



Administration et contrôle

GenAI

NetApp
October 06, 2025

Sommaire

Administration et contrôle	1
Gérer l'infrastructure GenAI	1
Afficher des informations sur l'infrastructure	1
Déposer l'infrastructure	1
Gérer les bases de connaissances GenAI	2
Afficher des informations sur une base de connaissances	2
Modifier une base de connaissances	2
Protégez une base de connaissances avec des instantanés	3
Ajouter des sources de données supplémentaires à une base de connaissances	5
Synchronisez vos sources de données avec une base de connaissances	10
Évaluez les modèles de chat avant de créer une base de connaissances	11
Annulez la publication de votre base de connaissances	11
Supprimer une base de connaissances	12
Gérer les connecteurs Amazon Q Business	12
Afficher des informations sur un connecteur	12
Editer un connecteur	13
Ajoutez des sources de données supplémentaires à un connecteur	13
Synchronisez vos sources de données avec un connecteur	19
Supprimer un connecteur	20
Gérez les sources de données GenAI	20
Afficher des informations sur une source de données	20
Modifier les paramètres de la source de données	21
Mettre à jour le contenu d'une source de données existante	21
Supprimer une source de données	22
Surveillez les opérations de charge de travail avec Tracker dans NetApp Workload Factory	22
Suivez et surveillez les opérations	23
Afficher une demande d'API	23
Réessayez une opération ayant échoué	23
Modifiez et réessayez une opération ayant échoué	23

Administration et contrôle

Gérer l'infrastructure GenAI

Vous pouvez afficher les détails sur votre infrastructure RAG GenAI déployée ou supprimer l'infrastructure de chatbot si vous n'en avez plus besoin.

Afficher des informations sur l'infrastructure

Vous pouvez afficher des informations sur l'infrastructure du chatbot

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des ["expériences de la console"](#) .
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Sélectionnez le menu **Infrastructure**.
4. Afficher des informations sur l'infrastructure, qui incluent des détails sur les composants suivants :
 - Paramètres AWS
 - Paramètres de l'infrastructure
 - Le moteur ai
 - La base de données vectorielle

Déposer l'infrastructure

Si vous n'avez plus besoin de l'infrastructure de chatbot que vous avez déployée pour un ou plusieurs chatbots, vous pouvez la supprimer de Workload Factory.



Tous les chatbots déployés sur cette infrastructure seront désactivés et tous les historiques de discussion seront supprimés.

Cette opération supprime uniquement les liens vers l'infrastructure AI de Workload Factory ; elle ne supprime pas tous les composants d'AWS. Vous devrez supprimer manuellement les composants d'infrastructure suivants d'AWS :

- L'instance de VM
- Terminaux privés
- Volume du système de fichiers FSX pour ONTAP qui contient les bases de données d'IA
- Le rôle IAM
- La politique
- Le groupe de sécurité

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des ["expériences de la console"](#) .
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Sélectionnez le menu **Infrastructure**.

4. **...** Sélectionnez et sélectionnez **Supprimer l'infrastructure chatbot**.
5. Confirmez que vous souhaitez supprimer l'infrastructure et sélectionnez **Supprimer**.

Résultat

Les composants de l'infrastructure du chatbot sont supprimés de Workload Factory.

Gérer les bases de connaissances GenAI

Après avoir créé une base de connaissances, vous pouvez afficher les détails de la base de connaissances, modifier la base de connaissances, intégrer d'autres sources de données ou supprimer la base de connaissances.

Afficher des informations sur une base de connaissances

Vous pouvez afficher des informations sur les paramètres d'une base de connaissances et sur la source de données intégrée.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des ["expériences de la console"](#) .
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Sélectionnez la base de connaissances que vous souhaitez afficher.

Si elle est définie, les amorces de conversation actuellement utilisées s'affichent dans le volet de droite.

4. Pour afficher les détails de la base de connaissances, sélectionnez **...** et sélectionnez **gérer la base de connaissances**.

Cette page affiche l'état publié, l'état d'intégration des sources de données, le mode d'intégration, la liste de toutes les sources de données intégrées, etc.

Le menu **actions** vous permet de gérer la base de connaissances si vous souhaitez apporter des modifications.

Modifier une base de connaissances

Vous pouvez mettre à jour une base de connaissances en modifiant certains paramètres ou ajouter ou supprimer des sources de données.

Chaque fois que vous ajoutez, modifiez ou supprimez des sources de données de la base de connaissances, vous devez synchroniser la source de données afin qu'elle soit réindexée dans la base de connaissances. La synchronisation est incrémentielle. Amazon Bedrock ne traite donc que les objets de votre volume FSX pour ONTAP qui ont été ajoutés, modifiés ou supprimés depuis la dernière synchronisation.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des ["expériences de la console"](#) .
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances, sélectionnez la base de connaissances que vous souhaitez mettre à jour.
4. Sélectionnez **...** et sélectionnez **gérer la base de connaissances**.

Cette page affiche l'état publié, l'état d'intégration des sources de données, le mode d'intégration, la liste de toutes les sources de données intégrées, etc.

