de AWS se puede adjuntar una política directamente a un recurso (en lugar de utilizar un rol como representante). Para obtener información acerca de la diferencia entre los roles y las políticas basadas en recursos para el acceso entre cuentas, consulte [Cómo los roles de IAM difieren de las políticas basadas en recursos](#) en la Guía del usuario de IAM.
- Acceso entre servicios: algunos Servicios de AWS utilizan características de otros Servicios de AWS. Por ejemplo, cuando realiza una llamada en un servicio, es común que ese servicio ejecute aplicaciones en Amazon EC2 o almacene objetos en Amazon S3. Es posible que un servicio haga esto usando los permisos de la entidad principal, usando un rol de servicio o usando un rol vinculado a servicios.
- Permisos principales: cuando utiliza un usuario o un rol de IAM para llevar a cabo acciones en AWS, se lo considera una entidad principal. Las políticas conceden permisos a una entidad principal. Cuando utiliza algunos servicios, es posible que realice una acción que desencadene otra acción en un servicio diferente. En este caso, debe tener permisos para realizar ambas acciones. Para ver si una acción requiere acciones dependientes adicionales en una política, consulte [Claves de acciones, recursos y condiciones para Amazon Kendra](#) en el Referencia de autorización de servicio.
- Rol de servicio: un rol de servicio es un [rol de IAM](#) que adopta un servicio para realizar acciones en su nombre. Un administrador de IAM puede crear, modificar y eliminar un rol de servicio desde IAM. Para obtener más información, consulte [Creación de un rol para delegar permisos a un Servicio de AWS](#) en la Guía del usuario de IAM.
- Rol vinculado a servicio: un rol vinculado a servicio es un tipo de rol de servicio que está vinculado a un Servicio de AWS. El servicio puede asumir el rol para realizar una acción en su nombre. Los roles vinculados a servicios aparecen en la Cuenta de AWS y son propiedad del servicio. Un administrador de IAM puede ver, pero no editar, los permisos de los roles vinculados a servicios.
- Aplicaciones que se ejecutan en Amazon EC2: puede utilizar un rol de IAM que le permita administrar credenciales temporales para las aplicaciones que se ejecutan en una instancia de EC2 y realizar solicitudes a la AWS CLI o a la API de AWS. Es preferible hacerlo de este modo a almacenar claves de acceso en la instancia de EC2. Para asignar un rol de AWS a una instancia de EC2 y ponerla a disposición de todas las aplicaciones, cree un perfil de instancia adjuntado a la instancia. Un perfil de instancia contiene el rol y permite a los programas que se ejecutan en la instancia de EC2 obtener credenciales temporales. Para obtener más información, consulte [Uso de un rol de IAM para conceder permisos a aplicaciones que se ejecutan en instancias Amazon EC2](#) en la Guía del usuario de IAM.

Para obtener información sobre el uso de los roles de IAM, consulte [Cuándo crear un rol de IAM \(en lugar de un usuario\)](#) en la Guía del usuario de IAM.

Administración de acceso mediante políticas

Para controlar el acceso en AWS, se crean políticas y se asocian a identidades o recursos de AWS. Una política es un objeto de AWS que, cuando se asocia a una identidad o un recurso, define sus permisos. AWS evalúa estas políticas cuando una entidad principal (sesión de rol, usuario o usuario raíz) realiza una solicitud. Los permisos en las políticas determinan si la solicitud se permite o se deniega. Las mayoría de las políticas se almacenan en AWS como documentos JSON. Para obtener más información sobre la estructura y el contenido de los documentos de política JSON, consulte [Información general de políticas JSON](#) en la Guía del usuario de IAM.

Los administradores pueden utilizar las políticas JSON de AWS para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y bajo qué condiciones.

De forma predeterminada, los usuarios y los roles no tienen permisos. Un administrador de IAM puede crear políticas de IAM para conceder permisos a los usuarios para realizar acciones en los recursos que

necesitan. A continuación, el administrador puede agregar las políticas de IAM a los roles y los usuarios pueden asumirlos.

Las políticas de IAM definen permisos para una acción independientemente del método que se utilice para realizar la operación. Por ejemplo, suponga que dispone de una política que permite la acción `iam:GetRole`. Un usuario con dicha política puede obtener información del usuario de la AWS Management Console, la AWS CLI o la API de AWS.

Políticas basadas en identidad

Las políticas basadas en identidad son documentos de políticas de permisos JSON que puede adjuntar a una identidad, como un usuario de IAM, un grupo de usuarios o un rol. Estas políticas controlan qué acciones pueden realizar los usuarios y los roles, en qué recursos y bajo qué condiciones. Para obtener más información sobre cómo crear una política basada en identidad, consulte [Creación de políticas de IAM](#) en la Guía del usuario de IAM.

Las políticas basadas en identidad pueden clasificarse además como políticas insertadas o políticas administradas. Las políticas insertadas se integran directamente en un único usuario, grupo o rol. Las políticas administradas son políticas independientes que puede adjuntar a varios usuarios, grupos y roles de su Cuenta de AWS. Las políticas administradas incluyen las políticas administradas por AWS y las políticas administradas por el cliente. Para obtener más información acerca de cómo elegir una política administrada o una política insertada, consulte [Elegir entre políticas administradas y políticas insertadas](#) en la Guía del usuario de IAM.

Políticas basadas en recursos

Las políticas basadas en recursos son documentos de política JSON que se adjuntan a un recurso. Ejemplos de políticas basadas en recursos son las políticas de confianza de roles de IAM y las políticas de bucket de Amazon S3. En los servicios que admiten políticas basadas en recursos, los administradores de servicios pueden utilizarlos para controlar el acceso a un recurso específico. Para el recurso al que se adjunta la política, la política define qué acciones puede realizar una entidad principal especificada en ese recurso y en qué condiciones. Debe [especificar una entidad principal](#) en una política basada en recursos. Las entidades principales pueden incluir cuentas, usuarios, roles, usuarios federados o Servicios de AWS.

Las políticas basadas en recursos son políticas insertadas que se encuentran en ese servicio. No se puede utilizar políticas de IAM administradas por AWS en una política basada en recursos.

Listas de control de acceso (ACL)

Las listas de control de acceso (ACL) controlan qué entidades principales (miembros de cuentas, usuarios o roles) tienen permisos para acceder a un recurso. Las ACL son similares a las políticas basadas en recursos, aunque no utilizan el formato de documento de política JSON.

Amazon S3, AWS WAF y Amazon VPC son ejemplos de servicios que admiten las ACL. Para obtener más información sobre las ACL, consulte [Información general de Lista de control de acceso \(ACL\)](#) en la Guía para desarrolladores de Amazon Simple Storage Service.

Otros tipos de políticas

AWS admite otros tipos de políticas adicionales menos frecuentes. Estos tipos de políticas pueden establecer el máximo de permisos que los tipos de políticas más frecuentes le otorgan.

- Límites de permisos: un límite de permisos es una característica avanzada que le permite establecer los permisos máximos que una política basada en identidad puede conceder a una entidad de IAM (usuario o rol de IAM). Puede establecer un límite de permisos para una identidad. Los permisos resultantes son la intersección de las políticas basadas en identidad de la entidad y los límites de permisos. Las políticas

basadas en recursos que especifiquen el usuario o rol en el campo Principal no estarán restringidas por el límite de permisos. Una denegación explícita en cualquiera de estas políticas anulará el permiso. Para obtener más información sobre los límites de los permisos, consulte [Límites de permisos para las entidades de IAM](#) en la Guía del usuario de IAM.

- Políticas de control de servicio (SCP): las SCP son políticas de JSON que especifican los permisos máximos de una organización o una unidad organizativa en AWS Organizations. AWS Organizations es un servicio que le permite agrupar y administrar de manera centralizada varias Cuentas de AWS que posea su empresa. Si habilita todas las características en una organización, entonces podrá aplicar políticas de control de servicio (SCP) a una o todas sus cuentas. Una SCP limita los permisos para las entidades de las cuentas de miembros, incluido cada Usuario raíz de la cuenta de AWS. Para obtener más información acerca de Organizations y las SCP, consulte [Funcionamiento de las SCP](#) en la Guía del usuario de AWS Organizations.
- Políticas de sesión: las políticas de sesión son políticas avanzadas que se pasan como parámetro cuando se crea una sesión temporal mediante programación para un rol o un usuario federado. Los permisos de la sesión resultantes son la intersección de las políticas basadas en identidad del rol y las políticas de la sesión. Los permisos también pueden proceder de una política basada en recursos. Una denegación explícita en cualquiera de estas políticas anulará el permiso. Para obtener más información, consulte [Políticas de sesión](#) en la Guía del usuario de IAM.

Varios tipos de políticas

Cuando se aplican varios tipos de políticas a una solicitud, los permisos resultantes son más complicados de entender. Para obtener información acerca de cómo AWS decide si permitir o no una solicitud cuando hay varios tipos de políticas implicados, consulte [Lógica de evaluación de políticas](#) en la Guía del usuario de IAM.

Cómo funciona Amazon Kendra con IAM

Antes de utilizar IAM para administrar el acceso a Amazon Kendra, debe entender qué funciones de IAM están disponibles para su uso con Amazon Kendra. Para obtener una visión general de cómo Amazon Kendra y otros AWS los servicios funcionan con IAM, consulte [AWS Servicios que funcionan con IAM](#) en la Guía del usuario de IAM.

Temas

- [Políticas basadas en la identidad de Amazon Kendra \(p. 647\)](#)
- [Políticas basadas en recursos de Amazon Kendra \(p. 649\)](#)
- [Listas de control de acceso \(ACL\) \(p. 649\)](#)
- [Autorización basada en etiquetas de Amazon Kendra \(p. 650\)](#)
- [Funciones de IAM de Amazon Kendra \(p. 650\)](#)

Políticas basadas en la identidad de Amazon Kendra

Con las políticas basadas en identidad de IAM, puede especificar las acciones y recursos permitidos o denegados, así como las condiciones en las que se permiten o deniegan las acciones. Amazon Kendra admite acciones, recursos y claves de condiciones específicos. Para obtener información sobre todos los elementos que utiliza en una política JSON, consulte [Referencia de los elementos de las políticas JSON de IAM](#) en la Guía del usuario de IAM.

Acciones

Los administradores pueden utilizar las políticas JSON de AWS para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y bajo qué condiciones.

El elemento Action de una política JSON describe las acciones que puede utilizar para permitir o denegar el acceso en una política. Las acciones de la política generalmente tienen el mismo nombre que la operación de API de AWS asociada. Hay algunas excepciones, como acciones de solo permiso que no tienen una operación de API coincidente. También hay algunas operaciones que requieren varias acciones en una política. Estas acciones adicionales se denominan acciones dependientes.

Incluya acciones en una política para conceder permisos y así llevar a cabo la operación asociada.

Las acciones políticas de Amazon Kendra utilizan el siguiente prefijo antes de la acción:kendra:. Por ejemplo, para conceder permiso a alguien para incluir los índices de Amazon Kendra con la [ListIndices](#)Operación de API, incluye la `kendra>ListIndices`acción en su política. Las instrucciones de la política deben incluir un elemento Action o un elemento NotAction. Amazon Kendra define su propio conjunto de acciones que describen las tareas que puede realizar con este servicio.

Para especificar varias acciones en una única instrucción, sepárelas con comas del siguiente modo:

```
"Action": [  
    "kendra:action1",  
    "kendra:action2"
```

Puede utilizar caracteres comodín para especificar varias acciones (*). Por ejemplo, para especificar todas las acciones que comiencen con la palabra `Describe`, incluya la siguiente acción:

```
"Action": "kendra:Describe*"
```

Para ver una lista de las acciones de Amazon Kendra, consulte [Acciones definidas por Amazon Kendra](#)en elGuía del usuario de IAM.

Recursos

Los administradores pueden utilizar las políticas JSON de AWS para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y bajo qué condiciones.

El elemento Resource de la política JSON especifica el objeto u objetos a los que se aplica la acción. Las instrucciones deben contener un elemento Resource o NotResource. Como práctica recomendada, especifique un recurso utilizando el [Nombre de recurso de Amazon \(ARN\)](#). Puede hacerlo para acciones que admitan un tipo de recurso específico, conocido como permisos de nivel de recurso.

Para las acciones que no admiten permisos de nivel de recurso, como las operaciones de descripción, utilice un carácter comodín (*) para indicar que la instrucción se aplica a todos los recursos.

```
"Resource": "*"
```

El recurso de índice de Amazon Kendra tiene el siguiente ARN:

```
arn:${Partition}:kendra:${Region}:${Account}:index/${IndexId}
```

Para obtener más información acerca del formato de los ARN, consulte [Nombres de recursos de Amazon \(ARN\) y espacios de nombres de servicios de AWS](#).

Por ejemplo, para especificar un índice en la sentencia, utilice el GUID del índice en el siguiente ARN:

```
"Resource": "arn:aws:kendra:${Region}:${Account}:index/${GUID}"
```

Para especificar todos los índices que pertenecen a una cuenta específica, utilice el comodín (*):

```
"Resource": "arn:aws:${Region}:${Account}:index/*"
```

Algunas acciones de Amazon Kendra, como las de creación de recursos, no se pueden realizar en un recurso específico. En dichos casos, debe utilizar el carácter comodín (*).

```
"Resource": "*"
```

Para ver una lista de los tipos de recursos de Amazon Kendra y sus ARN, consulte [Recursos definidos por Amazon Kendra](#) en el Guía del usuario de IAM. Para saber con qué acciones puede especificar el ARN de cada recurso, consulte [Acciones definidas por Amazon Kendra](#).

