



# **Administración y supervisión GenAI**

NetApp  
October 06, 2025

# Tabla de contenidos

- Administración y supervisión ..... 1
  - Gestione la infraestructura de GenAI ..... 1
    - Permite ver información sobre la infraestructura ..... 1
    - Elimine la infraestructura ..... 1
  - Gestione las bases de conocimientos de GenAI ..... 2
    - Ver información sobre una base de conocimientos ..... 2
    - Editar una base de conocimientos ..... 2
    - Proteger una base de conocimientos con instantáneas ..... 3
    - Agregue orígenes de datos adicionales a una base de conocimientos ..... 5
    - Sincronice sus orígenes de datos con una base de conocimientos ..... 10
    - Evalúe los modelos de chat antes de crear una base de conocimientos ..... 10
    - Anule la publicación de su base de conocimientos ..... 11
    - Suprimir una base de conocimientos ..... 11
  - Administre conectores de Amazon Q Business ..... 12
    - Ver información sobre un conector ..... 12
    - Edite un conector ..... 12
    - Agregue orígenes de datos adicionales a un conector ..... 13
    - Sincronice sus orígenes de datos con un conector ..... 18
    - Eliminar un conector ..... 18
  - Gestionar orígenes de datos de GenAI ..... 19
    - Ver información sobre un origen de datos ..... 19
    - Editar la configuración del origen de datos ..... 19
    - Actualizar el contenido de un origen de datos existente ..... 20
    - Suprimir un origen de datos ..... 20
  - Supervise las operaciones de carga de trabajo con Tracker en NetApp Workload Factory ..... 21
    - Seguimiento y supervisión de operaciones ..... 21
    - Ver solicitud de API ..... 21
    - Vuelva a intentar una operación fallida ..... 22
    - Edite y vuelva a intentar una operación con errores ..... 22

# Administración y supervisión

## Gestione la infraestructura de GenAI

Puede ver los detalles sobre su infraestructura RAG de GenAI implementada o eliminar la infraestructura de chatbot si ya no la necesita.

### Permite ver información sobre la infraestructura

Puede ver información sobre la infraestructura de chatbot.

#### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#).
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. Seleccione el menú **Infraestructura**.
4. Vea información sobre la infraestructura, que incluye detalles sobre los siguientes componentes:
  - Configuración de AWS
  - Configuración de infraestructura
  - El motor de la IA
  - La base de datos de vectores

### Elimine la infraestructura

Si ya no necesita la infraestructura de chatbot que implementó para uno o más chatbots, puede eliminarla de Workload Factory.



Todos los chatbots que se han implementado en esta infraestructura se deshabilitarán y se eliminará todo el historial de chat.

Esta operación solo elimina los vínculos a la infraestructura de IA de Workload Factory; no elimina todos los componentes de AWS. Necesitará eliminar manualmente los siguientes componentes de infraestructura de AWS:

- La instancia de VM
- Puntos finales privados
- El volumen del sistema de archivos FSx para ONTAP que contiene las bases de datos de IA
- El rol de IAM
- La política
- El grupo de seguridad

#### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#).
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. Seleccione el menú **Infraestructura**.

4. ... Seleccione y seleccione **Eliminar infraestructura de chatbot**.
5. Confirme que desea eliminar la infraestructura y seleccione **Eliminar**.

## Resultado

Los componentes de infraestructura del chatbot se eliminan de Workload Factory.

# Gestione las bases de conocimientos de GenAI

Después de crear una base de conocimientos, puede ver los detalles de la base de conocimientos, modificar la base de conocimientos, integrar orígenes de datos adicionales o suprimir la base de conocimientos.

## Ver información sobre una base de conocimientos

Puede ver información sobre la configuración de una base de conocimientos y el origen de datos que están integrados.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los "[experiencias de consola](#)".
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. Seleccione la base de conocimientos que desea ver.

Si se define, los inicios de conversación que se están utilizando se muestran en el panel derecho.

4. Para ver los detalles de la base de conocimientos, ... seleccione y seleccione **Administrar base de conocimientos**.

Esta página muestra el estado Publicado, el estado de incrustación de los orígenes de datos, el modo de incrustación, la lista de todos los orígenes de datos incrustados, etc.

El menú **Acciones** le permite administrar la base de conocimientos si desea realizar cambios.

## Editar una base de conocimientos

Puede actualizar una base de conocimientos cambiando algunas configuraciones o puede agregar o eliminar orígenes de datos.

Cada vez que agregue, modifique o elimine orígenes de datos de la base de conocimientos, debe sincronizar el origen de datos para que se vuelva a indexar en la base de conocimientos. La sincronización es incremental, por lo que Amazon Bedrock solo procesa los objetos del volumen FSx para ONTAP que se han agregado, modificado o eliminado desde la última sincronización.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los "[experiencias de consola](#)".
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimientos, seleccione la base de conocimientos que desea actualizar.
4. ... Selecciona y selecciona **Administrar base de conocimientos**.

Esta página muestra el estado Publicado, el estado de incrustación de los orígenes de datos, el modo de incrustación, la lista de todos los orígenes de datos incrustados, etc.

5. Selecciona el menú **Acciones** y selecciona **Editar base de conocimientos**.
6. En la página Editar base de conocimiento, puede cambiar el nombre de la base de conocimiento, la descripción, el modelo de inserción, el modelo de chat, la habilitación de funciones, elegir si los iniciadores de conversación se crean de manera automática o manual y la política de instantáneas utilizada para el volumen que contiene la base de conocimiento.

Si utiliza el modo Manual para iniciar una conversación, también puede cambiar los temas de inicio de la conversación aquí.



Cada escaneo de la base de conocimientos, que incluye incrustación, costos. Si las barreras de protección de datos están activadas después de crear una base de conocimientos, la base de conocimientos se vuelve a analizar e incurre en costes. Del mismo modo, si cambia los modelos de chat, GenAI volverá a analizar las fuentes de datos asociadas (lo que supondrá un coste).

7. Seleccione **Guardar** después de haber realizado los cambios.

## Proteger una base de conocimientos con instantáneas

Puede proteger los datos de la base de conocimientos realizando y restaurando instantáneas de sus volúmenes de base de conocimientos. Puede restaurar desde una instantánea para revertir a la versión anterior de la base de conocimientos en cualquier momento.

Las copias Snapshot pueden ser más rápidas y eficientes en cuanto al almacenamiento que los backups, y permiten proteger cada base de conocimientos mediante una política de protección diferente. Algunos de los escenarios en los que las instantáneas pueden ser útiles son:

- Pérdida o daños de datos accidentales
- Recuperación de datos incorrectos que se están ingiriendo en la base de conocimientos
- Probar diferentes fuentes de datos o estrategias de fragmentación, y revertir rápidamente una vez finalizada la prueba

## Tomar una copia Snapshot de un volumen de base de conocimientos

Puede guardar el estado de una base de conocimientos mediante una instantánea manual del volumen de la base de conocimientos.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#).
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimientos, seleccione la base de conocimientos que desea proteger.
4. Selecciona y selecciona **Administrar base de conocimientos**.