5. Sélectionnez le menu **actions** et sélectionnez **Modifier la base de connaissances**.
6. Dans la page Modifier la base de connaissances, vous pouvez modifier le nom de la base de connaissances, la description, le modèle d'intégration, le modèle de discussion, l'activation des fonctionnalités, choisir si les démarreurs de conversation sont créés automatiquement ou manuellement et la stratégie de capture instantanée utilisée pour le volume qui contient la base de connaissances.

Si vous utilisez le mode Manuel pour les amorces de conversation, vous pouvez également modifier les amorces de conversation ici.



Chaque analyse de la base de connaissances, qui inclut l'intégration, coûte. Si des rails de protection des données sont activés après la création d'une base de connaissances, celle-ci est analysée à nouveau et entraîne des coûts. De même, si vous modifiez les modèles de chat, GenAI analyse à nouveau les sources de données associées (ce qui engendre un coût).

7. Sélectionnez **Enregistrer** après avoir effectué vos modifications.

Protégez une base de connaissances avec des instantanés

Vous pouvez protéger les données de votre base de connaissances en prenant et en restaurant des instantanés des volumes de votre base de connaissances. Vous pouvez à tout moment restaurer à partir d'un instantané pour revenir à la version précédente de la base de connaissances.

Les copies Snapshot peuvent être plus rapides et plus efficaces en termes de stockage que les sauvegardes et vous permettent de protéger chaque base de connaissances à l'aide d'une règle de protection différente. Voici quelques-uns des scénarios dans lesquels les snapshots peuvent être utiles :

- Perte ou corruption accidentelle des données
- Récupération à partir de données incorrectes ingérées dans la base de connaissances
- Test de différentes sources de données ou stratégies de suppression, et rétablissement rapide une fois les tests terminés

Prenez un instantané d'un volume de la base de connaissances

Vous pouvez enregistrer l'état d'une base de connaissances en prenant un instantané manuel du volume de la base de connaissances.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances, sélectionnez la base de connaissances que vous souhaitez protéger.
4. Sélectionnez **...** et sélectionnez **gérer la base de connaissances**.

Cette page affiche l'état publié, l'état d'intégration des sources de données, le mode d'intégration, la liste de toutes les sources de données intégrées, etc.

5. Sélectionnez le menu **actions** et sélectionnez **instantané > Créer un nouvel instantané**.

6. Si vous le souhaitez, sélectionnez **définir le nom de l'instantané** et entrez un nom personnalisé pour l'instantané.

La définition d'un nom personnalisé peut vous aider à mieux déterminer le contenu d'un instantané si vous devez le restaurer ultérieurement.

7. Sélectionnez **Créer**.

Un instantané de la base de connaissances est créé.

Restaurer un snapshot d'un volume de la base de connaissances

Vous pouvez à tout moment restaurer un snapshot manuel ou planifié d'un volume de la base de connaissances.



Vous ne pouvez pas restaurer un snapshot à l'aide de l'interface utilisateur des workloads Generative ai si la base de données stockée sur le volume est corrompue ou supprimée. Pour contourner ce problème, vous pouvez restaurer le snapshot à l'aide de "[INTERFACE DE LIGNE DE COMMANDES DE ONTAP](#)" sur le cluster ONTAP où le volume est hébergé.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances, sélectionnez la base de connaissances que vous souhaitez restaurer.
4. Sélectionnez **...** et sélectionnez **gérer la base de connaissances**.

Cette page affiche l'état publié, l'état d'intégration des sources de données, le mode d'intégration, la liste de toutes les sources de données intégrées, etc.

5. Sélectionnez le menu **actions** et sélectionnez **instantané > Restaurer instantané**.

La boîte de dialogue de sélection d'instantané s'affiche, dans laquelle vous pouvez afficher la liste des instantanés créés pour cette base de connaissances.

6. (Facultatif) désélectionnez l'option **suspendre l'exécution et les analyses programmées après la restauration de l'instantané** si vous souhaitez que les analyses de source de données planifiées et en cours d'exécution continuent après la restauration de l'instantané.

Cette option est activée par défaut pour garantir qu'une analyse ne se produit pas lorsque la base de connaissances est en état partiellement restauré ou qu'une analyse ne met pas à jour une base de connaissances récemment restaurée avec des données plus anciennes.

7. Sélectionnez l'instantané à restaurer dans la liste.
8. Sélectionnez **Restaurer**.

Cloner une base de connaissances

Vous pouvez créer une nouvelle base de connaissances à partir d'un instantané de la base de connaissances. Ceci est utile si la base de connaissances d'origine est corrompue ou perdue.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances, sélectionnez la base de connaissances que vous souhaitez restaurer.
4. Sélectionnez **☰** et sélectionnez **gérer la base de connaissances**.

Cette page affiche l'état publié, l'état d'intégration des sources de données, le mode d'intégration, la liste de toutes les sources de données intégrées, etc.

5. Sélectionnez le menu **actions** et sélectionnez **instantané > Cloner la base de connaissances**.

La boîte de dialogue clone s'affiche.

6. Si vous le souhaitez, désélectionnez l'option **interrompre l'exécution et les analyses planifiées après le clonage de l'instantané** si vous souhaitez que les analyses de source de données planifiées et en cours d'exécution continuent après le clonage de l'instantané.