Claves de condición

Los administradores pueden utilizar las políticas JSON de AWS para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y bajo qué condiciones.

El elemento Condition (o bloque de Condition) permite especificar condiciones en las que entra en vigor una instrucción. El elemento Condition es opcional. Puede crear expresiones condicionales que utilicen [operadores de condición](#), tales como igual o menor que, para que la condición de la política coincida con los valores de la solicitud.

Si especifica varios elementos de Condition en una instrucción o varias claves en un único elemento de Condition, AWS las evalúa mediante una operación AND lógica. Si especifica varios valores para una única clave de condición, AWS evalúa la condición con una operación lógica OR. Se deben cumplir todas las condiciones antes de que se concedan los permisos de la instrucción.

También puede utilizar variables de marcador de posición al especificar condiciones. Por ejemplo, puede conceder un permiso de usuario de IAM para acceder a un recurso solo si está etiquetado con su nombre de usuario de IAM. Para obtener más información, consulte [Elementos de la política de IAM: variables y etiquetas](#) en la Guía del usuario de IAM.

AWS admite claves de condición globales y claves de condición específicas del servicio. Para ver todas las claves de condición globales de AWS, consulte [Claves de contexto de condición globales de AWS](#) en la Guía del usuario de IAM.

Amazon Kendra no proporciona ninguna clave de condición específica del servicio, pero sí admite el uso de algunas claves de condición globales. Para ver todas las claves de condición globales de AWS, consulte [Claves de contexto de condición globales de AWS](#) en la Guía del usuario de IAM.

Ejemplos

Para ver ejemplos de políticas basadas en la identidad de Amazon Kendra, consulte [Ejemplos de políticas basadas en la identidad de Amazon Kendra \(p. 651\)](#).

Políticas basadas en recursos de Amazon Kendra

Amazon Kendra no admite políticas basadas en recursos.

Listas de control de acceso (ACL)

Amazon Kendra no admite listas de control de acceso (ACL) para acceder a AWS servicios y recursos.

Autorización basada en etiquetas de Amazon Kendra

Puede asociar etiquetas a ciertos tipos de recursos de Amazon Kendra para autorizar el acceso a esos recursos. Para controlar el acceso en función de las etiquetas, proporcione la información de las etiquetas en el elemento de condición de una política mediante elaws :RequestTag/*key-name*, oaws :TagKeysclaves de estado.

La siguiente tabla muestra las acciones, los tipos de recursos correspondientes y las claves de condición para el control de acceso basado en etiquetas. Cada acción se autoriza en función de las etiquetas asociadas al tipo de recurso correspondiente.

Acción	Tipo de recurso	Claves de condición
CreateDataSource		aws :RequestTag, aws :TagKeys
CreateFaq		aws :RequestTag, aws :TagKeys
CreateIndex		aws :RequestTag, aws :TagKeys
API_ListTagsForResource	fuente de datos, preguntas frecuentes, índice	
TagResource	fuente de datos, preguntas frecuentes, índice	aws :RequestTag, aws :TagKeys
UntagResource	fuente de datos, preguntas frecuentes, índice	aws :TagKeys

Para obtener información sobre cómo etiquetar los recursos de Amazon Kendra, consulte [Etiquetas \(p. 11\)](#). Para ver un ejemplo de política basada en identidades que limita el acceso a un recurso en función de las etiquetas de recursos, consulte [Ejemplos de políticas basadas en etiquetas \(p. 653\)](#). Para obtener más información sobre el uso de etiquetas para limitar el acceso a los recursos, consulte [Controlar el acceso mediante etiquetas](#) en el Guía del usuario de IAM.

Funciones de IAM de Amazon Kendra

Un [rol de IAM](#) es una entidad de la cuenta de AWS que dispone de permisos específicos.

Uso de credenciales temporales con Amazon Kendra

Puede utilizar credenciales temporales para iniciar sesión con federación, asumir un rol de IAM o asumir un rol de acceso entre cuentas. Para obtener credenciales de seguridad temporales, llame AWS STS Operaciones de API como [AssumeRole](#) o [GetFederationToken](#).

Amazon Kendra admite el uso de credenciales temporales.

Roles de servicio

Esta característica permite que un servicio asuma un [rol de servicio](#) en su nombre. Este rol permite que el servicio obtenga acceso a los recursos de otros servicios para completar una acción en su nombre. Los roles de servicio aparecen en su cuenta de IAM y son propiedad de la cuenta. Esto significa que un administrador de IAM puede cambiar los permisos de este rol. Sin embargo, hacerlo podría deteriorar la funcionalidad del servicio.

Amazon Kendra admite funciones de servicio.

Elección de un rol de IAM en Amazon Kendra

Al crear un índice, llame alBatchPutDocumentoperación, cree una fuente de datos o cree una pregunta frecuente, debe proporcionar un rol de acceso Amazon Resource Name (ARN) que Amazon Kendra utilice para acceder a los recursos requeridos en su nombre. Si ha creado un rol anteriormente, la consola de Amazon Kendra le proporciona una lista de roles entre los que puede elegir. Es importante elegir un rol que permita el acceso a los recursos que necesita. Para obtener más información, consulte [IAMfunciones de acceso para Amazon Kendra \(p. 15\)](#).

Ejemplos de políticas basadas en la identidad de Amazon Kendra

De forma predeterminada, los usuarios y los roles no tienen permiso para crear o modificar los recursos de Amazon Kendra. Tampoco pueden realizar tareas mediante la AWS Management Console, la AWS CLI, o la API de AWS. Un administrador de IAM debe crear políticas de IAM que concedan permisos a los usuarios y a los roles para realizar operaciones de la API concretas en los recursos especificados que necesiten. El administrador debe adjuntar esas políticas a los usuarios o grupos que necesiten esos permisos.

Para obtener información acerca de cómo crear una política basada en identidad de IAM con estos documentos de políticas JSON de ejemplo, consulte [Creación de políticas en la pestaña JSON](#) en la Guía del usuario de IAM.

Temas

- [Prácticas recomendadas relativas a políticas \(p. 651\)](#)
- [AWSPolíticas gestionadas \(predefinidas\) para Amazon Kendra \(p. 652\)](#)
- [Permitir a los usuarios consultar sus propios permisos \(p. 652\)](#)
- [Acceder a un índice de Amazon Kendra \(p. 653\)](#)
- [Ejemplos de políticas basadas en etiquetas \(p. 653\)](#)

Prácticas recomendadas relativas a políticas

Las políticas basadas en la identidad determinan si alguien puede crear, acceder o eliminar recursos de Amazon Kendra en su cuenta. Estas acciones pueden generar costos adicionales para su Cuenta de AWS. Siga estas directrices y recomendaciones al crear o editar políticas basadas en identidad:

- Comience con las políticas administradas por AWS y continúe con los permisos de privilegio mínimo: a fin de comenzar a conceder permisos a los usuarios y las cargas de trabajo, utilice las políticas administradas por AWS, que conceden permisos para muchos casos de uso comunes. Están disponibles en la Cuenta de AWS. Se recomienda definir políticas administradas por el cliente de AWS específicas para sus casos de uso a fin de reducir aún más los permisos. Con el fin de obtener más información, consulte las [políticas administradas por AWS](#) o las [políticas administradas por AWS para funciones de trabajo](#) en la Guía de usuario de IAM.
- Aplique permisos de privilegio mínimo: cuando establezca permisos con políticas de IAM, conceda solo los permisos necesarios para realizar una tarea. Para ello, debe definir las acciones que se pueden llevar a cabo en determinados recursos en condiciones específicas, también conocidos como permisos de privilegios mínimos. Con el fin de obtener más información sobre el uso de IAM para aplicar permisos, consulte [Políticas y permisos en IAM](#) en la Guía de usuario de IAM.
- Use condiciones en las políticas de IAM para restringir aún más el acceso: puede agregar una condición a sus políticas para limitar el acceso a las acciones y los recursos. Por ejemplo, puede escribir una condición de política para especificar que todas las solicitudes deben enviarse utilizando SSL. También

puede usar condiciones para conceder acceso a acciones de servicios si se emplean a través de un Servicio de AWS determinado, como por ejemplo AWS CloudFormation. Para obtener más información, consulte [Elementos de la política JSON de IAM: condición](#) en la Guía del usuario de IAM.

- Use el Analizador de acceso de IAM para validar las políticas de IAM con el fin de garantizar la seguridad y funcionalidad de los permisos: el Analizador de acceso de IAM valida políticas nuevas y existentes para que respeten el lenguaje (JSON) de las políticas de IAM y las prácticas recomendadas de IAM. IAM Access Analyzer proporciona más de 100 verificaciones de políticas y recomendaciones procesables para ayudar a crear políticas seguras y funcionales. Para obtener más información, consulte la [política de validación del Analizador de acceso de IAM](#) en la Guía de usuario de IAM.
- Solicite la autenticación multifactor (MFA): si se encuentra en una situación en la que necesita usuarios raíz o de IAM en su Cuenta de AWS, active la MFA para mayor seguridad. Para solicitar la MFA cuando se invocan las operaciones de la API, agregue las condiciones de MFA a sus políticas. Para obtener más información, consulte [Configuración de acceso a una API protegida por MFA](#) en la Guía de usuario de IAM.

Para obtener más información sobre las prácticas recomendadas de IAM, consulte las [Prácticas recomendadas de seguridad en IAM](#) en la Guía de usuario de IAM.

AWS Políticas gestionadas (predefinidas) para Amazon Kendra

AWS aborda muchos casos de uso comunes proporcionando políticas de IAM independientes creadas y administradas por AWS. Estas políticas se denominan AWS Políticas gestionadas. Las políticas administradas le permiten asignar permisos a usuarios, grupos y roles con más facilidad que si tuviera que escribir las políticas usted mismo. Para obtener más información, consulte [Adición de permisos a un usuario](#) en la Guía del usuario de IAM.

Los siguientes AWS Políticas gestionadas, que puede adjuntar a los grupos y roles de su cuenta, son específicas de Amazon Kendra:

- `AmazonKendraReadOnly`— Otorga acceso de solo lectura a los recursos de Amazon Kendra.
- `AmazonKendraFullAccess`— Otorga acceso completo para crear, leer, actualizar, eliminar, etiquetar y ejecutar todos los recursos de Amazon Kendra.

En el caso de la consola, tu rol también debe tener `iam:CreateRole, iam:CreatePolicy, iam:AttachRolePolicy, ys3>ListBucket` permisos.

Note

Para revisar estos permisos, inicie sesión en la consola de IAM y busque políticas específicas.

También puede crear sus propias políticas personalizadas para permitir los permisos para las acciones de la API de Amazon Kendra. Puede asociar estas políticas personalizadas a los roles o grupos de IAM que requieran esos permisos. Para ver ejemplos de políticas de IAM para Amazon Kendra, consulte [Ejemplos de políticas basadas en la identidad de Amazon Kendra \(p. 651\)](#).

Permitir a los usuarios consultar sus propios permisos

En este ejemplo, se muestra cómo podría crear una política que permite a los usuarios de IAM ver las políticas administradas e insertadas que se adjuntan a la identidad de sus usuarios. Esta política incluye permisos para llevar a cabo esta acción en la consola o mediante programación con la AWS CLI o la API de 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": [  

```

Acceder a un índice de Amazon Kendra

En este ejemplo, desea conceder un usuario en suAWSacceso a la cuenta para consultar un índice.

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Sid": "QueryIndex",  
            "Effect": "Allow",  
            "Action": [  
                "kendra:Query"  
            ],  
            "Resource": "arn:aws:kendra:${Region}:${Account}:index/${Index ID}"  
        }  
    ]  
}
```

Ejemplos de políticas basadas en etiquetas

Las políticas basadas en etiquetas son documentos de políticas de JSON que especifican las acciones que un director puede realizar en los recursos etiquetados.

Ejemplo: usar una etiqueta para acceder a un recurso

Esta política de ejemplo otorga un usuario o un rol en suAWSpermiso de cuenta para usar elQueryoperación con cualquier recurso etiquetado con la clave**department** y el valor**finance**.

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

```
"Statement": [
    {
        "Effect": "Allow",
        "Action": [
            "kendra:Query"
        ],
        "Resource": "*",
        "Condition": {
            "StringEquals": {
                "aws:ResourceTag/department": "finance"
            }
        }
    }
]
```

Ejemplo: utilice una etiqueta para activar las operaciones de Amazon Kendra

Esta política de ejemplo otorga un usuario o un rol en suAWSpermiso de cuenta para utilizar cualquier operación de Amazon Kendra, exceptoTagResourceoperación con cualquier recurso etiquetado con la clave**departmenty** el valor**finance**.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra:*",
            "Resource": "*"
        },
        {
            "Effect": "Deny",
            "Action": [
                "kendra:TagResource"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceTag/department": "finance"
                }
            }
        }
    ]
}
```

Ejemplo: usar una etiqueta para restringir el acceso a una operación

Este ejemplo de política restringe el acceso de un usuario o rol en suAWScuenta para usar elCreateIndexoperación a menos que el usuario proporcione la**departmentetiqueta** y tiene los valores permitidos**financeyIT**.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kendra:CreateIndex",
            "Resource": "*"
        },
        {
            "Effect": "Deny",
            "Action": "kendra:CreateIndex",
            "Resource": "*"
        }
    ]
}
```

```
"Resource": "*",
"Condition": {
    "Null": {
        "aws:RequestTag/department": "true"
    }
},
{
    "Effect": "Deny",
    "Action": "kendra>CreateIndex",
    "Resource": "*",
    "Condition": {
        "ForAnyValue:StringNotEquals": {
            "aws:RequestTag/department": [
                "finance",
                "IT"
            ]
        }
    }
}
]
```

AWSpolíticas gestionadas para Amazon Kendra

Para agregar permisos a usuarios, grupos y roles, es más fácil utilizar las políticas administradas de AWS que escribirlas uno mismo. Se necesita tiempo y experiencia para [crear políticas de IAM administradas por el cliente](#) que proporcionen a su equipo solo los permisos necesarios. Para comenzar a hacerlo con rapidez, puede utilizar nuestras políticas administradas por AWS. Estas políticas cubren casos de uso comunes y están disponibles en su cuenta de AWS. Para obtener más información sobre las políticas administradas por AWS, consulte [Políticas administradas por AWS](#) en la Guía del usuario de IAM.