Esta página muestra el estado Publicado, el estado de incrustación de los orígenes de datos, el modo de incrustación, la lista de todos los orígenes de datos incrustados, etc.

5. Selecciona el menú **Acciones** y selecciona **Instantánea > Crear nueva instantánea**.

6. Opcionalmente, seleccione **Definir nombre de instantánea** e introduzca un nombre personalizado para la instantánea.

Definir un nombre personalizado puede ayudarle a determinar mejor el contenido de una instantánea si necesita restaurarla en el futuro.

7. Seleccione **Crear**.

Se crea una instantánea de la base de conocimientos.

## Restaurar una copia Snapshot de un volumen de base de conocimientos

Cualquier momento, puede restaurar una snapshot manual o programada de un volumen de base de conocimientos.



No puede restaurar una snapshot utilizando la interfaz de usuario de cargas de trabajo de IA generativas si la base de datos almacenada en el volumen está dañada o se ha eliminado. Como solución alternativa, puede restaurar la snapshot mediante "[CLI de ONTAP](#)" en el clúster de ONTAP donde se aloja el volumen.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los "[experiencias de consola](#)".
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimientos, seleccione la base de conocimientos que desea restaurar.
4. ... Seleccione y seleccione **Administrar base de conocimientos**.

Esta página muestra el estado Publicado, el estado de incrustación de los orígenes de datos, el modo de incrustación, la lista de todos los orígenes de datos incrustados, etc.

5. Seleccione el menú **Acciones** y seleccione **Instantánea > Restaurar instantánea**.

Aparece el cuadro de diálogo de selección de instantáneas, donde puede ver una lista de las instantáneas que se han creado para esta base de conocimientos.

6. (Opcional) Deseleccione la opción **Pausa en ejecución y escaneos programados después de restaurar la instantánea** si desea que los escaneos de origen de datos programados y actualmente en ejecución continúen después de restaurar la instantánea.

Esta opción está activada de forma predeterminada para garantizar que no se realice un análisis mientras la base de conocimientos se encuentre en un estado parcialmente restaurado o que un análisis no actualice una base de conocimientos recién restaurada con datos más antiguos.

7. Seleccione la copia Snapshot que desea restaurar de la lista.
8. Seleccione **Restaurar**.

## Clonar una base de conocimientos

Puede crear una nueva base de conocimientos a partir de una instantánea de la base de conocimientos. Esto es útil si la base de conocimientos original está dañada o perdida.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimientos, seleccione la base de conocimientos que desea restaurar.
4. ... Seleccione y seleccione **Administrar base de conocimientos**.

Esta página muestra el estado Publicado, el estado de incrustación de los orígenes de datos, el modo de incrustación, la lista de todos los orígenes de datos incrustados, etc.

5. Seleccione el menú **Acciones** y seleccione **Instantánea > Clonar base de conocimientos**.

Aparece el cuadro de diálogo del clon.

6. Opcionalmente, anule la selección de la opción **Pausa en ejecución y escaneos programados después de clonar la instantánea** si desea que los escaneos de origen de datos programados y actualmente en ejecución continúen después de clonar la instantánea.

Esta opción está activada de forma predeterminada para garantizar que no se realice un análisis mientras la base de conocimientos se encuentre en un estado parcialmente restaurado o que un análisis no actualice una base de conocimientos recién restaurada con datos más antiguos.

7. Seleccione la copia de Snapshot que desea clonar de la lista.
8. Seleccione **continuar**.
9. Introduzca un nombre para la nueva base de conocimientos.
10. Seleccione una SVM del sistema de archivos y un nombre de volumen para la nueva base de conocimientos.
11. Seleccione **Clonar**.

## Agregue orígenes de datos adicionales a una base de conocimientos

Puede incrustar orígenes de datos adicionales en su base de conocimientos para rellenarlos con datos adicionales de la organización.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimiento, seleccione la base de conocimientos donde desea agregar el origen de datos.
4. ... Seleccione y seleccione **Añadir fuente de datos**.
5. Seleccione el tipo de fuente de datos que desea agregar:
  - Agregar el sistema de archivos FSx para ONTAP (usar archivos de un volumen FSx para ONTAP existente)
  - Agregar sistema de archivos (utilizar archivos de un recurso compartido SMB o NFS genérico)

## Agregar un sistema de archivos FSx para ONTAP

1. **Seleccione un sistema de archivos:** Seleccione el sistema de archivos FSX for ONTAP donde residen sus archivos de origen de datos y seleccione **Siguiente**.
2. **Selecciona un volumen:** Selecciona el volumen en el que residen tus archivos de origen de datos y selecciona **Siguiente**.

Al seleccionar los archivos almacenados mediante el protocolo SMB, deberá introducir la información de Active Directory, que incluye el dominio, la dirección IP, el nombre de usuario y la contraseña.

3. **Seleccione una fuente de datos:** Seleccione la ubicación de la fuente de datos en función de dónde haya guardado los archivos. Esto puede ser un volumen completo, o simplemente una carpeta o subcarpeta específica en el volumen, y seleccione **Siguiente**.
4. **Configuraciones:** Configure cómo la fuente de datos ingiere información de sus archivos y qué archivos incluye en los escaneos:
  - **Definir fuente de datos:** En la sección **Estrategia de Chunking**, defina cómo el motor GenAI divide el contenido de la fuente de datos en fragmentos cuando la fuente de datos se integra con una base de conocimientos. Puede elegir una de las siguientes estrategias:
    - **\* Fragmentación de varias frases\*:** Organiza la información de su fuente de datos en fragmentos definidos por frases. Puedes elegir cuántas oraciones componen cada fragmento (hasta 100).
    - **\* Fragmento basado en superposición \*:** Organiza la información de su fuente de datos en fragmentos definidos por caracteres que pueden superponerse a fragmentos vecinos. Puedes elegir el tamaño de cada fragmento en caracteres, y cuánto se superpone cada fragmento con fragmentos adyacentes. Puede configurar un tamaño de fragmento de entre 50 y 3000 caracteres, y un porcentaje de superposición de entre 1 y 99%.



La elección de un alto porcentaje de superposición puede aumentar en gran medida los requisitos de almacenamiento con solo pequeñas mejoras en la precisión de la recuperación.

- **Filtrado de archivos:** Configure qué archivos se incluyen en los escaneos:
  - En la sección **File Types support**, elija incluir todos los tipos de archivos o seleccionar tipos de archivos individuales para incluirlos en los escaneos de fuentes de datos.

Si incluye imágenes o archivos PDF, NetApp Workload Factory for GenAI analiza el texto de las imágenes (incluidas las imágenes en documentos PDF), y esto implica un costo mayor.

Al incluir datos de texto de imágenes, GenAI no puede enmascarar la información de identificación personal (PII) de la imagen a medida que los datos de texto escaneados se envían desde su entorno a AWS. Sin embargo, una vez almacenados los datos, todos los PII se enmascaran en la base de datos de GenAI.



Su elección para incluir archivos de imagen en los escaneos está relacionada con el modelo de chat de la base de conocimientos. Si incluye archivos de imagen en los escaneos, el modelo de chat debe admitir imágenes. Si se seleccionan aquí tipos de archivo de imagen, no puede cambiar la base de conocimientos a un modelo de chat que no admita archivos de imagen.