Cette option est activée par défaut pour garantir qu'une analyse ne se produit pas lorsque la base de connaissances est en état partiellement restauré ou qu'une analyse ne met pas à jour une base de connaissances récemment restaurée avec des données plus anciennes.

7. Sélectionnez l'instantané à cloner dans la liste.
8. Sélectionnez **Continuer**.
9. Entrez un nom pour la nouvelle base de connaissances.
10. Choisir une SVM de système de fichiers et un nom de volume pour la nouvelle base de connaissances.
11. Sélectionnez **Clone**.

Ajouter des sources de données supplémentaires à une base de connaissances

Vous pouvez intégrer d'autres sources de données dans votre base de connaissances pour les remplir avec des données d'organisation supplémentaires.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances, sélectionnez la base de connaissances dans laquelle vous souhaitez ajouter la source de données.
4. Sélectionnez **☰** et sélectionnez **Ajouter une source de données**.
5. Sélectionnez le type de source de données que vous souhaitez ajouter :
 - Ajouter un système de fichiers FSx pour ONTAP (utiliser les fichiers d'un volume FSx pour ONTAP existant)
 - Ajouter un système de fichiers (utiliser des fichiers provenant d'un partage SMB ou NFS générique)

Ajouter un système de fichiers FSx pour ONTAP

1. **Sélectionnez un système de fichiers** : sélectionnez le système de fichiers FSX pour ONTAP dans lequel résident vos fichiers source de données et sélectionnez **Suivant**.
2. **Sélectionnez un volume** : sélectionnez le volume sur lequel vos fichiers de source de données résident et sélectionnez **Suivant**.

Lorsque vous sélectionnez des fichiers stockés à l'aide du protocole SMB, vous devez entrer les informations Active Directory, notamment le domaine, l'adresse IP, le nom d'utilisateur et le mot de passe.

3. **Sélectionnez une source de données** : sélectionnez l'emplacement de la source de données en fonction de l'emplacement d'enregistrement des fichiers. Il peut s'agir d'un volume entier, ou simplement d'un dossier ou d'un sous-dossier spécifique dans le volume, et sélectionnez **Suivant**.
4. **Configurations** : configurez la façon dont la source de données ingère les informations de vos fichiers et les fichiers qu'elle inclut dans les analyses :
 - **Définir la source de données** : dans la section **Stratégie de partage**, définissez la façon dont le moteur GenAI divise le contenu de la source de données en blocs lorsque la source de données est intégrée à une base de connaissances. Vous pouvez choisir l'une des stratégies suivantes :
 - **Chunking à plusieurs phrases** : organise les informations de votre source de données en blocs définis par des phrases. Vous pouvez choisir combien de phrases composent chaque morceau (jusqu'à 100).
 - **Chunking basé sur le chevauchement** : organise les informations de votre source de données en blocs définis par des caractères qui peuvent chevaucher des blocs voisins. Vous pouvez choisir la taille de chaque bloc en caractères et la quantité de chaque bloc qui chevauche les blocs adjacents. Vous pouvez configurer une taille de bloc comprise entre 50 et 3000 caractères et un pourcentage de chevauchement compris entre 1 et 99 %.



Le choix d'un pourcentage de chevauchement élevé peut considérablement augmenter les besoins de stockage avec seulement de légères améliorations de la précision de récupération.

- **Filtrage de fichiers** : configurez les fichiers inclus dans les analyses :
 - Dans la section **prise en charge des types de fichiers**, choisissez soit d'inclure tous les types de fichiers, soit de sélectionner des types de fichiers individuels à inclure dans les analyses de sources de données.

Si vous incluez des images ou des fichiers PDF, NetApp Workload Factory for GenAI analyse le texte dans les images (y compris les images dans les documents PDF), ce qui entraîne un coût plus élevé.

Lors de l'inclusion de données texte à partir d'images, GenAI ne peut pas masquer les informations à caractère personnel (PII) de l'image car les données texte numérisées sont envoyées de votre environnement vers AWS. Cependant, une fois les données stockées, toutes les PII sont masquées dans la base de données GenAI.



Votre choix d'inclure des fichiers image dans les analyses est lié au modèle de chat de la base de connaissances. Si vous incluez des fichiers image dans les numérisations, le modèle de chat doit prendre en charge les images. Si des types de fichiers d'image sont sélectionnés ici, vous ne pouvez pas passer de la base de connaissances à un modèle de chat qui ne prend pas en charge les fichiers d'image.

- Dans la section **filtre de temps de modification de fichier**, choisissez d'activer ou de désactiver l'inclusion de fichiers en fonction de leur heure de modification. Si vous activez le filtrage de l'heure de modification, sélectionnez une plage de dates dans la liste.



Si vous incluez des fichiers basés sur une plage de dates de modification, dès que la plage de dates n'est pas satisfaite (les fichiers n'ont pas été modifiés dans la plage de dates spécifiée), les fichiers seront exclus de l'analyse périodique et la source de données n'inclura pas ces fichiers.