Los servicios de AWS mantienen y actualizan las políticas administradas por AWS. No puede cambiar los permisos en las políticas administradas por AWS. En ocasiones, los servicios agregan permisos adicionales a una política administrada por AWS para admitir características nuevas. Este tipo de actualización afecta a todas las identidades (usuarios, grupos y roles) donde se asocia la política. Es más probable que los servicios actualicen una política administrada por AWS cuando se lanza una nueva característica o cuando se ponen a disposición nuevas operaciones. Los servicios no quitan permisos de una política administrada por AWS, por lo que las actualizaciones de políticas no deteriorarán los permisos existentes.

Además, AWS admite políticas administradas para funciones de trabajo que abarcan varios servicios. Por ejemplo, elReadOnlyAccess AWSIa política gestionada proporciona acceso de solo lectura a todosAWSservicios y recursos. Cuando un servicio lanza una nueva característica, AWS agrega permisos de solo lectura para las operaciones y los recursos nuevos. Para obtener una lista y descripciones de las políticas de funciones de trabajo, consulte [Políticas administradas de AWS para funciones de trabajo](#) en la Guía del usuario de IAM.

Política administrada por AWS: AmazonKendraReadOnly

Otorga acceso de solo lectura a los recursos de Amazon Kendra. Esta política incluye los siguientes permisos.

- **kendra**— Permite a los usuarios realizar acciones que devuelven una lista de elementos o detalles sobre un elemento. Esto incluye las operaciones de API que comienzan con `Describe`, `List`, `Query`, `BatchGetDocumentStatus`, `GetQuerySuggestions`, o `GetSnapshots`.

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Action": [  
                "kendra:Describe*",  
                "kendra>List*",  
                "kendra:Query",  
                "kendra:BatchGetDocumentStatus",  
                "kendra:GetQuerySuggestions",  
                "kendra:GetSnapshots"  
            ],  
            "Effect": "Allow",  
            "Resource": "*"  
        }  
    ]  
}
```

Política administrada por AWS: AmazonKendraFullAccess

Otorga acceso completo para crear, leer, actualizar, eliminar, etiquetar y ejecutar todos los recursos de Amazon Kendra. Esta política incluye los siguientes permisos.

- **kendra**—Permite a los directores acceder por lectura y escritura a todas las acciones de Amazon Kendra.
- **s3**—Permite a los directores obtener las ubicaciones de los buckets de Amazon S3 y listar los buckets.
- **iam**—Permite a los directores aprobar y enumerar funciones.
- **kms**—Permite a los directores describir y enumerar AWS KMS claves y alias.
- **secretsmanager**—Permite a los directores crear, describir y enumerar secretos.
- **ec2**—Permite a los directores describir los grupos de seguridad, las VCP (nube privada virtual) y las subredes.
- **cloudwatch**—Permite a los directores ver las métricas de Cloud Watch.

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Effect": "Allow",  
            "Action": "iam:PassRole",  
            "Resource": "*",  
            "Condition": {  
                "StringEquals": {  
                    "iam:PassedToService": "kendra.amazonaws.com"  
                }  
            }  
        },  
        {  
            "Effect": "Allow",  
            "Action": [  
                "iam>ListRoles"  
            ],  
            "Resource": "*"  
        },  
        {  
            "Effect": "Allow",  
            "Action": [  
                "kendra:BatchGetDocumentStatus",  
                "kendra:Describe*",  
                "kendra>List*",  
                "kendra:Query",  
                "kendra:GetQuerySuggestions",  
                "kendra:GetSnapshots"  
            ],  
            "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 se actualiza aAWSpolíticas gestionadas

Ver detalles sobre las actualizaciones deAWSgestionó las políticas de Amazon Kendra desde que este servicio comenzó a rastrear estos cambios. Para recibir alertas automáticas sobre los cambios en esta página, suscríbase a la fuente RSS de la página del historial de documentos de Amazon Kendra.

Cambio	Descripción	Fecha
AmazonKendraReadOnly —Añadir permiso al soporteGetSchemas, BatchGetDocuments y GetDocumentStatus	Amazon Kendra agregó nuevas API <code>GetSnapshots</code> y <code>BatchGetDocumentStatus</code> . <code>GetSnapshots</code> proporciona información sobre los resultados de la indexación de los documentos. <code>BatchGetDocumentStatus</code> supervisa el progreso de la indexación de sus documentos.	3 de enero de 2022
AmazonKendraReadOnly —Añadir permiso al soporteGetQuerySuggestions y GetDocumentStatus	Amazon Kendra ha añadido una nueva API <code>GetQuerySuggestions</code> que permite obtener sugerencias de consultas para consultas de búsqueda populares, lo que ayuda a guiar la búsqueda de los usuarios. Cuando los usuarios escriben su consulta de búsqueda, la consulta sugerida ayuda a completar automáticamente la búsqueda.	27 de mayo de 2021
Amazon Kendra comenzó a rastrear los cambios	Amazon Kendra comenzó a rastrear los cambios en su AWS Políticas gestionadas.	27 de mayo de 2021

Solución de problemas de identidad y acceso a Amazon Kendra

Utilice la siguiente información para diagnosticar y solucionar problemas comunes que puedan surgir al trabajar con Amazon Kendra e IAM.

Temas

- [No estoy autorizado a realizar una acción en Amazon Kendra \(p. 658\)](#)
- [No estoy autorizado a realizar iam:PassRole \(p. 659\)](#)
- [Soy administrador y quiero permitir que otras personas accedan a Amazon Kendra \(p. 659\)](#)
- [Quiero permitir que personas ajenas a mi AWS cuenta para acceder a mis recursos de Amazon Kendra \(p. 659\)](#)

No estoy autorizado a realizar una acción en Amazon Kendra

Si la AWS Management Console le indica que no está autorizado para llevar a cabo una acción, debe ponerse en contacto con su administrador para recibir ayuda. El administrador es la persona que le proporcionó las credenciales de inicio de sesión.

El siguiente ejemplo de error se produce cuando `mateojackson` el usuario intenta usar la consola para ver los detalles de un índice pero no tiene `kendra:DescribeIndex` permisos.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:  
kendra:DescribeIndex on resource: index ARN
```

En este caso, Mateo pide a su administrador que actualice sus políticas de forma que pueda obtener acceso al recurso `index` mediante la acción `kendra:DescribeIndex`.

No estoy autorizado a realizar `iam:PassRole`

Si recibes un error que indica que no estás autorizado a realizar la `iam:PassRole` acción, sus políticas deben actualizarse para permitirle transferir un rol a Amazon Kendra.

Algunos servicios de Servicios de AWS le permiten transferir un rol existente a dicho servicio en lugar de crear un nuevo rol de servicio o uno vinculado al servicio. Para ello, debe tener permisos para transferir el rol al servicio.