- En la sección **Filtro de tiempo de modificación de archivo**, elija habilitar o deshabilitar la



inclusión de archivos en función de su tiempo de modificación. Si activa el filtrado de tiempo de modificación, seleccione un intervalo de fechas de la lista.



Si incluye archivos basados en un rango de fechas de modificación, en cuanto el rango de fechas no se cumpla (los archivos no se han modificado dentro del rango de fechas especificado), los archivos se excluirán de la exploración periódica y el origen de datos no incluirá estos archivos.

5. En la sección **Permission Aware**, que solo está disponible cuando el origen de datos que seleccionó está en un volumen que utiliza el protocolo SMB, puede habilitar o deshabilitar las respuestas con permiso:
  - **Activado:** Los usuarios del chatbot que accedan a esta base de conocimientos solo obtendrán respuestas a las consultas de las fuentes de datos a las que tengan acceso.
  - **Deshabilitado:** Los usuarios del chatbot recibirán respuestas usando contenido de todas las fuentes de datos integradas.
6. Seleccione **Agregar** para agregar esta fuente de datos a su base de conocimientos.

### Agregar un sistema de archivos NFS genérico

1. **Seleccionar un sistema de archivos:** Ingrese la dirección IP o FQDN para el host del sistema de archivos donde residen los archivos de origen de datos, elija el protocolo NFS para el recurso compartido de red y seleccione **Siguiente**.
2. **Seleccione una fuente de datos:** Seleccione la ubicación de la fuente de datos en función de dónde haya guardado los archivos. Esto puede ser un volumen completo, o simplemente una carpeta o subcarpeta específica en el volumen, y seleccione **Siguiente**.



En algunos casos, podría ser necesario introducir manualmente el nombre de la exportación NFS y seleccionar **Recuperar directorios** para ver los directorios disponibles. Puede seleccionar la exportación completa o solo carpetas específicas.

3. **Configuraciones:** Configure cómo la fuente de datos ingiere información de sus archivos y qué archivos incluye en los escaneos:
  - **Definir fuente de datos:** En la sección **Estrategia de Chunking**, defina cómo el motor GenAI divide el contenido de la fuente de datos en fragmentos cuando la fuente de datos se integra con una base de conocimientos. Puede elegir una de las siguientes estrategias:
    - \* Fragmentación de varias frases\*: Organiza la información de su fuente de datos en fragmentos definidos por frases. Puedes elegir cuántas oraciones componen cada fragmento (hasta 100).
    - \* Fragmento basado en superposición \*: Organiza la información de su fuente de datos en fragmentos definidos por caracteres que pueden superponerse a fragmentos vecinos. Puedes elegir el tamaño de cada fragmento en caracteres, y cuánto se superpone cada fragmento con fragmentos adyacentes. Puede configurar un tamaño de fragmento de entre 50 y 3000 caracteres, y un porcentaje de superposición de entre 1 y 99%.



La elección de un alto porcentaje de superposición puede aumentar en gran medida los requisitos de almacenamiento con solo pequeñas mejoras en la precisión de la recuperación.

- **Filtrado de archivos:** Configure qué archivos se incluyen en los escaneos:

- En la sección **File Types support**, elija incluir todos los tipos de archivos o seleccionar tipos de archivos individuales para incluirlos en los escaneos de fuentes de datos.

Si incluye imágenes o archivos PDF, NetApp Workload Factory for GenAI analiza el texto de las imágenes (incluidas las imágenes en documentos PDF), y esto implica un costo mayor.

Al incluir datos de texto de imágenes, GenAI no puede enmascarar la información de identificación personal (PII) de la imagen a medida que los datos de texto escaneados se envían desde su entorno a AWS. Sin embargo, una vez almacenados los datos, todos los PII se enmascaran en la base de datos de GenAI.



Su elección para incluir archivos de imagen en los escaneos está relacionada con el modelo de chat de la base de conocimientos. Si incluye archivos de imagen en los escaneos, el modelo de chat debe admitir imágenes. Si se seleccionan aquí tipos de archivo de imagen, no puede cambiar la base de conocimientos a un modelo de chat que no admita archivos de imagen.

- En la sección **Filtro de tiempo de modificación de archivo**, elija habilitar o deshabilitar la inclusión de archivos en función de su tiempo de modificación. Si activa el filtrado de tiempo de modificación, seleccione un intervalo de fechas de la lista.



Si incluye archivos basados en un rango de fechas de modificación, en cuanto el rango de fechas no se cumpla (los archivos no se han modificado dentro del rango de fechas especificado), los archivos se excluirán de la exploración periódica y el origen de datos no incluirá estos archivos.

4. Seleccione **Agregar fuente de datos** para agregar esta fuente de datos a su base de conocimientos.

### Agregar un sistema de archivos SMB genérico

#### 1. Seleccionar sistema de archivos:

- a. Ingrese la dirección IP o FQDN del host del sistema de archivos donde residen los archivos de origen de datos.
- b. Seleccione el protocolo SMB para el recurso compartido de red.
- c. Ingrese la información de Active Directory, que incluye el dominio, la dirección IP, el nombre de usuario y la contraseña.
- d. Seleccione **Siguiente**.

2. **Seleccione una fuente de datos:** Seleccione la ubicación de la fuente de datos en función de dónde haya guardado los archivos. Esto puede ser un volumen completo, o simplemente una carpeta o subcarpeta específica en el volumen, y seleccione **Siguiente**.



En algunos casos, podría ser necesario introducir manualmente el nombre del recurso compartido SMB y seleccionar **Recuperar directorios** para ver los directorios disponibles. Puede seleccionar todo el recurso compartido o solo algunas carpetas.

3. **Configuraciones:** Configure cómo la fuente de datos ingiere información de sus archivos y qué archivos incluye en los escaneos:

- **Definir fuente de datos:** En la sección **Estrategia de Chunking**, defina cómo el motor GenAI divide el contenido de la fuente de datos en fragmentos cuando la fuente de datos se integra con una base de conocimientos. Puede elegir una de las siguientes estrategias:

- \* Fragmentación de varias frases\*: Organiza la información de su fuente de datos en fragmentos definidos por frases. Puedes elegir cuántas oraciones componen cada fragmento (hasta 100).
- \* Fragmento basado en superposición \*: Organiza la información de su fuente de datos en fragmentos definidos por caracteres que pueden superponerse a fragmentos vecinos. Puedes elegir el tamaño de cada fragmento en caracteres, y cuánto se superpone cada fragmento con fragmentos adyacentes. Puede configurar un tamaño de fragmento de entre 50 y 3000 caracteres, y un porcentaje de superposición de entre 1 y 99%.



La elección de un alto porcentaje de superposición puede aumentar en gran medida los requisitos de almacenamiento con solo pequeñas mejoras en la precisión de la recuperación.