5. Dans la section **permission Aware**, disponible uniquement lorsque la source de données que vous avez sélectionnée se trouve sur un volume qui utilise le protocole SMB, vous pouvez activer ou désactiver les réponses sensibles aux autorisations :
 - **Activé** : les utilisateurs du chatbot qui accèdent à cette base de connaissances n'obtiennent que les réponses aux requêtes des sources de données auxquelles ils ont accès.
 - **Désactivé** : les utilisateurs du chatbot recevront des réponses en utilisant le contenu de toutes les sources de données intégrées.
6. Sélectionnez **Ajouter** pour ajouter cette source de données à votre base de connaissances.

Ajouter un système de fichiers NFS générique

1. **Sélectionnez un système de fichiers** : saisissez l'adresse IP ou le nom de domaine complet de l'hôte du système de fichiers où résident vos fichiers source de données, choisissez le protocole NFS pour le partage réseau et sélectionnez **Suivant**.
2. **Sélectionnez une source de données** : sélectionnez l'emplacement de la source de données en fonction de l'emplacement d'enregistrement des fichiers. Il peut s'agir d'un volume entier, ou simplement d'un dossier ou d'un sous-dossier spécifique dans le volume, et sélectionnez **Suivant**.



Dans certains cas, vous devrez peut-être saisir manuellement le nom de l'exportation NFS et sélectionner « Récupérer les répertoires » pour afficher les répertoires disponibles. Vous pouvez sélectionner l'exportation dans son intégralité ou seulement certains dossiers.

3. **Configurations** : configurez la façon dont la source de données ingère les informations de vos fichiers et les fichiers qu'elle inclut dans les analyses :
 - **Définir la source de données** : dans la section **Stratégie de partage**, définissez la façon dont le moteur GenAI divise le contenu de la source de données en blocs lorsque la source de données est intégrée à une base de connaissances. Vous pouvez choisir l'une des stratégies suivantes :
 - **Chunking à plusieurs phrases** : organise les informations de votre source de données en blocs définis par des phrases. Vous pouvez choisir combien de phrases composent chaque morceau (jusqu'à 100).
 - **Chunking basé sur le chevauchement** : organise les informations de votre source de données en blocs définis par des caractères qui peuvent chevaucher des blocs voisins. Vous pouvez choisir la taille de chaque bloc en caractères et la quantité de chaque bloc qui chevauche les blocs adjacents. Vous pouvez configurer une taille de bloc comprise entre 50

et 3000 caractères et un pourcentage de chevauchement compris entre 1 et 99 %.



Le choix d'un pourcentage de chevauchement élevé peut considérablement augmenter les besoins de stockage avec seulement de légères améliorations de la précision de récupération.

- **Filtrage de fichiers** : configurez les fichiers inclus dans les analyses :

- Dans la section **prise en charge des types de fichiers**, choisissez soit d'inclure tous les types de fichiers, soit de sélectionner des types de fichiers individuels à inclure dans les analyses de sources de données.

Si vous incluez des images ou des fichiers PDF, NetApp Workload Factory for GenAI analyse le texte dans les images (y compris les images dans les documents PDF), ce qui entraîne un coût plus élevé.

Lors de l'inclusion de données texte à partir d'images, GenAI ne peut pas masquer les informations à caractère personnel (PII) de l'image car les données texte numérisées sont envoyées de votre environnement vers AWS. Cependant, une fois les données stockées, toutes les PII sont masquées dans la base de données GenAI.



Votre choix d'inclure des fichiers image dans les analyses est lié au modèle de chat de la base de connaissances. Si vous incluez des fichiers image dans les numérisations, le modèle de chat doit prendre en charge les images. Si des types de fichiers d'image sont sélectionnés ici, vous ne pouvez pas passer de la base de connaissances à un modèle de chat qui ne prend pas en charge les fichiers d'image.

- Dans la section **filtre de temps de modification de fichier**, choisissez d'activer ou de désactiver l'inclusion de fichiers en fonction de leur heure de modification. Si vous activez le filtrage de l'heure de modification, sélectionnez une plage de dates dans la liste.



Si vous incluez des fichiers basés sur une plage de dates de modification, dès que la plage de dates n'est pas satisfaite (les fichiers n'ont pas été modifiés dans la plage de dates spécifiée), les fichiers seront exclus de l'analyse périodique et la source de données n'inclura pas ces fichiers.

4. Sélectionnez **Ajouter une source de données** pour ajouter cette source de données à votre base de connaissances.

Ajouter un système de fichiers SMB générique

1. Sélectionnez le système de fichiers:

- a. Saisissez l'adresse IP ou le nom de domaine complet de l'hôte du système de fichiers où résident vos fichiers source de données.
- b. Choisissez le protocole SMB pour le partage réseau.
- c. Saisissez les informations Active Directory, qui incluent le domaine, l'adresse IP, le nom d'utilisateur et le mot de passe.
- d. Sélectionnez **Suivant**.

2. **Sélectionnez une source de données** : sélectionnez l'emplacement de la source de données en fonction de l'emplacement d'enregistrement des fichiers. Il peut s'agir d'un volume entier, ou simplement d'un dossier ou d'un sous-dossier spécifique dans le volume, et sélectionnez **Suivant**.



Dans certains cas, vous devrez peut-être saisir manuellement le nom du partage SMB et sélectionner « Récupérer les répertoires » pour afficher les répertoires disponibles. Vous pouvez sélectionner le partage entier ou seulement certains dossiers.