El siguiente ejemplo de error se produce cuando un usuario de IAM llamado `marymajor` intenta utilizar la consola para realizar una acción en Amazon Kendra. Sin embargo, la acción requiere que el servicio cuente con permisos que otorga un rol de servicio. Mary no tiene permisos para transferir el rol al servicio.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform: iam:PassRole
```

En este caso, las políticas de Mary se deben actualizar para permitirle realizar la acción `iam:PassRole`.

Si necesita ayuda, póngase en contacto con su administrador de AWS. El administrador es la persona que le proporcionó las credenciales de inicio de sesión.

Soy administrador y quiero permitir que otras personas accedan a Amazon Kendra

Para permitir que otras personas accedan a Amazon Kendra, debe crear una entidad de IAM (usuario o rol) para la persona o la aplicación que necesita acceso. Esta persona utilizará las credenciales de la entidad para acceder a AWS. A continuación, debe adjuntar una política a la entidad que le otorgue los permisos correctos en Amazon Kendra.

Para comenzar de inmediato, consulte [Creación del primer grupo y usuario delegado de IAM](#) en la Guía del usuario de IAM.

Quiero permitir que personas ajenas a mi AWS cuenta para acceder a mis recursos de Amazon Kendra

Puede crear un rol que los usuarios de otras cuentas o las personas externas a la organización puedan utilizar para acceder a sus recursos. Puede especificar una persona de confianza para que asuma el rol. En el caso de los servicios que admitan las políticas basadas en recursos o las listas de control de acceso (ACL), puede utilizar dichas políticas para conceder a las personas acceso a sus recursos.

Para obtener más información, consulte lo siguiente:

- Para saber si Amazon Kendra admite estas funciones, consulte [Cómo funciona Amazon Kendra con IAM \(p. 647\)](#).
- Para obtener información acerca de cómo proporcionar acceso a los recursos de las Cuentas de AWS de su propiedad, consulte [Proporcionar acceso a un usuario de IAM a otra Cuenta de AWS de la que es propietario](#) en la Guía del usuario de IAM.
- Para obtener información acerca de cómo proporcionar acceso a los recursos a Cuentas de AWS de terceros, consulte [Proporcionar acceso a Cuentas de AWS que son propiedad de terceros](#) en la Guía del usuario de IAM.
- Para obtener información sobre cómo proporcionar acceso mediante una identidad federada, consulte [Proporcionar acceso a usuarios autenticados externamente \(identidad federada\)](#) en la Guía del usuario de IAM.

- Para obtener información sobre la diferencia entre los roles y las políticas basadas en recursos para el acceso entre cuentas, consulte [Cómo los roles de IAM difieren de las políticas basadas en recursos](#) en la Guía del usuario de IAM.

Prácticas recomendadas de seguridad

Amazon Kendra ofrece una serie de funciones de seguridad que debe tener en cuenta a la hora de desarrollar e implementar sus propias políticas de seguridad. Las siguientes prácticas recomendadas son directrices generales y no suponen una solución de seguridad completa. Puesto que es posible que estas prácticas recomendadas no sean adecuadas o suficientes para el entorno, considérelas como consideraciones útiles en lugar de como normas.

Aplicar el principio del mínimo privilegio

Amazon Kendra proporciona una política de acceso pormenorizada para las aplicaciones que utilizan IAM roles. Recomendamos que se concedan a las funciones únicamente el conjunto mínimo de privilegios exigido por el trabajo, como cubrir su solicitud y el acceso al destino del registro. También recomendamos auditar los permisos de los trabajos de forma regular y ante cualquier cambio en la solicitud.

Permisos de control de acceso basados en roles (RBAC)

Los administradores deben controlar estrictamente los permisos de control de acceso basado en roles (RBAC) para las aplicaciones de Amazon Kendra.

Registro y monitoreo en Amazon Kendra

La monitorización es una parte importante del mantenimiento de la confiabilidad, la disponibilidad y el rendimiento de las aplicaciones de Amazon Kendra. Para supervisar las llamadas a la API de Amazon Kendra, puede utilizar AWS CloudTrail. Para supervisar el estado de tus trabajos, usa Amazon CloudWatch Registros.

- Amazon CloudWatch Alarms—Uso de CloudWatch Alarms, observas una sola métrica durante un período de tiempo que especifiques. Si la métrica supera una política, CloudWatch las alarmas no invocan acciones cuando una métrica se encuentra en un estado determinado. En su lugar, el estado debe haber cambiado y debe mantenerse durante el número de períodos especificado. Para obtener más información, consulte [Monitorización de Amazon Kendra con Amazon CloudWatch](#) (p. 631).
- AWS CloudTrail Registros—CloudTrail proporciona un registro de las acciones realizadas por un usuario, un rol o un AWS servicio en Amazon Kendra o Amazon Kendra Intelligent Ranking. Uso de la información recopilada por CloudTrail, puede determinar la solicitud que se hizo a Amazon Kendra, la dirección IP desde la que se realizó la solicitud, quién la hizo, cuándo se hizo y detalles adicionales. Para obtener más información, consulte [Registro de llamadas a la API de Amazon Kendra con AWS CloudTrail registros](#) (p. 627) y [Registro de llamadas a la API de clasificación inteligente de Amazon Kendra con AWS CloudTrail registros](#) (p. 629).

Validación de conformidad para Amazon Kendra

Los auditores externos evalúan la seguridad y el cumplimiento de Amazon Kendra como parte de varios programas de cumplimiento de Amazon Kendra. Amazon Kendra cumple con lo siguiente:

- Ley de Portabilidad y Responsabilidad de Seguros Médicos de EE. UU (Health Insurance Portability and Accountability Act, HIPAA).
- Controles de sistema y organización (SOC) 2
- Programa de evaluadores registrados de seguridad de la información (IRAP)
- El Programa federal de gestión de riesgos y autorizaciones (FedRAMP) es moderado en las regiones este/oeste de EE. UU.
- El programa federal de gestión de riesgos y autorizaciones (FedRAMP) ocupa un lugar destacado en AWS GovCloud Región (EE. UU.)

Para obtener una lista de los servicios de AWS en el ámbito de programas de conformidad específicos, consulte [AWS Services in Scope by Compliance Program \(Servicios en el ámbito de programas de conformidad\)](#). Para obtener información general, consulte [Programas de conformidad de AWS](#).

Puede descargar los informes de auditoría de terceros utilizando AWS Artifact. Para obtener más información, consulte [Descarga de informes en AWS Artifact](#).

Su responsabilidad de cumplimiento al utilizar Amazon Kendra está determinada por la confidencialidad de sus datos, los objetivos de cumplimiento de su empresa y las leyes y reglamentos aplicables. AWS proporciona los siguientes recursos para facilitar el cumplimiento:

- [Guías de inicio rápido de seguridad y cumplimiento](#)—Estas guías de implementación analizan las consideraciones arquitectónicas y proporcionan los pasos para implementar entornos de referencia centrados en la seguridad y el cumplimiento en AWS.
- [Documento técnico sobre arquitectura para la seguridad y el cumplimiento de la HIPAA](#)—Este documento técnico describe cómo las empresas pueden utilizar AWS para crear aplicaciones que cumplan con la HIPAA.
- [AWS Recursos de cumplimiento](#)—Esta colección de libros de trabajo y guías puede aplicarse a su sector y ubicación.
- [Evaluación de recursos con reglas](#) en el AWS Config Guía para desarrolladores—La AWS Config servicio evalúa en qué medida las configuraciones de sus recursos cumplen con las prácticas internas, las directrices del sector y las regulaciones.
- [AWS Security Hub](#)—Este AWS servicio proporciona una visión completa de su estado de seguridad en AWS que le ayuda a comprobar el cumplimiento de los estándares y las mejores prácticas del sector de la seguridad.

Resiliencia en Amazon Kendra

La infraestructura global de AWS se compone de regiones y zonas de disponibilidad de AWS. AWS Las regiones proporcionan varias zonas de disponibilidad físicamente independientes y aisladas que se encuentran conectadas mediante redes con un alto nivel de rendimiento y redundancia, además de baja latencia. Con las zonas de disponibilidad, puede diseñar y utilizar aplicaciones y bases de datos que realizan una comutación por error automática entre las zonas sin interrupciones. Las zonas de disponibilidad tienen una mayor disponibilidad, tolerancia a errores y escalabilidad que las infraestructuras tradicionales de centros de datos únicos o múltiples.

Para obtener más información sobre las regiones y zonas de disponibilidad de AWS, consulte [Infraestructura global de AWS](#).

Con AWS Infraestructura global, Amazon Kendra Enterprise Edition es tolerante a errores, escalable y altamente disponible. Actualmente no se permite volver a versiones anteriores de un índice, pero puedes actualizar o volver a crear partes del índice de la siguiente manera:[borrando y sumando](#) las fuentes de datos existentes vuelven a su índice.

Seguridad de infraestructura en Amazon Kendra

Como servicio gestionado, Amazon Kendra está protegido por AWS Seguridad de red global. Para obtener información sobre los servicios de seguridad de AWS y cómo AWS protege la infraestructura, consulte [Seguridad en la nube de AWS](#). Para diseñar su entorno de AWS con las prácticas recomendadas de seguridad de infraestructura, consulte [Protección de la infraestructura](#) en Portal de seguridad de AWS Well-Architected Framework.

Usa AWS llamadas a la API publicadas para acceder a Amazon Kendra a través de la red. Los clientes deben admitir lo siguiente:

- Seguridad de la capa de transporte (TLS). Nosotros exigimos TLS 1.2 y recomendamos TLS 1.3.
- Conjuntos de cifrado con confidencialidad directa total (PFS) tales como DHE (Ephemeral Diffie-Hellman) o ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). La mayoría de los sistemas modernos como Java 7 y posteriores son compatibles con estos modos.

Además, las solicitudes deben estar firmadas mediante un ID de clave de acceso y una clave de acceso secreta que esté asociada a una entidad de seguridad de IAM. También puede utilizar [AWS Security Token Service](#) (AWS STS) para generar credenciales de seguridad temporales para firmar solicitudes.

Configuración y análisis de vulnerabilidades en AWS Identity and Access Management

AWS gestiona las tareas de seguridad básicas, como la aplicación de parches en la base de datos y el sistema operativo (SO) de invitado, la configuración del firewall y la recuperación de desastres. Estos procedimientos han sido revisados y certificados por los terceros pertinentes. Para obtener más detalles, consulte los siguientes recursos de :

- [Modelo de responsabilidad compartida](#)
- AWS: [Descripción general de los procesos de seguridad](#) (libro blanco)

Los siguientes recursos también abordan la configuración y el análisis de vulnerabilidad en AWS Identity and Access Management (IAM):

- [Validación de cumplimiento para AWS Identity and Access Management](#)
- [Mejores prácticas de seguridad y casos de uso en AWS Identity and Access Management.](#)

Cuotas para Amazon Kendra

Regiones admitidas

Para obtener una lista de AWS las regiones en las que Amazon Kendra está disponible, consulte [Amazon Kendra regiones y puntos de conexión](#) en la Referencia general de Amazon Web Services.

Cuotas

Las cuotas de servicio, también denominadas límites, son la cantidad máxima de recursos de servicio para tu AWS cuenta. Para obtener más información, consulte [las cuotas Amazon Kendra de servicio](#) en la Referencia AWS general.

Descripción	Valor predeterminado	Edición	Ajustable
Cantidad máxima de índices por cuenta	5	Developer	Sí
Cantidad máxima de índices por cuenta	5	Enterprise	Sí
Número máximo de orígenes de datos por índice	5	Developer	No
Número máximo de orígenes de datos por índice	50	Enterprise	Sí
Cantidad de texto extraído para un índice en una sola unidad. No puedes añadir unidades adicionales para extraer texto para la Edición para desarrolladores.	3 GB	Developer	No
Cantidad de texto extraído para un índice en una sola unidad. Puedes añadir hasta 100 unidades adicionales para extraer texto para la edición empresarial o simplemente ponerte en contacto con el servicio de asistencia .	30 GB	Enterprise	Sí
Cantidad de consultas por segundo para un	0.05	Developer	No

Descripción	Valor predeterminado	Edición	Ajustable
índice en una sola unidad. No puedes añadir unidades adicionales para las consultas de la Edición para desarrolladores.			
Cantidad de consultas por segundo para un índice en una sola unidad. Puede añadir hasta 100 unidades adicionales para consultas sobre la edición empresarial o simplemente ponerse en contacto con el servicio de asistencia .	0.1	Enterprise	Sí
Número máximo de resultados de búsqueda por consulta. El valor predeterminado es 100. Para obtener más de 100 resultados, solo tienes que ponerte en contacto con el equipo de soporte .	100	Todas las ediciones	Sí
Cantidad máxima de resultados de búsqueda por página	100	Todas las ediciones	Sí
Cantidad máxima de documentos destacados por conjunto de resultados destacados	4	Enterprise	Sí
Número máximo de textos de consulta por conjunto de resultados destacados	49	Enterprise	No
Número máximo de caracteres por texto de consulta en un conjunto de resultados destacados	1 000	Enterprise	Sí
Número máximo de conjuntos de resultados destacados por índice	50	Enterprise	Sí

Descripción	Valor predeterminado	Edición	Ajustable
Número máximo de palabras simbólicas por texto de consulta antes del truncamiento. El valor predeterminado es 30. Para permitir más de 30 palabras, solo tienes que ponerte en contacto con el equipo de soporte .	30	Todas las ediciones	Sí
Tamaño máximo de un solo documento	50 MB	Todas las ediciones	Sí
Cantidad máxima de texto extraído de un solo documento	5 MB	Todas las ediciones	No
Tamaño máximo de lista de grupos de usuarios por atributo de consulta	10	Todas las ediciones	Sí
Tamaño máximo de lista de cadenas por atributo de consulta	10	Todas las ediciones	Sí
Cantidad máxima de atributos personalizados por índice	500	Todas las ediciones	No
Número máximo de preguntas frecuentes por índice	30	Todas las ediciones	Sí
Tamaño máximo de 1 FAQ	5 MB	Todas las ediciones	Sí
Número máximo de resultados devueltos para las preguntas frecuentes	4	Todas las ediciones	Sí
Número máximo de caracteres que se muestran en el resultado de una pregunta de preguntas frecuentes	200	Todas las ediciones	Sí
Número máximo de tesauros por índice	1	Todas las ediciones	No
Tamaño máximo de un archivo de sinónimos	5 MB	Todas las ediciones	Sí

Descripción	Valor predeterminado	Edición	Ajustable
Número máximo de reglas de sinónimos por tesauro	10 000	Todas las ediciones	Sí
Número máximo de sinónimos por término en todos los tesauros de un índice	10	Todas las ediciones	No
Número máximo de sugerencias de consulta devueltas por GetQuerySuggestions llamada	10	Todas las ediciones	Sí
Número máximo de campos/atributos para sugerencias de consulta por llamada GetQuerySuggestions	10	Todas las ediciones	Sí
Cantidad máxima de campos/atributos adicionales para sugerencias de consulta por llamada GetQuerySuggestions	5	Todas las ediciones	Sí
Número máximo de listas bloqueadas por índice	1	Todas las ediciones	No
Tamaño máximo de un archivo de texto de lista de bloques	2 MB	Todas las ediciones	Sí
Número máximo de elementos (palabras o frases) en una lista de bloqueo	20 000	Todas las ediciones	Sí
Cantidad máxima de Amazon Kendra experiencias por índice	50	Todas las ediciones	Sí
Número máximo de sugerencias de consultas con corrección ortográfica que se pueden devolver en una Query llamada a la API.	1	Todas las ediciones	Sí

Descripción	Valor predeterminado	Edición	Ajustable
Cantidad máxima de Rescore solicitudes por segundo para un plan de ejecución de rescore o una sola unidad de capacidad. Puedes añadir hasta 1000 unidades adicionales.	0.01	Enterprise	No
Cantidad máxima de planes de ejecución de rescore por cuenta.	50	Enterprise	Sí
Cantidad máxima de identificadores Title para un documento en una Rescore solicitud.	100	Enterprise	No
Cantidad máxima de identificadores Body para un documento en una Rescore solicitud.	200	Enterprise	No
Cantidad máxima de documentos en una Rescore solicitud.	25	Enterprise	No
Cantidad máxima de documentos por grupo en una Rescore solicitud.	3	Enterprise	No

Para obtener más información sobre las cuotas de Amazon Kendra servicio y solicitar un aumento de cuota, consulte [Cuotas de servicio](#).

Solución de problemas

Esta sección puede ayudarlo a resolver problemas comunes que pueda encontrar al trabajar con Amazon Kendra.

Temas

- [Solución de problemas de fuentes de datos \(p. 668\)](#)
- [Solución de problemas de resultados de búsqueda de documentos \(p. 671\)](#)
- [Solución de problemas generales \(p. 672\)](#)

Solución de problemas de fuentes de datos

Esta sección puede ayudarle a solucionar problemas con Amazon Kendra conectores de fuentes de datos.

Mis documentos no estaban indexados

Al sincronizar su Amazon Kendra indexar con una fuente de datos, es posible que tenga problemas que impidan que los documentos se indexen. La indexación es un proceso de dos pasos. En primer lugar, se comprueba la fuente de datos para ver si hay documentos nuevos y actualizados que se van a indexar y para buscar documentos para eliminarlos del índice. En segundo lugar, a nivel de documento, se accede a cada documento y se indexa.

Puede producirse un error en cualquiera de estos pasos. Los errores a nivel de fuente de datos se notifican en la consola del Historial de ejecución de sincronización sección de la página de detalles de la fuente de datos. El estado del trabajo de sincronización puede ser Tuvo éxito, Incompleto, o Falló. También puede ver la cantidad de documentos indexados y eliminados durante el trabajo. Si el estado es Falló, se muestra un mensaje en Detalles columna.

Los errores a nivel de documento se indican en Amazon CloudWatch Logs. Puede ver los errores utilizando el CloudWatch consola.

Mi trabajo de sincronización ha fallado

Un trabajo de sincronización suele fallar cuando hay un error de configuración en el índice o la fuente de datos. En la consola, encontrará el mensaje de error en Historial de ejecución de sincronización sección de la página de detalles de la fuente de datos, en Detalles columna. Los errores a nivel de documento se indican en Amazon CloudWatch Logs. El mensaje de error proporciona información sobre lo que salió mal. El problema suele ser que el índice o la fuente de datos no tienen la información adecuada IAM permisos. El mensaje de error describe los permisos que faltan. Estos son algunos de los mensajes de error que puede recibir:

`Failed to create log group for job. Please make sure that the IAM role provided has sufficient permissions.`

Si su rol de índice no tiene permiso para usar CloudWatch, la fuente de datos no podrá crear un CloudWatch registro. Si recibes este error, debes añadir CloudWatch permisos para la función de índice.

`Failed to access Amazon S3 file prefix (bucket name) while trying to crawl your metadata files. Please make sure the IAM role (ARN) provided has sufficient permissions.`

Cuando se utiliza un Amazon S3 fuente de datos, Amazon Kendra debe tener permiso para acceder al depósito que contiene los documentos. Debes añadir un permiso para Amazon Kendra para leer el bucket de la fuente de datos IAM rol.

The provided IAM role ([ARN](#)) could not be assumed. Please make sure Amazon Kendra is a trusted entity that is allowed to assume the role.

Amazon Kendra necesita permiso para asumir el índice y la fuente de datos IAM roles. Debe añadir una política de confianza a las funciones con permiso para la acción `:AssumeRole`.

Para las [políticas IAM](#) que Amazon Kendra necesita indexar una fuente de datos, consulte [IAM policies](#).

Mi trabajo de sincronización está incompleto

Los trabajos suelen estar incompletos cuando han completado el proceso a nivel de fuente de datos, pero tienen algún error durante el proceso a nivel de documento. Cuando un trabajo está incompleto, es posible que algunos de los documentos no se hayan indexado correctamente. Para un Amazon S3 fuente de datos, un trabajo incompleto suele deberse a:

- Los metadatos de uno o más documentos no eran válidos.
- Cuando se envían documentos para su indexación pero no se ha enviado al menos un documento.
- Cuando se envían documentos para eliminarlos del índice pero no se ha enviado al menos un documento.

Para solucionar un trabajo de sincronización incompleto, consulte primero su CloudWatch logs.

1. En la columna de detalles, selecciona Ver detalles en CloudWatch.
2. Revise los mensajes de error para ver qué causó el error en el documento.

Mi trabajo de sincronización se ha realizado correctamente, pero no hay documentos indexados

Ocasionalmente, una ejecución de un trabajo de sincronización de índices se marcará como Tuvo éxito pero no hay documentos nuevos o actualizados indexados cuando se espera que lleguen. Las posibles razones incluyen:

- Comprobar CloudWatch Metrics `CloudWatch Documents SubmittedForIndexing Failed` para ver si algún documento no se pudo sincronizar. Comprueba tu CloudWatch logs para obtener más detalles.
- Para un Amazon S3 fuente de datos, puede que haya proporcionado Amazon Kendra el nombre o el prefijo del bucket son incorrectos. Asegúrese de que el balde que Amazon Kendra está utilizando es el que contiene los documentos que se van a indexar.
- Al volver a indexar un documento que no se pudo indexar en un trabajo anterior, Amazon Kendra lo indexará a menos que haya cambiado el documento o su archivo de metadatos asociado.

Tengo problemas de formato de archivo al sincronizar mi fuente de datos

Si tiene problemas de formato de archivo al añadir archivos a la fuente de datos o sincronizar la fuente de datos, asegúrese de que los tipos de documentos sean Amazon Kendra soportado. Para obtener una lista de los tipos de documentos admitidos por Amazon Kendra, véase [Tipos o formatos de documentos](#).

Si está utilizando el `BatchPutDocument` API con archivos de texto plano, especifique `PLAIN_TEXT` como tipo de contenido.

¿Cuánto tiempo lleva sincronizar una fuente de datos?

Si no hay actualizaciones en los documentos, sincronice la hora de unAmazon Kendrael índice aumenta en proporción lineal al número de documentos. Por ejemplo, 1000 documentos sin ninguna actualización tardarían unos cinco minutos en sincronizarse y 2000 documentos sin ninguna actualización tardarían unos 10 minutos. Si hay alguna actualización en los documentos, el tiempo de sincronización aumentará en función de la cantidad de documentos actualizados.

¿Cuánto cuesta sincronizar una fuente de datos?

Al sincronizar el índice, tardará dos minutos en calentarse y activarseAmazon EC2para establecer las conexiones necesarias. No se le cobrará nada durante este proceso. El medidor de uso solo comienza después de que se inicie el trabajo de sincronización. Para obtener más información sobreAmazon Kendraprecios, consulte[Amazon Kendraprecios](#).

Voy a recibir unAmazon EC2error de autorización

Si unAmazon EC2se produce un error de operación no autorizada durante la sincronización de una fuente de datos de nube privada virtual (VPC), es probable que su VPCIAMel rol carece de los permisos necesarios. Compruebe que ellAMel rol que utiliza para la fuente de datos tiene los permisos adjuntos. Para obtener más información, consulte[Nube privada virtualAMpapel](#).

No puedo usar los enlaces del índice de búsqueda para abrir miAmazon S3objetos

TuAmazon KendraEl índice solo puede acceder a los archivos queAmazon S3la fuente de datos le otorga permisos de acceso. Por ejemplo,Amazon Kendranolo puede modificar elAmazon S3permisos que determinan si un objeto debe ser público o estar cifrado. Amazon Kendratampoco tiene los permisos predeterminados para crear o devolver un enlace firmado paraAmazon S3objetos. Si quieres activar la vinculación firmada paraAmazon S3objetos en unAmazon Kendraíndice, tiene dos opciones:

- Puede firmar los resultados de la consulta de índice con el objeto uri de origen antes de devolver el resultado a la página de búsqueda. Para unstep-by-steptutorial de este proceso, consulte[Compartir objetos mediante URL prefirmadas](#).
- Puede anular elAmazon S3obtén metadatos fuente uri y haz que tu servicio esté disponible a través de unCloudFrontred de entrega de contenido (CDN) conectada a unAmazon S3balde. O bien, puede utilizar unAPI Gatewaypunto final de proxy que devuelve una URL prefirmada y la redirige a ella.

Voy a recibir unAccessDeniedAl utilizar un archivo de certificado SSLmensaje de error

Si aparece un error de acceso denegado al utilizar un certificado SSL con su fuente de datos, asegúrese de quelAMel rol tiene permiso para acceder al archivo de certificado SSL en la ubicación especificada. Si el certificado está cifrado con unaAWS KMSclave, tulAMel rol también debe tener permiso para descifrar mediante elAWS KMSclave. Para obtener más información, consulte[Autenticación y control de acceso paraAWS KMS](#).

Recibo un error de autorización al utilizar unSharePointfuente de datos

Si recibes un error de autorización al sincronizar tu índice con unSharePointfuente de datos, confirme que tiene asignada una función de administrador del sitio enSharePoint.

Mi índice no rastrea los documentos de mi fuente de datos de Confluence

Si tuAmazon Kendraindex no rastrea documentos de tu fuente de datos de Confluence durante el proceso de sincronización, confirma que formas parte de los grupos de administradores de Confluence.

Solución de problemas de resultados de búsqueda de documentos

Esta sección puede ayudarle a solucionar problemas en suAmazon Kendraresultados de búsqueda.

Los resultados de mi búsqueda no son relevantes para mi consulta de búsqueda

Si los resultados de la búsqueda parecen irrelevantes, puede ser por los siguientes motivos:

- Resultados conLOWla confianza se incluye en los resultados. Puede filtrar los resultados conLOWconfianza mediante el uso deQueryResultItemesScoreAttributecampo para excluir cualquier resultado con un valor deLOW.Amazon Kendraasigna a cada resultado un valor de segmento de confianza de cualquiera de los dosVERY_HIGH,HIGH,MEDIUMyLOW. Estos valores indican el nivel de confianza de que un resultado es relevante para una consulta. Además, independientemente de los grupos de confianza,Amazon Kendradevuelve tres tipos de resultados en el siguiente orden:ANSWER(extracto de respuesta sugerido),QUESTION_ANSWER(Preguntas frecuentes) yDOCUMENT(extracto del documento). Por lo tanto, es posible queLOWconfidenciaQUESTION_ANSWERresultado para posicionarse por encima de unVERY_HIGHconfidenciaDOCUMENTresultado. Sin embargo, no siempre es necesariamente cierto queLOWconfidenciaQUESTION_ANSWERes un resultado mejor que elVERY_HIGHconfidenciaDOCUMENT.
- Algunos campos o atributos de metadatos se elevan a un valor muy alto, lo que afecta a la clasificación de los resultados.Amazon Kendrabusca en el índice mediante varios parámetros, como el título del documento, el texto, la fecha y los campos o atributos de texto personalizados. Puedes experimentar con diferentes valores de refuerzo para obtener los mejores resultados en todas las consultas. También puedes usar dinámica[ajuste de relevancia](#)en el nivel de consulta para utilizar diferentes valores de refuerzo para cada consulta.
- Los usuarios utilizan términos especializados cuando solicitan información y no hay sinónimos personalizados configurados en el índice para gestionar estos términos especializados. Para obtener más información sobre cómo y cuándo usar sinónimos, consulte[Añadir sinónimos personalizados a un índice](#).

¿Por qué solo veo 100 resultados?

Amazon Kendra devuelve el recuento total de los documentos pertinentes. De forma predeterminada, se devuelven los 100 primeros por consulta. Los resultados están paginados. Puedes usar `PageNumber` para acceder a diferentes páginas.

Puede configurar Amazon Kendra para devolver hasta 1000 documentos o resultados de búsqueda por consulta, con hasta 100 resultados por página. Para obtener más de 100 resultados, puede solicitarlo comunicándose con [Soporte de cuotas](#). Aumentar el número de resultados de búsqueda podría afectar a la latencia.

¿Por qué faltan los documentos que espero ver?

Amazon Kendra admite listas de control de acceso (ACL) basadas en usuarios y grupos. Amazon Kendra implementa las políticas de ACL a través de conectores. Si un índice no configura una ACL, solo se mostrarán los documentos que coincidan con el filtro de atributos para el usuario y el grupo. Si se proporciona un filtro de atributos de usuario o grupo, no se mostrarán los documentos sin una ACL.

Si utiliza un control de acceso basado en tokens, se mostrarán los documentos sin una política de ACL y los documentos que coincidan con el usuario y los grupos.

¿Por qué veo documentos que tienen una política de ACL?

Si un índice no configura una política de control de acceso, el filtro puede proporcionar usuarios y grupos. Si no se aplica ningún filtro de usuario y grupo, se devolverán todos los documentos relacionados. Se ignorará cualquier política de ACL.

Solución de problemas generales

Amazon Kendra usa CloudWatch métricas y registros para proporcionar información sobre la sincronización de las fuentes de datos. Puedes usar las métricas y los registros para determinar qué ha fallado en una ejecución de sincronización y cómo solucionarlo.

Para solucionar problemas generales, comience con su CloudWatch métricas.

- Compruebe el `DocumentsCrawled` métrica para ver cuántos documentos ha comprobado su fuente de datos. Para un Amazon S3 depósito, si el número es inferior al esperado, compruebe que la fuente de datos apunta al depósito correcto.
- Compruebe el `DocumentsSkippedNoChange` métrica para ver cuántos documentos se omitieron porque no han cambiado desde la última sincronización. Si el número no coincide con lo esperado, comprueba que tu repositorio se haya actualizado correctamente.
- Compruebe el `DocumentsSkippedInvalidMetadata` métrica para ver cuántos documentos tenían metadatos no válidos. Comprueba tu CloudWatch regis tra para ver los errores específicos que se han producido.
- Compruebe el `DocumentsSubmittedForIndexingFailed` métrica para ver cuántos documentos se enviaron desde la fuente de datos al índice pero no se pudieron indexar. Por ejemplo, si usa un atributo de metadatos en un Amazon S3 fuente de datos que no se haya definido como un campo de índice personalizado, el documento no se indexará. Comprueba tu CloudWatch regis tra para ver los errores específicos que se han producido.
- Compruebe el `DocumentsSubmittedForDeletionFailed` métrica para ver cuántos documentos que la fuente de datos intentó eliminar del índice no se pudieron eliminar del índice. Comprueba tu CloudWatch regis tra para ver los errores específicos que se han producido.

Puedes mirar el CloudWatch registra una ejecución de sincronización determinada para obtener detalles de los errores que se produjeron durante la ejecución. Para obtener más información sobre CloudWatch registros con Amazon Kendra, consulte [CloudWatch Logs](#).

Amazon KendraClasificación inteligente

Amazon KendraLa clasificación inteligente utiliza funciones de búsqueda Amazon Kendra semántica para volver a clasificar de forma inteligente los resultados de un servicio de búsqueda.

Temas

- [Amazon KendraClasificación inteligente para autogestionados OpenSearch \(p. 674\)](#)
- [Clasificación semántica de los resultados de un servicio de búsqueda \(p. 683\)](#)

Amazon KendraClasificación inteligente para autogestionados OpenSearch

Puede aprovechar las capacidades Amazon Kendra de búsqueda semántica para mejorar los resultados de búsqueda desde [OpenSearchel](#) servicio de búsqueda de código abierto autogestionado basado en la licencia Apache 2.0. El complemento Amazon Kendra Intelligent Ranking vuelve a clasificar semánticamente los resultados OpenSearch utilizando. Amazon Kendra Para ello, comprende el significado y el contexto de una consulta de búsqueda mediante campos específicos, como el cuerpo o el título del documento, de los resultados de OpenSearch búsqueda predeterminados.

Tomemos, por ejemplo, esta consulta: «discurso principal». Dado que «dirección» tiene varios significados, Amazon Kendra puede inferir el significado detrás de la consulta para devolver información relevante alineada con el significado deseado. En este contexto, es el discurso de apertura de una conferencia. Un servicio de búsqueda más sencillo podría no tener en cuenta la intención y podría arrojar resultados para una dirección postal en Main Street, por ejemplo.

El complemento Intelligent Ranking for OpenSearch está disponible para la versión 2.4.0 OpenSearch (autogestionada) y versiones posteriores. Puede instalar el complemento mediante un script Bash de inicio rápido para crear una nueva imagen de Docker OpenSearch con el complemento Intelligent Ranking incluido. Consulte[Configuración del complemento de búsqueda inteligente \(p. 675\)](#): este es un ejemplo de una configuración que le permitirá empezar a trabajar rápidamente.

Cómo funciona el complemento de búsqueda inteligente

El proceso general del complemento Intelligent Ranking para OpenSearch (autogestionado) es el siguiente:

1. Un OpenSearch usuario emite una consulta y OpenSearch proporciona una respuesta a la consulta o una lista de documentos relevantes para la consulta.
2. El complemento Intelligent Ranking toma la respuesta a la consulta y extrae la información de los documentos.
3. El complemento Intelligent Ranking hace una llamada a la API [Rescore](#) de Amazon Kendra Intelligent Ranking.
4. La Rescore API toma la información extraída de los documentos y vuelve a clasificar semánticamente los resultados de la búsqueda.
5. La Rescore API envía los resultados de búsqueda reclasificados al complemento. El complemento reorganiza los resultados de la búsqueda en la respuesta de OpenSearch búsqueda para reflejar la nueva clasificación semántica.

El complemento Intelligent Ranking vuelve a clasificar los resultados mediante los campos «cuerpo» y «título». Estos campos del complemento se pueden asignar a los campos del OpenSearch índice que mejor se ajusten a la definición del cuerpo y el título de un documento. Por ejemplo, si el índice contiene capítulos de un libro con campos como «encabezado del capítulo» y «contenido del capítulo», puede asignar el primero a «título» y el segundo a «cuerpo» para obtener los mejores resultados.

Configuración del complemento de búsqueda inteligente

A continuación se describe cómo configurar rápidamente OpenSearch (autogestionado) con el complemento Intelligent Ranking.

Configuración OpenSearch (autogestionada) con el complemento Intelligent Ranking (configuración rápida)

Si ya está utilizando una imagen de Dockeropensearch:2.4.0, puede usar este [Dockerfile](#) para crear una nueva imagen de OpenSearch 2.4.0 con el complemento Intelligent Ranking. Incluye un contenedor para la nueva imagen en el archivo [docker-compose.yml](#) o en el [archivo opensearch.yml](#). También incluye el ID del plan de ejecución de rescore generado al crear un plan de ejecución de rescore, junto con la información de tu región y punto final; consulta el paso 2 para crear un plan de ejecución de rescore.

Si ya descargaste una versión de la imagen de opensearch Docker anterior a la 2.4.0, debes usar la imagen de Docker opensearch:2.4.0 o una versión posterior y crear una nueva imagen con el complemento Intelligent Ranking incluido.

1. Descargue e instale [Docker Desktop](#) para su sistema operativo. Docker Desktop incluye Docker Compose y Docker Engine. Se recomienda comprobar si el equipo cumple con los requisitos del sistema mencionados en los detalles de instalación de Docker.

También puede aumentar sus requisitos de uso de memoria en la configuración de su escritorio Docker. Usted es responsable de los requisitos de uso de Docker fuera de los límites de uso disponibles gratuitamente para los servicios de Docker. Consulte [Suscripciones a Docker](#).

Compruebe que el estado de Docker Desktop esté «en ejecución».

2. Aproveche la clasificación Amazon Kendra inteligente y sus requisitos [de capacidad](#). Una vez Amazon Kendra aprovisione la Clasificación inteligente, se le cobrará por hora en función de las unidades de capacidad establecidas. Consulte la [información sobre precios y niveles gratuitos](#).

Utiliza la [CreateRescoreExecutionPlanAPI](#) para aprovisionar elRescore API. Si no necesitas más unidades de capacidad que las predeterminadas de una sola unidad, no agregues más unidades y solo proporciones un nombre para tu plan de ejecución de rescore. También puedes actualizar tus requisitos de capacidad mediante la [UpdateRescoreExecutionPlanAPI](#). Para obtener más información, consulta [Clasificar semánticamente los resultados de un servicio de búsqueda](#).

Si lo desea, puede ir al paso 3 para crear un plan de ejecución de rescore predeterminado al ejecutar el script Bash de inicio rápido.

Para el paso 4, anote el ID del plan de ejecución de rescore incluido en la respuesta.

CLI