- **Consciente de permisos:** habilitar o deshabilitar respuestas que tengan en cuenta los permisos:
  - **Activado:** Los usuarios del chatbot que accedan a esta base de conocimientos solo obtendrán respuestas a las consultas de las fuentes de datos a las que tengan acceso.
  - **Deshabilitado:** Los usuarios del chatbot recibirán respuestas usando contenido de todas las fuentes de datos integradas.
- **Filtrado de archivos:** Configure qué archivos se incluyen en los escaneos:
  - En la sección **File Types support**, elija incluir todos los tipos de archivos o seleccionar tipos de archivos individuales para incluirlos en los escaneos de fuentes de datos.

Si incluye imágenes o archivos PDF, NetApp Workload Factory for GenAI analiza el texto de las imágenes (incluidas las imágenes en documentos PDF), y esto implica un costo mayor.

Al incluir datos de texto de imágenes, GenAI no puede enmascarar la información de identificación personal (PII) de la imagen a medida que los datos de texto escaneados se envían desde su entorno a AWS. Sin embargo, una vez almacenados los datos, todos los PII se enmascaran en la base de datos de GenAI.



Su elección para incluir archivos de imagen en los escaneos está relacionada con el modelo de chat de la base de conocimientos. Si incluye archivos de imagen en los escaneos, el modelo de chat debe admitir imágenes. Si se seleccionan aquí tipos de archivo de imagen, no puede cambiar la base de conocimientos a un modelo de chat que no admita archivos de imagen.

- En la sección **Filtro de tiempo de modificación de archivo**, elija habilitar o deshabilitar la inclusión de archivos en función de su tiempo de modificación. Si activa el filtrado de tiempo de modificación, seleccione un intervalo de fechas de la lista.



Si incluye archivos basados en un rango de fechas de modificación, en cuanto el rango de fechas no se cumpla (los archivos no se han modificado dentro del rango de fechas especificado), los archivos se excluirán de la exploración periódica y el origen de datos no incluirá estos archivos.

4. Seleccione **Agregar fuente de datos** para agregar esta fuente de datos a su base de conocimientos.

## Resultado


La fuente de datos está integrada en su base de conocimientos.

## Sincronice sus orígenes de datos con una base de conocimientos

Las fuentes de datos se sincronizan automáticamente con la base de conocimientos asociada una vez al día para que cualquier cambio en la fuente de datos se refleje en el chatbot. Si realiza cambios en cualquiera de sus orígenes de datos y desea sincronizar los datos de inmediato, puede realizar una sincronización a petición.

La sincronización es incremental, por lo que Amazon Bedrock solo procesa los objetos de los orígenes de datos que se han agregado, modificado o eliminado desde la última sincronización.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#).
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimientos, seleccione la base de conocimientos que desea sincronizar.
4.  Seleccione y seleccione **Administrar base de conocimientos**.
5. Seleccione el menú **Acciones** y seleccione **Escanear ahora**.

Verá un mensaje que indica que se están escaneando las fuentes de datos y un mensaje final cuando se complete el análisis.


### Resultado

La base de conocimientos se sincroniza con las fuentes de datos adjuntas y cualquier chatbot activo comenzará a utilizar la información más reciente de sus fuentes de datos.

## Pausar o reanudar una sincronización programada

Si desea pausar o reanudar la siguiente sincronización (escaneo) de los orígenes de datos, puede hacerlo en cualquier momento. Es posible que deba pausar la siguiente sincronización programada si va a realizar cambios en un origen de datos y no desea que la sincronización se produzca durante la ventana de cambio.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#).
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En el menú Bases de conocimiento y conectores, seleccione la base de conocimiento cuyos análisis desea pausar o reanudar.
4.  Seleccione y seleccione **Administrar base de conocimientos**.
5. Seleccione el menú **Acciones** y seleccione **Escanear > Pausar escaneo programado** o **Escanear > Reanudar escaneo programado**.

Verá un mensaje que indica que el siguiente análisis programado se ha pausado o reanudado.

## Evalúe los modelos de chat antes de crear una base de conocimientos

Puede evaluar los modelos de chat básicos disponibles antes de crear una base de conocimientos para poder ver qué modelo funciona mejor para su implementación. Dado que el soporte de modelos varía según la región de AWS, consulte ["Esta página de documentación de AWS"](#) para verificar qué modelos puede utilizar en las regiones en las que planea implementar su base de conocimientos.



Esta funcionalidad sólo está disponible cuando no se han creado bases de conocimiento, cuando no existen bases de conocimiento en la página de inventario de bases de conocimiento.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de Bases de Conocimientos, verá la opción de seleccionar el modelo de chat en el lado derecho de la página para el Chatbot.
4. Seleccione el modelo de chat de la lista e introduzca un conjunto de preguntas en el área de mensajes para ver cómo responde el chatbot.
5. Pruebe varios modelos para ver qué modelo es el mejor para su implementación.

### Resultado


Utilice ese modelo de chat cuando cree su base de conocimientos.

## Anule la publicación de su base de conocimientos

Una vez que hayas publicado tu base de conocimientos para que pueda integrarse con una aplicación de chatbot, puedes anular la publicación si quieres que la aplicación de chatbot no acceda a la base de conocimientos.

La anulación de la publicación de la base de conocimientos impide que las aplicaciones de chat funcionen. El punto final de API único en el que se puede acceder a la base de conocimientos está desactivado.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimientos, seleccione la base de conocimientos que desea anular la publicación.
4.  Seleccione y seleccione **Administrar base de conocimientos**.

Esta página muestra el estado Publicado, el estado de incrustación de los orígenes de datos, el modo de incrustación y la lista de todos los orígenes de datos incrustados.

5. Seleccione el menú **Acciones** y seleccione **Anular publicación**.

### Resultado


La base de conocimientos está desactivada y ya no es accesible por una aplicación de chatbot.

## Suprimir una base de conocimientos

Si ya no necesita una base de conocimientos, puede eliminarla. Cuando se elimina una base de conocimientos, ésta se elimina de Workload Factory y se elimina el volumen que contiene la base de conocimientos. Cualquier aplicación o chatbot que utilice la base de conocimientos dejará de funcionar. La eliminación de una base de conocimientos no es reversible.

Al suprimir una base de conocimientos, también debe desasociar la base de conocimientos de los agentes a los que está asociada para suprimir por completo todos los recursos asociados a la base de conocimientos.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimientos, seleccione la base de conocimientos que desea suprimir.
4.  Seleccione y seleccione **Administrar base de conocimientos**.
5. Seleccione el menú **Acciones** y seleccione **Eliminar base de conocimientos**.
6. En el cuadro de diálogo Eliminar base de conocimientos, confirme que desea eliminarla y seleccione **Eliminar**.

### Resultado

La base de conocimientos se elimina de Workload Factory y se elimina su volumen asociado.


## Administre conectores de Amazon Q Business

Después de crear un conector para Amazon Q Business, puede ver los detalles del conector, modificar el conector, integrar orígenes de datos adicionales o eliminar el conector.

### Ver información sobre un conector

Puede ver información sobre la configuración de un conector y los orígenes de datos integrados.

#### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. Desde la página de inventario de Bases de conocimiento y conectores, seleccione el conector que desea ver.
4. Para ver los detalles del conector,  seleccione y seleccione **Administrar conector**.