3. **Configurations** : configurez la façon dont la source de données ingère les informations de vos fichiers et les fichiers qu'elle inclut dans les analyses :

◦ **Définir la source de données** : dans la section **Stratégie de partage**, définissez la façon dont le moteur GenAI divise le contenu de la source de données en blocs lorsque la source de données est intégrée à une base de connaissances. Vous pouvez choisir l'une des stratégies suivantes :

- **Chunking à plusieurs phrases** : organise les informations de votre source de données en blocs définis par des phrases. Vous pouvez choisir combien de phrases composent chaque morceau (jusqu'à 100).
- **Chunking basé sur le chevauchement** : organise les informations de votre source de données en blocs définis par des caractères qui peuvent chevaucher des blocs voisins. Vous pouvez choisir la taille de chaque bloc en caractères et la quantité de chaque bloc qui chevauche les blocs adjacents. Vous pouvez configurer une taille de bloc comprise entre 50 et 3000 caractères et un pourcentage de chevauchement compris entre 1 et 99 %.



Le choix d'un pourcentage de chevauchement élevé peut considérablement augmenter les besoins de stockage avec seulement de légères améliorations de la précision de récupération.

◦ **Conscient des autorisations** : Activer ou désactiver les réponses conscientes des autorisations :

- **Activé** : les utilisateurs du chatbot qui accèdent à cette base de connaissances n'obtiennent que les réponses aux requêtes des sources de données auxquelles ils ont accès.
- **Désactivé** : les utilisateurs du chatbot recevront des réponses en utilisant le contenu de toutes les sources de données intégrées.

◦ **Filtrage de fichiers** : configurez les fichiers inclus dans les analyses :

- Dans la section **prise en charge des types de fichiers**, choisissez soit d'inclure tous les types de fichiers, soit de sélectionner des types de fichiers individuels à inclure dans les analyses de sources de données.

Si vous incluez des images ou des fichiers PDF, NetApp Workload Factory for GenAI analyse le texte dans les images (y compris les images dans les documents PDF), ce qui entraîne un coût plus élevé.

Lors de l'inclusion de données texte à partir d'images, GenAI ne peut pas masquer les informations à caractère personnel (PII) de l'image car les données texte numérisées sont envoyées de votre environnement vers AWS. Cependant, une fois les données stockées, toutes les PII sont masquées dans la base de données GenAI.



Votre choix d'inclure des fichiers image dans les analyses est lié au modèle de chat de la base de connaissances. Si vous incluez des fichiers image dans les numérisations, le modèle de chat doit prendre en charge les images. Si des types de fichiers d'image sont sélectionnés ici, vous ne pouvez pas passer de la base de connaissances à un modèle de chat qui ne prend pas en charge les fichiers d'image.

◦ Dans la section **filtre de temps de modification de fichier**, choisissez d'activer ou de désactiver l'inclusion de fichiers en fonction de leur heure de modification. Si vous activez le filtrage de

l'heure de modification, sélectionnez une plage de dates dans la liste.



Si vous incluez des fichiers basés sur une plage de dates de modification, dès que la plage de dates n'est pas satisfaite (les fichiers n'ont pas été modifiés dans la plage de dates spécifiée), les fichiers seront exclus de l'analyse périodique et la source de données n'inclura pas ces fichiers.

4. Sélectionnez **Ajouter une source de données** pour ajouter cette source de données à votre base de connaissances.

Résultat

La source de données est intégrée à votre base de connaissances.

Synchronisez vos sources de données avec une base de connaissances

Les sources de données sont automatiquement synchronisées avec la base de connaissances associée une fois par jour, de sorte que les modifications éventuelles des sources de données soient répercutées dans le chatbot. Si vous modifiez l'une de vos sources de données et que vous souhaitez synchroniser les données immédiatement, vous pouvez effectuer une synchronisation à la demande.

La synchronisation est incrémentielle. Amazon Bedrock ne traite donc que les objets de vos sources de données qui ont été ajoutés, modifiés ou supprimés depuis la dernière synchronisation.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances, sélectionnez la base de connaissances que vous souhaitez synchroniser.
4. Sélectionnez **...** et sélectionnez **gérer la base de connaissances**.
5. Sélectionnez le menu **actions** et sélectionnez **Rechercher maintenant**.

Vous verrez un message indiquant que vos sources de données sont en cours d'analyse et un message final lorsque l'analyse est terminée.

Résultat

La base de connaissances est synchronisée avec les sources de données jointes et tout chatbot actif commencera à utiliser les informations les plus récentes provenant de vos sources de données.

Mettre en pause ou reprendre une synchronisation planifiée

Si vous souhaitez interrompre ou reprendre la prochaine synchronisation (analyse) des sources de données, vous pouvez le faire à tout moment. Vous devrez peut-être interrompre la prochaine synchronisation programmée si vous allez apporter des modifications à une source de données et que vous ne souhaitez pas que la synchronisation se produise pendant la fenêtre de modification.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans le menu Bases de connaissances et connecteurs, sélectionnez la base de connaissances pour

laquelle vous souhaitez suspendre ou reprendre les analyses.