```
aws kendra-ranking create-rescore-execution-plan \
--name MyRescoreExecutionPlan \
--capacity-units '{"RescoreCapacityUnits":<integer number of additional capacity units>}'
```

Response:

```
{  
    "Id": "<rescore execution plan ID>",  
    "Arn": "arn:aws:kendra-ranking:<region>:<account-id>:rescore-execution-plan/<rescore-execution-plan-id>"  
}
```

Python

```
import boto3  
from botocore.exceptions import ClientError  
import pprint  
import time  
  
kendra_ranking = boto3.client("kendra-ranking")  
  
print("Create a rescore execution plan.")  
  
# Provide a name for the rescore execution plan  
name = "MyRescoreExecutionPlan"  
# Set your required additional capacity units  
# Don't set capacity units if you don't require more than 1 unit given by default  
capacity_units = 1  
  
try:  
    rescore_execution_plan_response = kendra_ranking.create_rescore_execution_plan(  
        Name = name,  
        CapacityUnits = {"RescoreCapacityUnits":capacity_units}  
    )  
  
    pprint.pprint(rescore_execution_plan_response)  
    rescore_execution_plan_id = rescore_execution_plan_response["Id"]  
  
    print("Wait for Amazon Kendra to create the rescore execution plan.")  
  
    while True:  
        # Get the details of the rescore execution plan, such as the status  
        rescore_execution_plan_description =  
            kendra_ranking.describe_rescore_execution_plan(  
                Id = rescore_execution_plan_id  
            )  
        # When status is not CREATING quit.  
        status = rescore_execution_plan_description["Status"]  
        print(" Creating rescore execution plan. Status: "+status)  
        time.sleep(60)  
        if status != "CREATING":  
            break  
  
    except ClientError as e:  
        print("%s" % e)  
  
print("Program ends.")
```

3. Descargue el [script de inicio rápido de Bash](#) GitHub para su versión de OpenSearch seleccionando la rama de versión en el menú desplegable de la rama principal.
4. Abre tu terminal y, en el directorio del script de Bash, ejecuta el siguiente comando.

```
bash search_processing_kendra_quickstart.sh -p <execution-plan-id> -r <region>
```

Al ejecutar este comando, proporciona el ID del plan de ejecución de rescore que anotó en el paso 2 al aprovisionar Amazon Kendra Intelligent Ranking, junto con la información de su región. Si lo desea, puede aprovisionar la Clasificación Amazon Kendra inteligente en su lugar mediante la --create-

execution-plan opción. Esto crea un plan de ejecución de rescore con un nombre y una capacidad predeterminados. Si AWS las credenciales no se pueden obtener de las variables de entorno, el perfil predeterminado o la Amazon EC2 instancia, proporcione sus credenciales mediante la --profile opción.

Para no perder el índice cuando se elimina el contenedor efímero predeterminado, puede hacer que el índice persista en todas las ejecuciones proporcionando el nombre del volumen de datos mediante la --volume-name opción. Si ya creó un índice, puede especificar el volumen en el archivo docker-compose.yml u opensearch.yml.

Este script usa imágenes de Docker OpenSearch y OpenSearch paneles con la versión que seleccionó en el GitHub repositorio para el script. Descarga un archivo zip para el complemento Intelligent Ranking y genera una imagen nueva de Docker Dockerfile para crear una nueva imagen de Docker OpenSearch que incluya el complemento. También crea un archivo [docker-compose.yml](#) que incluye contenedores para el complemento Intelligent Ranking y los paneles OpenSearch de control. OpenSearch El script agrega el ID del plan de ejecución de Rescore, la información de la región y el punto final (usa la región) al archivo docker-compose.yml. A continuación, el script se ejecuta docker-compose up para iniciar los contenedores OpenSearch con la clasificación inteligente incluida y los OpenSearch paneles de control. Para detener los recipientes sin retirarlos, corradocker-compose stop. Para retirar los contenedores, corradocker-compose down. Para dejar sus volúmenes intactos, no los ejecutedocker-compose down -v.

Ejemplo de docker-compose.yml

Un ejemplo de un archivo docker-compose.yml que usa la OpenSearch versión 2.4.0 o posterior con el complemento Intelligent Ranking y Dashboards 2.4.0 o versiones posteriores. OpenSearch

```
version: '3'
networks:
  opensearch-net:
volumes:
<volume-name>:
services:
  opensearch-node:
    image: <Docker image tag name of OpenSearch with Intelligent Ranking plugin>
    container_name: opensearch-node
    environment:
      - cluster.name=opensearch-cluster
      - node.name=opensearch-node
      - discovery.type=single-node
      - kendra_intelligent_ranking.service.endpoint=https://kendra-ranking.<region>.api.aws
      - kendra_intelligent_ranking.service.region=<region>
      - kendra_intelligent_ranking.service.execution_plan_id=<rescore-execution-plan-id>
    ulimits:
      memlock:
        soft: -1
        hard: -1
      nofile:
        soft: 65536
        hard: 65536
    ports:
      - 9200:9200
      - 9600:9600
    networks:
      - opensearch-net
  volumes:
    <docker-volume-name>:/usr/share/opensearch/data
  opensearch-dashboard:
    image: opensearchproject/opensearch-dashboards:<your-version>
    container_name: opensearch-dashboards
```

```
ports:  
  - 5601:5601  
environment:  
  OPENSEARCH_HOSTS: '['https://opensearch-node:9200"]'  
networks:  
  - opensearch-net
```

Ejemplo de un Dockerfile y creación de una imagen

Un ejemplo de uso de Dockerfile la OpenSearch versión 2.4.0 o posterior con el complemento Intelligent Ranking.

```
FROM opensearchproject/opensearch:<your-version>  
RUN /usr/share/opensearch/bin/opensearch-plugin install --batch https://github.com/  
opensearch-project/search-processor/releases/download/<your-version>/search-processor.zip
```

Creación de una imagen de Docker para OpenSearch con el complemento Intelligent Ranking.

```
docker build --tag=<Docker image tag name of OpenSearch with Intelligent Ranking plugin>
```

Interactuar con el complemento de búsqueda inteligente

Una vez que lo haya configurado OpenSearch (autogestionado) con el complemento Intelligent Ranking, podrá interactuar con el complemento mediante los comandos curl o las bibliotecas OpenSearch cliente. Las credenciales predeterminadas para acceder OpenSearch con el complemento Intelligent Ranking son el nombre de usuario «admin» y la contraseña «admin».

Para aplicar la configuración del complemento Intelligent Ranking a un OpenSearch índice:

Curl

```
curl -XPUT "https://localhost:9200/<your-docs-index>/_settings" -u 'admin:admin' --  
insecure -H 'Content-Type: application/json' -d'  
{  
  "index": {  
    "plugin" : {  
      "searchrelevance" : {  
        "result_transformer" : {  
          "kendra_intelligent_ranking": {  
            "order": 1,  
            "properties": {  
              "title_field": "title_field_name_here",  
              "body_field": "body_field_name_here"  
            }  
          }  
        }  
      }  
    }  
  }  
}
```

Python

```
pip install opensearch-py
```

```
from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

setting_body = {
    "index": {
        "plugin" : {
            "searchrelevance" : {
                "result_transformer" : {
                    "kendra_intelligent_ranking": {
                        "order": 1,
                        "properties": {
                            "title_field": "title_field_name_here",
                            "body_field": "body_field_name_here"
                        }
                    }
                }
            }
        }
    }
}

response = client.indices.put_settings(index_name, body=setting_body)
```

Debe incluir el nombre del campo de texto principal que desea utilizar para cambiar la clasificación, como el cuerpo del documento o el campo de contenido del documento. También puede incluir otros campos de texto, como el título del documento o el resumen del documento.

Ahora puede realizar cualquier consulta y los resultados se clasifican con el complemento Intelligent Ranking.

Curl

```
curl -XGET "https://localhost:9200/<your-docs-index>/_search?pretty" -u 'admin:admin' --insecure -H 'Content-Type: application/json' -d'
{
    "query" : {
        "match" : {
            "body_field_name_here": "intelligent systems"
        }
    }
}'
```

Python

```
from opensearchpy import OpenSearch
```

```
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

query = {
    'size': 10,
    "query" : {
        "match" : {
            "body_field_name_here": "intelligent systems"
        }
    }
}

response = client.search(
    body = query,
    index = index_name
)

print('\nSearch results:')
print(response)
```

Para eliminar la configuración del complemento Intelligent Ranking de un OpenSearch índice, sigue estos pasos:

Curl

```
curl -XPUT "http://localhost:9200/<your-docs-index>/_settings" -H 'Content-Type: application/json' -d'
{
  "index": {
    "plugin": {
      "searchrelevance": {
        "result_transformer": {
          "kendra_intelligent_ranking.*": null
        }
      }
    }
  }
}'
```

Python

```
from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
auth = ('admin', 'admin')

client = OpenSearch(
```

```
hosts = [{'host': host, 'port': port}],
http_compress = True, # enables gzip compression for request bodies
http_auth = auth,
# client_cert = client_cert_path,
# client_key = client_key_path,
use_ssl = True,
verify_certs = False,
ssl_assert_hostname = False,
ssl_show_warn = False,
ca_certs = ca_certs_path
)

setting_body = {
    "index": {
        "plugin": {
            "searchrelevance": {
                "result_transformer": {
                    "kendra_intelligent_ranking.*": null
                }
            }
        }
    }
}

response = client.indices.put_settings(index_name, body=setting_body)
```

Para probar el complemento Intelligent Ranking en una consulta determinada o en ciertos campos del cuerpo y el título:

Curl

```
curl -XGET "https://localhost:9200/<your-docs-index>/_search?pretty" -u 'admin:admin'
--insecure -H 'Content-Type: application/json' -d'
{
    "query": {
        "multi-match": {
            "query": "intelligent systems",
            "fields": ["body_field_name_here", "title_field_name_here"]
        }
    },
    "size": 25,
    "ext": {
        "search_configuration": {
            "result_transformer": {
                "kendra_intelligent_ranking": {
                    "order": 1,
                    "properties": {
                        "title_field": "title_field_name_here",
                        "body_field": "body_field_name_here"
                    }
                }
            }
        }
    }
}'
```

Python

```
from opensearchpy import OpenSearch
host = 'localhost'
port = 9200
```

```
auth = ('admin', 'admin')

client = OpenSearch(
    hosts = [{'host': host, 'port': port}],
    http_compress = True, # enables gzip compression for request bodies
    http_auth = auth,
    # client_cert = client_cert_path,
    # client_key = client_key_path,
    use_ssl = True,
    verify_certs = False,
    ssl_assert_hostname = False,
    ssl_show_warn = False,
    ca_certs = ca_certs_path
)