Esta página muestra el estado Publicado, el estado de incrustación de los orígenes de datos, el modo de incrustación, la lista de todos los orígenes de datos incrustados, etc.

El menú **Acciones** le permite administrar el conector si desea realizar cambios.

### Edite un conector

Puede actualizar un conector cambiando algunos ajustes o puede agregar o eliminar orígenes de datos.

Cada vez que agrega, modifica o elimina fuentes de datos del conector, GenAI necesita enviar la información de origen de datos a Amazon Q Business para que se vuelva a indexar. La sincronización es incremental, por lo que Amazon Q Business solo procesa los objetos de tu volumen de FSx para ONTAP que se han agregado, modificado o eliminado desde la última sincronización.

#### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimientos y conectores, seleccione el conector que desea

actualizar.

4. ... Selecciona y selecciona **Administrar conector**.

Esta página muestra el estado Publicado, el estado de incrustación de los orígenes de datos, el modo de incrustación, la lista de todos los orígenes de datos incrustados, etc.

5. Seleccione el menú **Acciones** y seleccione **Editar conector**.
6. En la página Editar conector, puede cambiar el nombre del conector, la descripción, el modelo de incrustación, la habilitación de barandillas de datos y la política de instantáneas utilizada para el volumen que contiene el conector.



Cada escaneo de la fuente de datos, que incluye la incrustación, incurre en un costo. Si habilita las barandillas de datos después de crear un conector, la fuente de datos se escanea de nuevo y conlleva costes.

7. Seleccione **Guardar** después de haber realizado los cambios.

## Agregue orígenes de datos adicionales a un conector

Puede incrustar orígenes de datos adicionales en el conector para rellenarlo con datos adicionales de la organización.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#).
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de Knowledge Bases & Connectors, seleccione el conector en el que desea agregar el origen de datos.
4. ... Selecciona y selecciona **Añadir fuente de datos**.
5. Seleccione el tipo de fuente de datos que desea agregar:
  - Agregar el sistema de archivos FSx para ONTAP (usar archivos de un volumen FSx para ONTAP existente)
  - Agregar sistema de archivos (utilizar archivos de un recurso compartido SMB o NFS genérico)



## Agregar un sistema de archivos FSx para ONTAP

1. **Seleccione un sistema de archivos:** Seleccione el sistema de archivos FSX for ONTAP donde residen sus archivos de origen de datos y seleccione **Siguiente**.
2. **Selecciona un volumen:** Selecciona el volumen en el que residen tus archivos de origen de datos y selecciona **Siguiente**.

Al seleccionar los archivos almacenados mediante el protocolo SMB, deberá introducir la información de Active Directory, que incluye el dominio, la dirección IP, el nombre de usuario y la contraseña.

3. **Seleccione una fuente de datos:** Seleccione la ubicación de la fuente de datos en función de dónde haya guardado los archivos. Esto puede ser un volumen completo, o simplemente una carpeta o subcarpeta específica en el volumen, y seleccione **Siguiente**.
4. **Configuraciones:** Configure cómo la fuente de datos ingiere información de sus archivos y qué archivos incluye en los escaneos:
  - **Definir fuente de datos:** En la sección **Estrategia de Chunking**, defina cómo el motor GenAI divide el contenido de la fuente de datos en fragmentos cuando la fuente de datos se integra con una base de conocimientos. Puede elegir una de las siguientes estrategias:
    - \* Fragmentación de varias frases\*: Organiza la información de su fuente de datos en fragmentos definidos por frases. Puedes elegir cuántas oraciones componen cada fragmento (hasta 100).
    - \* Fragmento basado en superposición \*: Organiza la información de su fuente de datos en fragmentos definidos por caracteres que pueden superponerse a fragmentos vecinos. Puedes elegir el tamaño de cada fragmento en caracteres, y cuánto se superpone cada fragmento con fragmentos adyacentes. Puede configurar un tamaño de fragmento de entre 50 y 3000 caracteres, y un porcentaje de superposición de entre 1 y 99%.



La elección de un alto porcentaje de superposición puede aumentar en gran medida los requisitos de almacenamiento con solo pequeñas mejoras en la precisión de la recuperación.

- **Filtrado de archivos:** Configure qué archivos se incluyen en los escaneos:
  - En la sección **File Types support**, elija incluir todos los tipos de archivos o seleccionar tipos de archivos individuales para incluirlos en los escaneos de fuentes de datos.

Si incluye imágenes o archivos PDF, NetApp Workload Factory for GenAI analiza el texto de las imágenes (incluidas las imágenes en documentos PDF), y esto implica un costo mayor.

Al incluir datos de texto de imágenes, GenAI no puede enmascarar la información de identificación personal (PII) de la imagen a medida que los datos de texto escaneados se envían desde su entorno a AWS. Sin embargo, una vez almacenados los datos, todos los PII se enmascaran en la base de datos de GenAI.



Su elección para incluir archivos de imagen en los escaneos está relacionada con el modelo de chat de la base de conocimientos. Si incluye archivos de imagen en los escaneos, el modelo de chat debe admitir imágenes. Si se seleccionan aquí tipos de archivo de imagen, no puede cambiar la base de conocimientos a un modelo de chat que no admita archivos de imagen.

- En la sección **Filtro de tiempo de modificación de archivo**, elija habilitar o deshabilitar la



inclusión de archivos en función de su tiempo de modificación. Si activa el filtrado de tiempo de modificación, seleccione un intervalo de fechas de la lista.



Si incluye archivos basados en un rango de fechas de modificación, en cuanto el rango de fechas no se cumpla (los archivos no se han modificado dentro del rango de fechas especificado), los archivos se excluirán de la exploración periódica y el origen de datos no incluirá estos archivos.

5. En la sección **Permission Aware**, que solo está disponible cuando el origen de datos que seleccionó está en un volumen que utiliza el protocolo SMB, puede habilitar o deshabilitar las respuestas con permiso:
  - **Activado:** Los usuarios del chatbot que accedan a esta base de conocimientos solo obtendrán respuestas a las consultas de las fuentes de datos a las que tengan acceso.
  - **Deshabilitado:** Los usuarios del chatbot recibirán respuestas usando contenido de todas las fuentes de datos integradas.
6. Seleccione **Agregar** para agregar esta fuente de datos a su base de conocimientos.

### Agregar un sistema de archivos NFS genérico

1. **Seleccionar un sistema de archivos:** Ingrese la dirección IP o FQDN para el host del sistema de archivos donde residen los archivos de origen de datos, elija el protocolo NFS para el recurso compartido de red y seleccione **Siguiente**.
2. **Seleccione una fuente de datos:** Seleccione la ubicación de la fuente de datos en función de dónde haya guardado los archivos. Esto puede ser un volumen completo, o simplemente una carpeta o subcarpeta específica en el volumen, y seleccione **Siguiente**.



En algunos casos, podría ser necesario introducir manualmente el nombre de la exportación NFS y seleccionar **Recuperar directorios** para ver los directorios disponibles. Puede seleccionar la exportación completa o solo carpetas específicas.