4. Sélectionnez **...** et sélectionnez **gérer la base de connaissances**.
5. Sélectionnez le menu **actions** et sélectionnez **Scan > Pause scan planifié** ou **Scan > reprendre scan planifié**.

Un message vous indique que la prochaine analyse programmée a été interrompue ou reprise.

Évaluez les modèles de chat avant de créer une base de connaissances

Vous pouvez évaluer les modèles de chat de base disponibles avant de créer une base de connaissances afin de déterminer le modèle le mieux adapté à votre implémentation. Étant donné que la prise en charge des modèles varie selon les régions AWS, reportez-vous à la section "[Cette page de documentation AWS](#)" pour vérifier les modèles que vous pouvez utiliser dans les régions où vous prévoyez de déployer votre base de connaissances.



Cette fonctionnalité n'est disponible que lorsqu'aucune base de connaissances n'a été créée — lorsqu'aucune base de connaissances n'existe dans la page d'inventaire des bases de connaissances.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Sur la page d'inventaire des bases de connaissances, vous verrez l'option de sélectionner le modèle de chat sur le côté droit de la page pour le Chatbot.
4. Sélectionnez le modèle de chat dans la liste et entrez un ensemble de questions dans la zone d'invite pour voir comment le chatbot répond.
5. Essayez plusieurs modèles pour identifier le modèle le mieux adapté à votre implémentation.

Résultat

Utilisez ce modèle de chat lorsque vous créez votre base de connaissances.

Annulez la publication de votre base de connaissances

Après avoir publié votre base de connaissances afin qu'elle puisse être intégrée à une application chatbot, vous pouvez annuler la publication si vous souhaitez désactiver l'application chatbot pour l'accès à la base de connaissances.

L'annulation de la publication de la base de connaissances empêche toute application de chat de fonctionner. Le point de terminaison unique de l'API auquel la base de connaissances était accessible est désactivé.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances, sélectionnez la base de connaissances que vous souhaitez annuler la publication.
4. Sélectionnez **...** et sélectionnez **gérer la base de connaissances**.

Cette page affiche l'état publié, l'état d'intégration des sources de données, le mode d'intégration et la liste de toutes les sources de données intégrées.

5. Sélectionnez le menu **actions** et sélectionnez **Annuler la publication**.

Résultat

La base de connaissances est désactivée et n'est plus accessible par une application chatbot.

Supprimer une base de connaissances

Si vous n'avez plus besoin d'une base de connaissances, vous pouvez la supprimer. Lorsque vous supprimez une base de connaissances, elle est supprimée de Workload Factory et le volume qui contient la base de connaissances est supprimé. Toutes les applications ou chatbots utilisant la base de connaissances cesseront de fonctionner. La suppression d'une base de connaissances n'est pas réversible.

Lorsque vous supprimez une base de connaissances, vous devez également dissocier la base de connaissances de tous les agents auxquels elle est associée afin de supprimer entièrement toutes les ressources associées à la base de connaissances.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances, sélectionnez la base de connaissances que vous souhaitez supprimer.
4. Sélectionnez **...** et sélectionnez **gérer la base de connaissances**.
5. Sélectionnez le menu **actions** et sélectionnez **Supprimer la base de connaissances**.
6. Dans la boîte de dialogue Supprimer la base de connaissances, confirmez que vous souhaitez la supprimer et sélectionnez **Supprimer**.

Résultat

La base de connaissances est supprimée de Workload Factory et son volume associé est supprimé.

Gérer les connecteurs Amazon Q Business

Après avoir créé un connecteur pour Amazon Q Business, vous pouvez afficher les détails du connecteur, modifier le connecteur, intégrer d'autres sources de données ou supprimer le connecteur.

Afficher des informations sur un connecteur

Vous pouvez afficher des informations sur les paramètres d'un connecteur et les sources de données intégrées.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances et des connecteurs, sélectionnez le connecteur que vous souhaitez afficher.
4. Pour afficher les détails du connecteur, sélectionnez **...** et sélectionnez **gérer le connecteur**.

Cette page affiche l'état publié, l'état d'intégration des sources de données, le mode d'intégration, la liste

de toutes les sources de données intégrées, etc.

Le menu **actions** vous permet de gérer le connecteur si vous souhaitez apporter des modifications.

Editer un connecteur

Vous pouvez mettre à jour un connecteur en modifiant certains paramètres ou ajouter ou supprimer des sources de données.

Chaque fois que vous ajoutez, modifiez ou supprimez des sources de données du connecteur, GenAI doit envoyer les informations de la source de données à Amazon Q Business afin qu'elles soient réindexées. La synchronisation est incrémentielle. Amazon Q Business ne traite donc que les objets de votre volume FSX pour ONTAP qui ont été ajoutés, modifiés ou supprimés depuis la dernière synchronisation.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire bases de connaissances et connecteurs, sélectionnez le connecteur à mettre à jour.
4. Sélectionnez **...** et sélectionnez **gérer le connecteur**.

Cette page affiche l'état publié, l'état d'intégration des sources de données, le mode d'intégration, la liste de toutes les sources de données intégrées, etc.