# Index settings null for kendra_intelligent_ranking

query = {
    "query": {
        "multi_match": {
            "query": "intelligent systems",
            "fields": ["body_field_name_here", "title_field_name_here"]
        }
    },
    "size": 25,
    "ext": {
        "search_configuration": {
            "result_transformer": {
                "kendra_intelligent_ranking": {
                    "order": 1,
                    "properties": {
                        "title_field": "title_field_name_here",
                        "body_field": "body_field_name_here"
                    }
                }
            }
        }
    }
}

response = client.search(
    body = query,
    index = index_name
)

print('\nSearch results:')
print(response)
```

Comparar OpenSearch los resultados con Amazon Kendra los resultados

Puede comparar los resultados clasificados side-by-side OpenSearch (autogestionados) con Amazon Kendra los resultados reclasificados. OpenSearchLa versión 2.4.0 y versiones posteriores de Dashboards ofrecen side-by-side resultados para que pueda comparar la forma OpenSearch en que se clasifican los documentos con la forma Amazon Kendra en que el complemento clasifica los documentos para una consulta de búsqueda.

Antes de poder comparar los resultados OpenSearch clasificados con los resultados Amazon Kendra reclasificados, asegúrese de que sus OpenSearch paneles estén respaldados por un OpenSearch servidor con el complemento Clasificación inteligente. Puede configurar esto usando Docker y un script Bash de inicio rápido. Consulte [Configuración del complemento de búsqueda inteligente \(p. 675\)](#).

A continuación se describe cómo comparar OpenSearch y Amazon Kendra buscar los resultados en los OpenSearch paneles. Para obtener más información, consulte la [documentación de OpenSearch](#).

Comparación de los resultados de búsqueda en los OpenSearch paneles

1. Abra `http://localhost:5601` e inicie sesión en los OpenSearch paneles. Las credenciales predeterminadas son el nombre de usuario «admin» y la contraseña «admin».
2. Selecciona Relevancia de búsqueda en los OpenSearch complementos del menú de navegación.
3. Introduzca el texto de búsqueda en la barra de búsqueda.
4. Seleccione su índice para la consulta 1 e introduzca una consulta en el DSL de OpenSearch consultas. Puede utilizar la `%SearchText%` variable para hacer referencia al texto de búsqueda que ha introducido en la barra de búsqueda. Para ver un ejemplo de esta consulta, consulte [OpenSearchla documentación](#). Los resultados devueltos para esta consulta son los OpenSearch resultados sin utilizar el complemento Intelligent Ranking.
5. Seleccione el mismo índice para la consulta 2 e introduzca la misma consulta en el DSL de OpenSearch consultas. Además, incluye la extensión con `kendra_intelligent_ranking` y especifica el obligatorio en el `body_field` que se debe clasificar. También puede especificar el campo de título, pero el campo cuerpo es obligatorio. Para ver un ejemplo de esta consulta, consulte [OpenSearchla documentación](#). Los resultados devueltos para esta consulta son los resultados Amazon Kendra reclasificados mediante el complemento Intelligent Ranking. El complemento clasifica hasta 25 resultados.
6. Selecciona Buscar para obtener y comparar los resultados.

Clasificación semántica de los resultados de un servicio de búsqueda

Amazon KendraIntelligent Ranking utiliza las capacidades Amazon Kendra de búsqueda semántica de un servicio de búsqueda para volver a clasificar los resultados de un servicio de búsqueda. Para ello, tiene en cuenta el contexto de la consulta de búsqueda, además de toda la información disponible en los documentos del servicio de búsqueda. Amazon Kendra La clasificación inteligente puede mejorar la coincidencia simple de palabras clave.

La [CreateRescoreExecutionPlanAPI](#) crea un recurso de clasificación Amazon Kendra inteligente que se utiliza para aprovisionar la API [Rescore](#). La Rescore API vuelve a clasificar los resultados de búsqueda de un servicio de búsqueda como [OpenSearch\(autogestionado\)](#).

Cuando llamas `CreateRescoreExecutionPlan`, configuras las unidades de capacidad necesarias para volver a clasificar los resultados de un servicio de búsqueda. Si no necesitas más unidades de capacidad más allá del valor predeterminado de una sola unidad, no cambies el valor predeterminado. Proporcione solo un nombre para su plan de ejecución de rescore. Puedes configurar hasta 1000 unidades adicionales. Para obtener información sobre lo que se incluye en una sola unidad de capacidad, consulte [Ajustar la capacidad](#). Una vez Amazon Kendra aprovisione la Clasificación inteligente, se le cobrará por hora en función de las unidades de capacidad establecidas. Consulte la [información sobre precios y niveles gratuitos](#).

Se genera un ID del plan de ejecución de rescore que se devuelve en la respuesta cuando se llama `CreateRescoreExecutionPlan`. La Rescore API utiliza el ID del plan de ejecución de rescore para volver a clasificar los resultados de un servicio de búsqueda con la capacidad que hayas establecido. Incluya el ID del plan de ejecución de rescore en los archivos de configuración de su servicio de búsqueda. [Por ejemplo, si utilizas OpenSearch \(autogestionado\), incluyes el ID del plan de ejecución de Rescore en tu archivo docker-compose.yml u opensearch.yml; consulta los resultados de clasificación inteligente \(autoservicio\). OpenSearch](#)

Cuando llama, también se genera un nombre de recurso de Amazon (ARN) en la respuesta `CreateRescoreExecutionPlan`. Puede usar este ARN para crear una política de permisos en

AWS Identity and Access Management (IAM) que restrinja el acceso de los usuarios a un ARN específico para un plan de ejecución de rescore específico.

El siguiente es un ejemplo de cómo crear un plan de ejecución de rescore con unidades de capacidad configuradas en 1.

CLI

```
aws kendra-ranking create-rescore-execution-plan \
--name MyRescoreExecutionPlan \
--capacity-units '{"RescoreCapacityUnits":1}'
```

Response:

```
{ "Id": "<rescore execution plan ID>", "Arn": "arn:aws:kendra-ranking:<region>:<account-id>:rescore-execution-plan/<rescore-execution-plan-id>" }
```

Python

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

kendra_ranking = boto3.client("kendra-ranking")

print("Create a rescore execution plan.")

# Provide a name for the rescore execution plan
name = "MyRescoreExecutionPlan"
# Set your required additional capacity units
# Don't set capacity units if you don't require more than 1 unit given by default
capacity_units = 1

try:
    rescore_execution_plan_response = kendra_ranking.create_rescore_execution_plan(
        Name = name,
        CapacityUnits = {"RescoreCapacityUnits":capacity_units}
    )

    pprint.pprint(rescore_execution_plan_response)

    rescore_execution_plan_id = rescore_execution_plan_response["Id"]

    print("Wait for Amazon Kendra to create the rescore execution plan.")

    while True:
        # Get the details of the rescore execution plan, such as the status
        rescore_execution_plan_description =
kendra_ranking.describe_rescore_execution_plan(
            Id = rescore_execution_plan_id
        )
        # When status is not CREATING quit.
        status = rescore_execution_plan_description["Status"]
        print(" Creating rescore execution plan. Status: "+status)
        time.sleep(60)
        if status != "CREATING":
            break

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

```
print("Program ends.")
```

Java

```
import java.util.concurrent.TimeUnit;

import software.amazon.awssdk.services.kendraranking.KendraRankingClient;
import software.amazon.awssdk.services.kendraranking.model.CapacityUnitsConfiguration;
import
    software.amazon.awssdk.services.kendraranking.model.CreateRescoreExecutionPlanRequest;
import
    software.amazon.awssdk.services.kendraranking.model.CreateRescoreExecutionPlanResponse;
import
    software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanRequest;
import
    software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanResponse;
import software.amazon.awssdk.services.kendraranking.model.RescoreExecutionPlanStatus;

public class CreateRescoreExecutionPlanExample {

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

        String rescoreExecutionPlanName = "MyRescoreExecutionPlan";
        int capacityUnits = 1;

        KendraRankingClient kendraRankingClient = KendraRankingClient.builder().build();

        System.out.println(String.format("Creating a rescore execution plan named %s",
            rescoreExecutionPlanName));

        CreateRescoreExecutionPlanResponse createResponse =
kendraRankingClient.createRescoreExecutionPlan(
            CreateRescoreExecutionPlanRequest.builder()
                .name(rescoreExecutionPlanName)
                .capacityUnits(
                    CapacityUnitsConfiguration.builder()
                        .rescoreCapacityUnits(capacityUnits)
                        .build()
                )
                .build()
        );
        String rescoreExecutionPlanId = createResponse.id();
        System.out.println(String.format("Waiting for rescore execution plan with id %s to
            finish creating.", rescoreExecutionPlanId));
        while (true) {
            DescribeRescoreExecutionPlanResponse describeResponse =
kendraRankingClient.describeRescoreExecutionPlan(
                DescribeRescoreExecutionPlanRequest.builder()
                    .id(rescoreExecutionPlanId)
                    .build()
            );
            RescoreExecutionPlanStatus rescoreExecutionPlanStatus =
describeResponse.status();
            if (rescoreExecutionPlanStatus != RescoreExecutionPlanStatus.CREATING) {
                break;
            }
            TimeUnit.SECONDS.sleep(60);
        }

        System.out.println("Rescore execution plan creation is complete.");
    }
}
```

El siguiente es un ejemplo de actualización de un plan de ejecución de rescore para establecer las unidades de capacidad en 2.

CLI

```
aws kendra-ranking update-rescore-execution-plan \
--id <rescore execution plan ID> \
--capacity-units '{"RescoreCapacityUnits":2}'
```

Python

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

kendra_ranking = boto3.client("kendra-ranking")

print("Update a rescore execution plan.")

# Provide the ID of the rescore execution plan
id = <rescore execution plan ID>
# Re-set your required additional capacity units
capacity_units = 2

try:
    kendra_ranking.update_rescore_execution_plan(
        Id = id,
        CapacityUnits = {"RescoreCapacityUnits":capacity_units}
    )

    print("Wait for Amazon Kendra to update the rescore execution plan.")

    while True:
        # Get the details of the rescore execution plan, such as the status
        rescore_execution_plan_description =
kendra_ranking.describe_rescore_execution_plan(
            Id = id
        )
        # When status is not UPDATING quit.
        status = rescore_execution_plan_description["Status"]
        print(" Updating rescore execution plan. Status: "+status)
        time.sleep(60)
        if status != "UPDATING":
            break

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

print("Program ends.")
```

Java

```
import java.util.concurrent.TimeUnit;

import software.amazon.awssdk.services.kendraranking.KendraRankingClient;
import software.amazon.awssdk.services.kendraranking.model.CapacityUnitsConfiguration;
import
software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanRequest;
import
software.amazon.awssdk.services.kendraranking.model.DescribeRescoreExecutionPlanResponse;
import software.amazon.awssdk.services.kendraranking.model.RescoreExecutionPlanStatus;
```

```
import software.amazon.awssdk.services.kendraranking.model.UpdateRescoreExecutionPlanRequest;
import software.amazon.awssdk.services.kendraranking.model.UpdateRescoreExecutionPlanResponse;

public class UpdateRescoreExecutionPlanExample {

    public static void main(String[] args) throws InterruptedException {
        String rescoreExecutionPlanId = <rescore execution plan ID>;
        int newCapacityUnits = 2;

        KendraRankingClient kendraRankingClient = KendraRankingClient.builder().build();

        System.out.println(String.format("Updating a rescore execution plan named %s",
                rescoreExecutionPlanId));

        UpdateRescoreExecutionPlanResponse updateResponse =
                kendraRankingClient.updateRescoreExecutionPlan(
                        UpdateRescoreExecutionPlanRequest.builder()
                                .id(rescoreExecutionPlanId)
                                .capacityUnits(
                                        CapacityUnitsConfiguration.builder()
                                                .rescoreCapacityUnits(newCapacityUnits)
                                                .build()
                                )
                                .build()
                );
        System.out.println(String.format("Waiting for rescore execution plan with id %s to
                finish updating.", rescoreExecutionPlanId));
        while (true) {
            DescribeRescoreExecutionPlanResponse describeResponse =
                    kendraRankingClient.describeRescoreExecutionPlan(
                            DescribeRescoreExecutionPlanRequest.builder()
                                    .id(rescoreExecutionPlanId)
                                    .build()
                    );
            RescoreExecutionPlanStatus rescoreExecutionPlanStatus =
                    describeResponse.status();
            if (rescoreExecutionPlanStatus != RescoreExecutionPlanStatus.UPDATING) {
                break;
            }
            TimeUnit.SECONDS.sleep(60);
        }
        System.out.println("Rescore execution plan update is complete.");
    }
}
```

El siguiente es un ejemplo del uso de la Rescore API.