3. **Configuraciones:** Configure cómo la fuente de datos ingiere información de sus archivos y qué archivos incluye en los escaneos:
  - **Definir fuente de datos:** En la sección **Estrategia de Chunking**, defina cómo el motor GenAI divide el contenido de la fuente de datos en fragmentos cuando la fuente de datos se integra con una base de conocimientos. Puede elegir una de las siguientes estrategias:
    - \* Fragmentación de varias frases\*: Organiza la información de su fuente de datos en fragmentos definidos por frases. Puedes elegir cuántas oraciones componen cada fragmento (hasta 100).
    - \* Fragmento basado en superposición \*: Organiza la información de su fuente de datos en fragmentos definidos por caracteres que pueden superponerse a fragmentos vecinos. Puedes elegir el tamaño de cada fragmento en caracteres, y cuánto se superpone cada fragmento con fragmentos adyacentes. Puede configurar un tamaño de fragmento de entre 50 y 3000 caracteres, y un porcentaje de superposición de entre 1 y 99%.



La elección de un alto porcentaje de superposición puede aumentar en gran medida los requisitos de almacenamiento con solo pequeñas mejoras en la precisión de la recuperación.

- **Filtrado de archivos:** Configure qué archivos se incluyen en los escaneos:

- En la sección **File Types support**, elija incluir todos los tipos de archivos o seleccionar tipos de archivos individuales para incluirlos en los escaneos de fuentes de datos.

Si incluye imágenes o archivos PDF, NetApp Workload Factory for GenAI analiza el texto de las imágenes (incluidas las imágenes en documentos PDF), y esto implica un costo mayor.

Al incluir datos de texto de imágenes, GenAI no puede enmascarar la información de identificación personal (PII) de la imagen a medida que los datos de texto escaneados se envían desde su entorno a AWS. Sin embargo, una vez almacenados los datos, todos los PII se enmascaran en la base de datos de GenAI.



Su elección para incluir archivos de imagen en los escaneos está relacionada con el modelo de chat de la base de conocimientos. Si incluye archivos de imagen en los escaneos, el modelo de chat debe admitir imágenes. Si se seleccionan aquí tipos de archivo de imagen, no puede cambiar la base de conocimientos a un modelo de chat que no admita archivos de imagen.

- En la sección **Filtro de tiempo de modificación de archivo**, elija habilitar o deshabilitar la inclusión de archivos en función de su tiempo de modificación. Si activa el filtrado de tiempo de modificación, seleccione un intervalo de fechas de la lista.



Si incluye archivos basados en un rango de fechas de modificación, en cuanto el rango de fechas no se cumpla (los archivos no se han modificado dentro del rango de fechas especificado), los archivos se excluirán de la exploración periódica y el origen de datos no incluirá estos archivos.

4. Seleccione **Agregar fuente de datos** para agregar esta fuente de datos a su base de conocimientos.

### Agregar un sistema de archivos SMB genérico

#### 1. Seleccionar sistema de archivos:

- a. Ingrese la dirección IP o FQDN del host del sistema de archivos donde residen los archivos de origen de datos.
- b. Seleccione el protocolo SMB para el recurso compartido de red.
- c. Ingrese la información de Active Directory, que incluye el dominio, la dirección IP, el nombre de usuario y la contraseña.
- d. Seleccione **Siguiente**.

2. **Seleccione una fuente de datos:** Seleccione la ubicación de la fuente de datos en función de dónde haya guardado los archivos. Esto puede ser un volumen completo, o simplemente una carpeta o subcarpeta específica en el volumen, y seleccione **Siguiente**.



En algunos casos, podría ser necesario introducir manualmente el nombre del recurso compartido SMB y seleccionar **Recuperar directorios** para ver los directorios disponibles. Puede seleccionar todo el recurso compartido o solo algunas carpetas.

3. **Configuraciones:** Configure cómo la fuente de datos ingiere información de sus archivos y qué archivos incluye en los escaneos:

- **Definir fuente de datos:** En la sección **Estrategia de Chunking**, defina cómo el motor GenAI divide el contenido de la fuente de datos en fragmentos cuando la fuente de datos se integra con una base de conocimientos. Puede elegir una de las siguientes estrategias:

- \* Fragmentación de varias frases\*: Organiza la información de su fuente de datos en fragmentos definidos por frases. Puedes elegir cuántas oraciones componen cada fragmento (hasta 100).
- \* Fragmento basado en superposición \*: Organiza la información de su fuente de datos en fragmentos definidos por caracteres que pueden superponerse a fragmentos vecinos. Puedes elegir el tamaño de cada fragmento en caracteres, y cuánto se superpone cada fragmento con fragmentos adyacentes. Puede configurar un tamaño de fragmento de entre 50 y 3000 caracteres, y un porcentaje de superposición de entre 1 y 99%.



La elección de un alto porcentaje de superposición puede aumentar en gran medida los requisitos de almacenamiento con solo pequeñas mejoras en la precisión de la recuperación.

- **Consciente de permisos:** habilitar o deshabilitar respuestas que tengan en cuenta los permisos:
  - **Activado:** Los usuarios del chatbot que accedan a esta base de conocimientos solo obtendrán respuestas a las consultas de las fuentes de datos a las que tengan acceso.
  - **Deshabilitado:** Los usuarios del chatbot recibirán respuestas usando contenido de todas las fuentes de datos integradas.
- **Filtrado de archivos:** Configure qué archivos se incluyen en los escaneos:
  - En la sección **File Types support**, elija incluir todos los tipos de archivos o seleccionar tipos de archivos individuales para incluirlos en los escaneos de fuentes de datos.

Si incluye imágenes o archivos PDF, NetApp Workload Factory for GenAI analiza el texto de las imágenes (incluidas las imágenes en documentos PDF), y esto implica un costo mayor.

Al incluir datos de texto de imágenes, GenAI no puede enmascarar la información de identificación personal (PII) de la imagen a medida que los datos de texto escaneados se envían desde su entorno a AWS. Sin embargo, una vez almacenados los datos, todos los PII se enmascaran en la base de datos de GenAI.



Su elección para incluir archivos de imagen en los escaneos está relacionada con el modelo de chat de la base de conocimientos. Si incluye archivos de imagen en los escaneos, el modelo de chat debe admitir imágenes. Si se seleccionan aquí tipos de archivo de imagen, no puede cambiar la base de conocimientos a un modelo de chat que no admita archivos de imagen.

- En la sección **Filtro de tiempo de modificación de archivo**, elija habilitar o deshabilitar la inclusión de archivos en función de su tiempo de modificación. Si activa el filtrado de tiempo de modificación, seleccione un intervalo de fechas de la lista.



Si incluye archivos basados en un rango de fechas de modificación, en cuanto el rango de fechas no se cumpla (los archivos no se han modificado dentro del rango de fechas especificado), los archivos se excluirán de la exploración periódica y el origen de datos no incluirá estos archivos.

4. Seleccione **Agregar fuente de datos** para agregar esta fuente de datos a su base de conocimientos.

## Resultado


El origen de datos está integrado en el conector.

## Sincronice sus orígenes de datos con un conector

Las fuentes de datos se sincronizan automáticamente con el conector asociado una vez al día para que cualquier cambio en la fuente de datos se refleje en Amazon Q Business. Si realiza cambios en cualquiera de sus orígenes de datos y desea sincronizar (escanear) los datos inmediatamente, puede realizar una sincronización a petición.