5. Sélectionnez le menu **actions** et sélectionnez **Modifier le connecteur**.
6. Dans la page Modifier le connecteur, vous pouvez modifier le nom du connecteur, la description, le modèle d'intégration, l'activation des rails de données et la stratégie de snapshot utilisée pour le volume qui contient le connecteur.



Chaque analyse de source de données, qui inclut l'intégration, implique un coût. Si vous activez les rails de données après la création d'un connecteur, la source de données est à nouveau analysée et génère des coûts supplémentaires.

7. Sélectionnez **Enregistrer** après avoir effectué des modifications.

Ajoutez des sources de données supplémentaires à un connecteur

Vous pouvez intégrer d'autres sources de données dans votre connecteur pour les remplir avec des données d'organisation supplémentaires.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire bases de connaissances et connecteurs, sélectionnez le connecteur dans lequel vous souhaitez ajouter la source de données.
4. Sélectionnez **...** et sélectionnez **Ajouter une source de données**.
5. Sélectionnez le type de source de données que vous souhaitez ajouter :
 - Ajouter un système de fichiers FSx pour ONTAP (utiliser les fichiers d'un volume FSx pour ONTAP existant)

- Ajouter un système de fichiers (utiliser des fichiers provenant d'un partage SMB ou NFS générique)

Ajouter un système de fichiers FSx pour ONTAP

1. **Sélectionnez un système de fichiers** : sélectionnez le système de fichiers FSX pour ONTAP dans lequel résident vos fichiers source de données et sélectionnez **Suivant**.
2. **Sélectionnez un volume** : sélectionnez le volume sur lequel vos fichiers de source de données résident et sélectionnez **Suivant**.

Lorsque vous sélectionnez des fichiers stockés à l'aide du protocole SMB, vous devez entrer les informations Active Directory, notamment le domaine, l'adresse IP, le nom d'utilisateur et le mot de passe.

3. **Sélectionnez une source de données** : sélectionnez l'emplacement de la source de données en fonction de l'emplacement d'enregistrement des fichiers. Il peut s'agir d'un volume entier, ou simplement d'un dossier ou d'un sous-dossier spécifique dans le volume, et sélectionnez **Suivant**.
4. **Configurations** : configurez la façon dont la source de données ingère les informations de vos fichiers et les fichiers qu'elle inclut dans les analyses :
 - **Définir la source de données** : dans la section **Stratégie de partage**, définissez la façon dont le moteur GenAI divise le contenu de la source de données en blocs lorsque la source de données est intégrée à une base de connaissances. Vous pouvez choisir l'une des stratégies suivantes :
 - **Chunking à plusieurs phrases** : organise les informations de votre source de données en blocs définis par des phrases. Vous pouvez choisir combien de phrases composent chaque morceau (jusqu'à 100).
 - **Chunking basé sur le chevauchement** : organise les informations de votre source de données en blocs définis par des caractères qui peuvent chevaucher des blocs voisins. Vous pouvez choisir la taille de chaque bloc en caractères et la quantité de chaque bloc qui chevauche les blocs adjacents. Vous pouvez configurer une taille de bloc comprise entre 50 et 3000 caractères et un pourcentage de chevauchement compris entre 1 et 99 %.



Le choix d'un pourcentage de chevauchement élevé peut considérablement augmenter les besoins de stockage avec seulement de légères améliorations de la précision de récupération.

- **Filtrage de fichiers** : configurez les fichiers inclus dans les analyses :
 - Dans la section **prise en charge des types de fichiers**, choisissez soit d'inclure tous les types de fichiers, soit de sélectionner des types de fichiers individuels à inclure dans les analyses de sources de données.

Si vous incluez des images ou des fichiers PDF, NetApp Workload Factory for GenAI analyse le texte dans les images (y compris les images dans les documents PDF), ce qui entraîne un coût plus élevé.

Lors de l'inclusion de données texte à partir d'images, GenAI ne peut pas masquer les informations à caractère personnel (PII) de l'image car les données texte numérisées sont envoyées de votre environnement vers AWS. Cependant, une fois les données stockées, toutes les PII sont masquées dans la base de données GenAI.



Votre choix d'inclure des fichiers image dans les analyses est lié au modèle de chat de la base de connaissances. Si vous incluez des fichiers image dans les numérisations, le modèle de chat doit prendre en charge les images. Si des types de fichiers d'image sont sélectionnés ici, vous ne pouvez pas passer de la base de connaissances à un modèle de chat qui ne prend pas en charge les fichiers d'image.

- Dans la section **filtre de temps de modification de fichier**, choisissez d'activer ou de désactiver l'inclusion de fichiers en fonction de leur heure de modification. Si vous activez le filtrage de l'heure de modification, sélectionnez une plage de dates dans la liste.



Si vous incluez des fichiers basés sur une plage de dates de modification, dès que la plage de dates n'est pas satisfaite (les fichiers n'ont pas été modifiés dans la plage de dates spécifiée), les fichiers seront exclus de l'analyse périodique et la source de données n'inclura pas ces fichiers.