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_response["ResultItems"])

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

print("Program ends.")
```

Java

```
import java.util.ArrayList;
import java.util.List;

import software.amazon.awssdk.services.kendraranking.KendraRankingClient;
import software.amazon.awssdk.services.kendraranking.model.RescoreRequest;
import software.amazon.awssdk.services.kendraranking.model.RescoreResponse;
import software.amazon.awssdk.services.kendraranking.model.Document;

public class RescoreExample {

    public static void main(String[] args) {

        String rescoreExecutionPlanId = <rescore execution plan ID>;
        String query = "intelligent systems";

        List<Document> documentList = new ArrayList<>();
        documentList.add(
            Document.builder()
                .id("DocId1")
                .originalScore(2.0F)
                .body("intelligent systems in everyday life")
                .title("Smart systems")
                .build()
        );
        documentList.add(
```

```
Document.builder()  
    .id("DocId2")  
    .originalScore(1.0F)  
    .body("living with intelligent systems")  
    .title("Smarter systems")  
    .build()  
);  
  
KendraRankingClient kendraRankingClient = KendraRankingClient.builder().build();  
  
RescoreResponse rescoreResponse = kendraRankingClient.rescore(  
    RescoreRequest.builder()  
        .rescoreExecutionPlanId(rescoreExecutionPlanId)  
        .searchQuery(query)  
        .documents(documentList)  
        .build()  
);  
  
System.out.println(rescoreResponse.rescoreId());  
System.out.println(rescoreResponse.resultItems());  
}  
}
```

Historial de revisión de Amazon Kendra

- Última actualización de la documentación: 22 de junio de 2023

En la siguiente tabla se describen los cambios importantes en cada versión de Amazon Kendra. Para obtener notificaciones sobre las actualizaciones de esta documentación, puede suscribirse a la [fuente RSS](#).

Cambio	Descripción	Fecha
Nueva característica	Recupere pasajes semánticamente relevantes mediante la Amazon Kendra API Retrieve para sistemas de recuperación y generación aumentada (RAG).	22 de junio de 2023
Nueva característica	Amazon Kendra ahora admite una versión actualizada del conector de fuentes de datos de Amazon Kendra Web Crawler. Para obtener más información, consulte Amazon Kendra Web Crawler v2.0 .	21 de junio de 2023
Expansión regional	Amazon Kendra ya está disponible en Europa (Londres) (eu-west-2).	5 de junio de 2023
Nueva característica	Amazon Kendra ahora es compatible con una versión actualizada del conector de fuente de datos de Alfresco. Para obtener más información, consulte Alfresco .	16 de mayo de 2023
Nueva característica	Amazon Kendra ahora proporciona un conector de fuentes de datos para Adobe Experience Manager. Para obtener más información, consulte Adobe Experience Manager .	11 de mayo de 2023
Nueva característica	Amazon Kendra ahora admite la configuración de los campos/atributos del documento cuando se llama. GetQuerySuggestions Ahora puede basar las sugerencias de consulta en el contenido de los campos del documento. Para obtener	2 de mayo de 2023

	más información, consulte Sugerencias de consultas .	
Nueva característica	Amazon Kendraahora proporciona un conector de fuente de datos para Gmail. Para obtener más información, consulta Gmail .	13 de abril de 2023
Nueva característica	Amazon Kendraahora admite una versión actualizada del conector de fuentes de OneDrive datos de Microsoft. Para obtener más información, consulte Microsoft OneDrive v2.0 .	3 de abril de 2023
Nueva característica	Mejore la visibilidad de los documentos nuevos o promocione ciertos documentos cuando sus usuarios escriban determinadas consultas mediante los resultados destacados .	30 de marzo de 2023
Nueva característica	Amazon Kendraahora admite un conector de fuente de datos actualizado para MicrosoftSharePoint. Para obtener más información, consulte Microsoft SharePoint .	2 de marzo de 2023
Nueva característica	Amazon Kendraahora admite una versión actualizada del conector de fuente de datos de Confluence. Para obtener más información, consulta Confluence .	1 de marzo de 2023
Expansión regional	Amazon Kendraya está disponible en Asia Pacífico (Tokio) (ap-northeast-1).	7 de febrero de 2023
Nueva característica	Amazon Kendraahora proporciona un conector de fuente de datos para Microsoft Exchange. Para obtener más información, consulte Microsoft Exchange .	12 de enero de 2023
Nueva característica	Amazon Kendraahora proporciona un conector de fuente de datos para Microsoft Yammer. Para obtener más información, consulte Microsoft Yammer .	12 de enero de 2023

<u>Nueva característica</u>	Amazon Kendraahora admite la indexación de tipos de documentos RTF, XML, XSLT, MS_EXCEL, CSV, JSON y MD. Para obtener más información, consulte Tipos de documentos .	11 de enero de 2023
<u>Nueva característica</u>	Amazon Kendraahora admite una versión actualizada del conector de fuente de Amazon S3 datos. Para obtener más información, consulte Amazon S3 .	10 de enero de 2023
<u>Nueva característica</u>	OpenSearch Los resultados de búsqueda (autogestionados) se pueden clasificar semánticamente mediante la clasificación Amazon Kendra inteligente .	9 de enero de 2023
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuentes de datos para Microsoft Teams. Para obtener más información, consulte Microsoft Teams .	5 de enero de 2023
<u>Nueva característica</u>	Amazon Kendraahora tiene un conector de fuente de datos actualizado para Google Drive. Para obtener más información, consulta Google Drive .	5 de enero de 2023
<u>Nueva característica</u>	Amazon Kendraahora tiene un conector de fuente de datos actualizado para ServiceNow. Para obtener más información, consulte ServiceNow .	21 de diciembre de 2022
<u>Nueva característica</u>	Amazon Kendraahora tiene un conector de fuente de datos actualizado para Salesforce. Para obtener más información, consulte Salesforce .	21 de diciembre de 2022
<u>Expansión regional</u>	Amazon Kendraya está disponible en Asia-Pacífico (Bombay) (ap-south-1).	14 de diciembre de 2022
<u>Nueva característica</u>	Amazon KendraLa función de búsqueda tabular puede buscar y extraer respuestas de tablas incrustadas en documentos HTML.	27 de noviembre de 2022
<u>Nueva característica</u>	Amazon Kendraadmite la búsqueda semántica de un conjunto selecto de idiomas .	27 de noviembre de 2022

<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuentes de datos para Dropbox. Para obtener más información, consulte Dropbox .	27 de septiembre de 2022
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para Zendesk. Para obtener más información, consulte Zendesk .	17 de agosto de 2022
<u>Nueva característica</u>	El control de acceso a nivel de documento ahora se puede reconfigurar después de indexar los documentos. Para obtener más información, consulte Configuración del control de acceso .	14 de julio de 2022
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para Alfresco. Para obtener más información, consulte Alfresco .	30 de junio de 2022
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para GitHub. Para obtener más información, consulte GitHub .	2 de junio de 2022
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para Jira. Para obtener más información, consulte Jira .	12 de mayo de 2022
<u>Nueva característica</u>	Las facetas anidadas dentro de una faceta se pueden mostrar en los resultados de la búsqueda. Para obtener más información, consulte Facetas .	5 de mayo de 2022
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para Quip. Para obtener más información, consulte Quip .	19 de abril de 2022
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para Box. Para obtener más información, consulte Box .	6 de abril de 2022

<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para Slack. Para obtener más información, consulte Slack .	14 de marzo de 2022
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para Amazon FSx. Para obtener más información, consulte Amazon FSx .	8 de febrero de 2022
<u>AWSactualizaciones de políticas gestionadas: políticas nuevas (p. 690)</u>	Amazon Kendrase agregaron nuevas políticas AWS administradas. Para obtener más información, consulte Políticas AWS gestionadas para Amazon Kendra .	3 de enero de 2022
<u>Nueva característica</u>	Amazon Kendrala aplicación de búsqueda se puede implementar en unos pocos clics sin necesidad de ningún código de interfaz. Para obtener más información, consulte Implementación de una aplicación de búsqueda sin código .	December 1, 2021
<u>Nueva característica</u>	Los metadatos y el contenido del documento se pueden enriquecer durante el proceso de ingestión del documento. Para obtener más información, consulte Customizing document metadata during the ingestion process (Personalización de los metadatos del documento durante el proceso de ingesta).	December 1, 2021
<u>Nueva característica</u>	Amazon Kendraofrece análisis de búsqueda para obtener información útil sobre su aplicación de búsqueda. Para obtener más información, consulte Obtener información con el análisis de búsqueda .	December 1, 2021
<u>Expansión regional</u>	Amazon Kendraahora está disponible en AWS GovCloud (US-Oeste) (us-gov-west-1).	13 de octubre de 2021

<u>Nueva característica</u>	Amazon Kendraahora puede indexar documentos en varios idiomas y filtrar los resultados de búsqueda por idioma. Consulte Añadir documentos en idiomas distintos del inglés y Buscar en idiomas .	7 de octubre de 2021
<u>Nueva característica</u>	Amazon Kendraahora se integra con el directorio Identity Center para obtener los niveles de acceso de grupos y usuarios para filtrar el contexto de los usuarios . Consulte Configuración de grupos de usuarios para IAM Identity Center .	6 de octubre de 2021
<u>Nuevo tutorial</u>	Amazon Kendraahora ofrece un tutorial que explica cómo crear una solución de búsqueda enriquecida con metadatos. Consulte Creación de una solución de búsqueda inteligente .	13 de agosto de 2021
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para Amazon WorkDocs. Para obtener más información, consulte Amazon WorkDocs .	20 de julio de 2021
<u>Nueva característica</u>	Amazon Kendraahora proporciona un rastreador web para rastrear e indexar páginas web. Para obtener más información, consulte Rastreador web .	17 de junio de 2021
<u>Expansión regional</u>	Amazon Kendraya está disponible en Canadá (Central) (ca-central-1).	16 de junio de 2021
<u>Expansión regional</u>	Amazon Kendraya está disponible en el este de EE. UU. (Ohio) (us-east-2).	7 de junio de 2021
<u>Nueva característica</u>	Amazon Kendraahora admite sugerencias de consultas, en las que se sugieren a los usuarios consultas populares relevantes para su búsqueda. Para obtener más información, consulta Sugerir consultas de búsqueda populares .	27 de mayo de 2021

<u>AWSactualizaciones de políticas gestionadas: políticas nuevas (p. 690)</u>	Amazon Kendrase agregaron nuevas políticas AWS administradas. Para obtener más información, consulte Políticas AWS gestionadas para Amazon Kendra .	27 de mayo de 2021
<u>Expansión regional</u>	Amazon Kendraya está disponible en Asia Pacifico (Singapur) (ap-southeast-1).	5 de mayo de 2021
<u>Nueva característica</u>	Amazon Kendraahora permite ajustar la relevancia de la búsqueda en la consulta anulando las configuraciones de ajuste establecidas a nivel de índice. Para obtener más información, consulte Ajustar la relevancia de la búsqueda y Ajustar las respuestas .	20 de abril de 2021
<u>Nueva característica</u>	Amazon Kendraahora admite la autenticación OAuth 2.0 y el uso de ServiceNow consultas para seleccionar documentos para su indexación. Para obtener más información, consulte ServiceNow .	1 de abril de 2021
<u>Nueva característica</u>	Amazon Kendraahora admite el aprendizaje incremental para los documentos de preguntas frecuentes. Para obtener más información, consulte Enviar comentarios para un aprendizaje incremental .	17 de febrero de 2021
<u>Nueva característica</u>	Amazon Kendraahora admite sinónimos de índice. Para obtener más información, consulte Añadir sinónimos a un índice .	10 de diciembre de 2020
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de base de datos para Google Workspace Drive. Para obtener más información, consulte Uso de un origen de datos de Google Workspace Drive .	8 de diciembre de 2020
<u>Nueva característica</u>	Amazon Kendraahora proporciona una JavaScript biblioteca que facilita el envío de comentarios sobre las consultasAmazon Kendra. Para obtener más información, consulta Enviar comentarios .	8 de diciembre de 2020

<u>Nueva característica</u>	Amazon Kendraahora admite el control de acceso de usuarios basado en tokens. Para obtener más información, consulte <u>Controlar el acceso a los documentos de un índice</u> .	5 de noviembre de 2020
<u>Nueva característica</u>	El conector de fuentes Amazon Kendra de datos de Confluence ahora funciona con Confluence Cloud. Para obtener más información, consulte <u>Using a Confluence data source</u> (Uso de un origen de datos de Confluence).	5 de noviembre de 2020
<u>Expansión regional</u>	Amazon Kendraya está disponible en Asia Pacífico (Sídney) (ap-southeast-2).	2 de noviembre de 2020
<u>Nueva característica</u>	Amazon Kendraahora proporciona un conector de fuente de datos para el servidor Confluence. Para obtener más información, consulte <u>Using a Confluence data source</u> (Uso de un origen de datos de Confluence).	26 de octubre de 2020
<u>Nueva característica</u>	Amazon Kendraahora proporciona una fuente de datos que puede utilizar para generar estadísticas para sus conectores personalizados. Para obtener más información, consulte <u>Uso de una fuente de datos personalizada</u> .	21 de octubre de 2020
<u>Nueva característica</u>	Amazon Kendraahora admite atributos personalizados para las preguntas frecuentes. Para obtener más información, consulte <u>Añadir preguntas y respuestas</u> .	17 de septiembre de 2020
<u>Nueva característica</u>	Amazon Kendraahora devuelve las puntuaciones de confianza de los resultados de la consulta. Para obtener más información, consulte <u>QueryResultItem</u> .	15 de septiembre de 2020
<u>Nueva característica</u>	AWS CloudFormation ya admite Amazon Kendra. Para obtener más información, consulte <u>referencia Amazon Kendra de tipos de recursos - AWS CloudFormation</u> .	10 de septiembre de 2020

<u>Nueva característica</u>	Amazon Kendra añade soporte para AWS PrivateLink. Para obtener más información, consulte Amazon Kendra y puntos de conexión de VPC de tipo interfaz (AWS PrivateLink) .	7 de julio de 2020
<u>Nueva guía</u>	Esta es la primera versión de la Guía para desarrolladores de Amazon Kendra.	11 de mayo de 2020

Referencia de la API

[El Documentación de referencia de la API](#) ahora es una guía independiente.

Glosario de AWS

Para ver la terminología más reciente de AWS, consulte el [Glosario de AWS](#) en la Referencia de Glosario de AWS.