La sincronización es incremental, por lo que Amazon Q Business solo procesa los objetos de los orígenes de datos que se han agregado, modificado o eliminado desde la última sincronización.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#).
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. Desde el menú Bases de conocimiento y conectores, seleccione el conector que desea sincronizar.
4.  Seleccione y seleccione **Administrar conector**.
5. Seleccione el menú **Acciones** y seleccione **Escanear ahora**.

Verá un mensaje que indica que se están escaneando las fuentes de datos y un mensaje final cuando se complete el análisis.


### Resultado

El conector se sincroniza con las fuentes de datos adjuntas y Amazon Q Business comenzará a utilizar la información más reciente de sus fuentes de datos.

## Pausar o reanudar una sincronización programada

Si desea pausar o reanudar la siguiente sincronización (escaneo) de los orígenes de datos, puede hacerlo en cualquier momento. Es posible que deba pausar la siguiente sincronización programada si va a realizar cambios en un origen de datos y no desea que la sincronización se produzca durante la ventana de cambio.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#).
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de conectores, seleccione el conector para el que desea pausar o reanudar las exploraciones.
4.  Seleccione y seleccione **Administrar conector**.
5. Seleccione el menú **Acciones** y seleccione **Escanear > Pausar escaneo programado o Escanear > Reanudar escaneo programado**.


Verá un mensaje que indica que el siguiente análisis programado se ha pausado o reanudado.

## Eliminar un conector

Si ya no necesita un conector, puede eliminarlo. Cuando se elimina un conector, se elimina de Workload Factory y se elimina el volumen que contiene el conector. La eliminación de un conector no es reversible.

Al suprimir un conector, también debe desasociar el conector de cualquier agente con el que esté asociado para suprimir completamente todos los recursos asociados al conector.

## Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimientos y conectores, seleccione el conector que desea suprimir.
4.  Seleccione y seleccione **Administrar conector**.
5. Seleccione el menú **Acciones** y seleccione **Borrar conector**.
6. En el cuadro de diálogo Eliminar conector, confirme que desea eliminarlo y seleccione **Eliminar**.

## Resultado

Se elimina el conector de Workload Factory y se elimina su volumen asociado.


# Gestionar orígenes de datos de GenAI

Después de crear una base de conocimientos o un conector utilizando orígenes de datos en el sistema de archivos FSx para ONTAP, puede ver los detalles de la fuente de datos, actualizar o cambiar el contenido de la fuente de datos, editar la configuración de la fuente de datos o eliminar la fuente de datos.


## Ver información sobre un origen de datos

Puede ver información sobre el contenido de un origen de datos y puede ver su estado de incrustación con la base de conocimientos o el conector. Dado que las fuentes de datos están asociadas a una base de conocimientos o conector, primero tendrá que elegir la base de conocimientos o el conector para poder ver los detalles de la fuente de datos.

## Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. Seleccione la base de conocimientos o el conector donde reside la fuente de datos y, a continuación,  seleccione \* Administrar base de conocimientos \* o \* Administrar conector \*.

La parte inferior de la página muestra los orígenes de datos asociados.

4. Expanda cada fila seleccionando la  para ver información detallada sobre cada origen de datos, como el sistema de archivos FSx para ONTAP, el volumen y la ruta donde reside el origen de datos.

También muestra la información de incrustación y si ese origen de datos está actualmente incrustado en la base de conocimientos o en el conector.

## Editar la configuración del origen de datos

Puede editar información sobre un origen de datos que haya integrado con una base de conocimientos o conector. La mayor parte de la información se corrige después de agregar un origen de datos, pero puede realizar cambios en alguna de la configuración (como la definición de fragmentación o el reconocimiento de permisos).

## Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de bases de conocimiento, seleccione la base de conocimientos donde reside el origen de datos y, a continuación, **...** seleccione y seleccione **Gestionar base de conocimientos**.

La parte inferior de la página muestra los orígenes de datos que forman parte de esta base de conocimientos.

4. En la fila de la fuente de datos que desea editar, **...** seleccione y seleccione **Editar fuente de datos**.
5. En la página Editar origen de datos, seleccione **▼** para ampliar la fila para la definición de fragmentos.
6. Actualice la configuración y la estrategia de fragmentación, así como el reconocimiento de permisos (para volúmenes SMB) y seleccione **Guardar**.

## Resultado

La configuración del origen de datos se actualiza y el sistema AI sincroniza el origen de datos para que se vuelva a indexar en la base de conocimientos.

## Actualizar el contenido de un origen de datos existente

Puede cambiar el contenido de un origen de datos en cualquier momento para agregar o actualizar los datos de su organización. Si este origen de datos se está utilizando activamente en una base de conocimientos, debe sincronizar el origen de datos para que se vuelva a indexar en la base de conocimientos. La sincronización es incremental, por lo que Amazon Bedrock solo procesa los objetos del volumen FSx para ONTAP que se han agregado, modificado o eliminado desde la última sincronización.

Las fuentes de datos se sincronizan automáticamente con la base de conocimientos una vez al día para que cualquier cambio en la fuente de datos se refleje en el chatbot. Si realiza cambios en un origen de datos y desea sincronizar los datos inmediatamente, puede ["realice una sincronización bajo demanda"](#).

## Suprimir un origen de datos

Si ya no necesita un origen de datos para formar parte de su base de conocimientos, puede eliminarlo.

### Pasos

1. Inicie sesión en Workload Factory utilizando uno de los ["experiencias de consola"](#) .
2. En el icono Cargas de trabajo de IA, seleccione **Implementar y gestionar**.
3. En la página de inventario de la base de conocimientos, seleccione la base de conocimientos donde reside el origen de datos y, a continuación, **...** seleccione y seleccione **Gestionar base de conocimientos**.

La parte inferior de la página muestra los orígenes de datos que forman parte de esta base de conocimientos.

4. En la fila de la fuente de datos que desea eliminar, **...** seleccione y seleccione **Eliminar fuente de datos**.
5. En el cuadro de diálogo Eliminar origen de datos, confirme que desea eliminarlo y seleccione **Confirmar**.

## Resultado

El origen de datos se elimina de la base de conocimientos y el sistema AI elimina la información indexada sobre este origen de datos de la base de conocimientos. Cualquier información de esa fuente de datos ya no estará disponible para los chatbots que estén utilizando la base de conocimientos.

# Supervise las operaciones de carga de trabajo con Tracker en NetApp Workload Factory

Supervise y realice un seguimiento de la ejecución de las operaciones de carga de trabajo y supervise el progreso de las tareas con Tracker en NetApp Workload Factory.

## Acerca de esta tarea

NetApp Workload Factory ofrece Tracker, una función de monitoreo, para que pueda monitorear y rastrear el progreso y el estado de las operaciones de carga de trabajo, revisar detalles de tareas y subtareas de operación y diagnosticar cualquier problema o falla.

Hay varias acciones disponibles en Tracker. Puede filtrar trabajos por intervalo de tiempo (últimas 24 horas, 7 días, 14 días o 30 días), carga de trabajo, estado y usuario; buscar trabajos con la función de búsqueda; y descargar la tabla de trabajos como archivo CSV. Puede actualizar el Tracker en cualquier momento y volver a intentar rápidamente una operación fallida o editar parámetros para una operación fallida y volver a intentar la operación.

Tracker admite dos niveles de supervisión dependiendo de la operación. Cada tarea, como el despliegue del sistema de archivos, muestra la descripción de la tarea, el estado, la hora de inicio, la duración de la tarea, el usuario, la región, el recurso proxy, el identificador de tarea y todas las subtareas relacionadas. Puede ver las respuestas de la API para entender lo que ha pasado durante la operación.

## Niveles de tareas de Tracker con ejemplos

- Nivel 1 (tarea): Realiza un seguimiento de la implementación del sistema de archivos.
- Nivel 2 (subtarea): Realiza un seguimiento de las subtareas relacionadas con el despliegue del sistema de archivos.

## Estado de la operación

El estado de la operación en Tracker es el siguiente *in progress*, *success* y *failed*.

## Frecuencia de operación

La frecuencia de operación se basa en el tipo de trabajo y la programación de trabajo.

## Retención de eventos

Los eventos se conservan en la interfaz de usuario durante 30 días.

## Seguimiento y supervisión de operaciones

Realice un seguimiento y monitoree las operaciones en la consola de Workload Factory con Tracker.

### Pasos

1. Inicie sesión con uno de los ["experiencias de consola"](#) botones .
2. Desde el menú de carga de trabajo, seleccione **Administración** y luego seleccione **Rastreador**.
3. En Tracker, utilice los filtros o la búsqueda para limitar los resultados de trabajo. También puedes descargar un informe de trabajos.

## Ver solicitud de API

Vea la solicitud API en CodeBox para una tarea en Tracker.

## Pasos

1. En Rastreador, seleccione una tarea.
2. Seleccione el menú de acciones y luego seleccione **Ver solicitud de API**.

## Vuelva a intentar una operación fallida

Vuelva a intentar una operación fallida en Tracker. También puede copiar el mensaje de error de una operación con un error.



Puede volver a intentar una operación con fallos hasta 10 veces.

## Pasos

1. En Tracker, seleccione una operación fallida.
2. Seleccione el menú de acciones y luego seleccione **Reintentar**.

## Resultado

Se vuelve a iniciar la operación.

## Edite y vuelva a intentar una operación con errores

Edite los parámetros de la operación fallida y vuelva a intentar la operación fuera de Tracker.

## Pasos

1. En Tracker, seleccione una operación fallida.
2. Seleccione el menú de acciones y luego seleccione **Editar y reintentar**.

Se le redirigirá a la página de operaciones, donde puede editar los parámetros y volver a intentar la operación.

## Resultado

Se vuelve a iniciar la operación. Vaya a Rastreador para ver el estado de la operación.



## Información de copyright

Copyright © 2025 NetApp, Inc. Todos los derechos reservados. Imprimido en EE. UU. No se puede reproducir este documento protegido por copyright ni parte del mismo de ninguna forma ni por ningún medio (gráfico, electrónico o mecánico, incluidas fotocopias, grabaciones o almacenamiento en un sistema de recuperación electrónico) sin la autorización previa y por escrito del propietario del copyright.

El software derivado del material de NetApp con copyright está sujeto a la siguiente licencia y exención de responsabilidad:

ESTE SOFTWARE LO PROPORCIONA NETAPP «TAL CUAL» Y SIN NINGUNA GARANTÍA EXPRESA O IMPLÍCITA, INCLUYENDO, SIN LIMITAR, LAS GARANTÍAS IMPLÍCITAS DE COMERCIALIZACIÓN O IDONEIDAD PARA UN FIN CONCRETO, CUYA RESPONSABILIDAD QUEDA EXIMIDA POR EL PRESENTE DOCUMENTO. EN NINGÚN CASO NETAPP SERÁ RESPONSABLE DE NINGÚN DAÑO DIRECTO, INDIRECTO, ESPECIAL, EJEMPLAR O RESULTANTE (INCLUYENDO, ENTRE OTROS, LA OBTENCIÓN DE BIENES O SERVICIOS SUSTITUTIVOS, PÉRDIDA DE USO, DE DATOS O DE BENEFICIOS, O INTERRUPTIÓN DE LA ACTIVIDAD EMPRESARIAL) CUALQUIERA SEA EL MODO EN EL QUE SE PRODUJERON Y LA TEORÍA DE RESPONSABILIDAD QUE SE APLIQUE, YA SEA EN CONTRATO, RESPONSABILIDAD OBJETIVA O AGRAVIO (INCLUIDA LA NEGLIGENCIA U OTRO TIPO), QUE SURJAN DE ALGÚN MODO DEL USO DE ESTE SOFTWARE, INCLUSO SI HUBIEREN SIDO ADVERTIDOS DE LA POSIBILIDAD DE TALES DAÑOS.

NetApp se reserva el derecho de modificar cualquiera de los productos aquí descritos en cualquier momento y sin aviso previo. NetApp no asume ningún tipo de responsabilidad que surja del uso de los productos aquí descritos, excepto aquello expresamente acordado por escrito por parte de NetApp. El uso o adquisición de este producto no lleva implícita ninguna licencia con derechos de patente, de marcas comerciales o cualquier otro derecho de propiedad intelectual de NetApp.

Es posible que el producto que se describe en este manual esté protegido por una o más patentes de EE. UU., patentes extranjeras o solicitudes pendientes.

LEYENDA DE DERECHOS LIMITADOS: el uso, la copia o la divulgación por parte del gobierno están sujetos a las restricciones establecidas en el subpárrafo (b)(3) de los derechos de datos técnicos y productos no comerciales de DFARS 252.227-7013 (FEB de 2014) y FAR 52.227-19 (DIC de 2007).

Los datos aquí contenidos pertenecen a un producto comercial o servicio comercial (como se define en FAR 2.101) y son propiedad de NetApp, Inc. Todos los datos técnicos y el software informático de NetApp que se proporcionan en este Acuerdo tienen una naturaleza comercial y se han desarrollado exclusivamente con fondos privados. El Gobierno de EE. UU. tiene una licencia limitada, irrevocable, no exclusiva, no transferible, no sublicenciable y de alcance mundial para utilizar los Datos en relación con el contrato del Gobierno de los Estados Unidos bajo el cual se proporcionaron los Datos. Excepto que aquí se disponga lo contrario, los Datos no se pueden utilizar, desvelar, reproducir, modificar, interpretar o mostrar sin la previa aprobación por escrito de NetApp, Inc. Los derechos de licencia del Gobierno de los Estados Unidos de América y su Departamento de Defensa se limitan a los derechos identificados en la cláusula 252.227-7015(b) de la sección DFARS (FEB de 2014).

## Información de la marca comercial

NETAPP, el logotipo de NETAPP y las marcas que constan en <http://www.netapp.com/TM> son marcas comerciales de NetApp, Inc. El resto de nombres de empresa y de producto pueden ser marcas comerciales de sus respectivos propietarios.