5. Dans la section **permission Aware**, disponible uniquement lorsque la source de données que vous avez sélectionnée se trouve sur un volume qui utilise le protocole SMB, vous pouvez activer ou désactiver les réponses sensibles aux autorisations :
 - **Activé** : les utilisateurs du chatbot qui accèdent à cette base de connaissances n'obtiennent que les réponses aux requêtes des sources de données auxquelles ils ont accès.
 - **Désactivé** : les utilisateurs du chatbot recevront des réponses en utilisant le contenu de toutes les sources de données intégrées.
6. Sélectionnez **Ajouter** pour ajouter cette source de données à votre base de connaissances.

Ajouter un système de fichiers NFS générique

1. **Sélectionnez un système de fichiers** : saisissez l'adresse IP ou le nom de domaine complet de l'hôte du système de fichiers où résident vos fichiers source de données, choisissez le protocole NFS pour le partage réseau et sélectionnez **Suivant**.
2. **Sélectionnez une source de données** : sélectionnez l'emplacement de la source de données en fonction de l'emplacement d'enregistrement des fichiers. Il peut s'agir d'un volume entier, ou simplement d'un dossier ou d'un sous-dossier spécifique dans le volume, et sélectionnez **Suivant**.



Dans certains cas, vous devrez peut-être saisir manuellement le nom de l'exportation NFS et sélectionner « Récupérer les répertoires » pour afficher les répertoires disponibles. Vous pouvez sélectionner l'exportation dans son intégralité ou seulement certains dossiers.

3. **Configurations** : configurez la façon dont la source de données ingère les informations de vos fichiers et les fichiers qu'elle inclut dans les analyses :
 - **Définir la source de données** : dans la section **Stratégie de partage**, définissez la façon dont le moteur GenAI divise le contenu de la source de données en blocs lorsque la source de données est intégrée à une base de connaissances. Vous pouvez choisir l'une des stratégies suivantes :
 - **Chunking à plusieurs phrases** : organise les informations de votre source de données en blocs définis par des phrases. Vous pouvez choisir combien de phrases composent chaque morceau (jusqu'à 100).
 - **Chunking basé sur le chevauchement** : organise les informations de votre source de données en blocs définis par des caractères qui peuvent chevaucher des blocs voisins. Vous pouvez choisir la taille de chaque bloc en caractères et la quantité de chaque bloc qui chevauche les blocs adjacents. Vous pouvez configurer une taille de bloc comprise entre 50

et 3000 caractères et un pourcentage de chevauchement compris entre 1 et 99 %.



Le choix d'un pourcentage de chevauchement élevé peut considérablement augmenter les besoins de stockage avec seulement de légères améliorations de la précision de récupération.

- **Filtrage de fichiers** : configurez les fichiers inclus dans les analyses :

- Dans la section **prise en charge des types de fichiers**, choisissez soit d'inclure tous les types de fichiers, soit de sélectionner des types de fichiers individuels à inclure dans les analyses de sources de données.

Si vous incluez des images ou des fichiers PDF, NetApp Workload Factory for GenAI analyse le texte dans les images (y compris les images dans les documents PDF), ce qui entraîne un coût plus élevé.

Lors de l'inclusion de données texte à partir d'images, GenAI ne peut pas masquer les informations à caractère personnel (PII) de l'image car les données texte numérisées sont envoyées de votre environnement vers AWS. Cependant, une fois les données stockées, toutes les PII sont masquées dans la base de données GenAI.



Votre choix d'inclure des fichiers image dans les analyses est lié au modèle de chat de la base de connaissances. Si vous incluez des fichiers image dans les numérisations, le modèle de chat doit prendre en charge les images. Si des types de fichiers d'image sont sélectionnés ici, vous ne pouvez pas passer de la base de connaissances à un modèle de chat qui ne prend pas en charge les fichiers d'image.

- Dans la section **filtre de temps de modification de fichier**, choisissez d'activer ou de désactiver l'inclusion de fichiers en fonction de leur heure de modification. Si vous activez le filtrage de l'heure de modification, sélectionnez une plage de dates dans la liste.



Si vous incluez des fichiers basés sur une plage de dates de modification, dès que la plage de dates n'est pas satisfaite (les fichiers n'ont pas été modifiés dans la plage de dates spécifiée), les fichiers seront exclus de l'analyse périodique et la source de données n'inclura pas ces fichiers.

4. Sélectionnez **Ajouter une source de données** pour ajouter cette source de données à votre base de connaissances.

Ajouter un système de fichiers SMB générique

1. Sélectionnez le système de fichiers:

- a. Saisissez l'adresse IP ou le nom de domaine complet de l'hôte du système de fichiers où résident vos fichiers source de données.
- b. Choisissez le protocole SMB pour le partage réseau.
- c. Saisissez les informations Active Directory, qui incluent le domaine, l'adresse IP, le nom d'utilisateur et le mot de passe.
- d. Sélectionnez **Suivant**.

2. **Sélectionnez une source de données** : sélectionnez l'emplacement de la source de données en fonction de l'emplacement d'enregistrement des fichiers. Il peut s'agir d'un volume entier, ou simplement d'un dossier ou d'un sous-dossier spécifique dans le volume, et sélectionnez **Suivant**.



Dans certains cas, vous devrez peut-être saisir manuellement le nom du partage SMB et sélectionner « Récupérer les répertoires » pour afficher les répertoires disponibles. Vous pouvez sélectionner le partage entier ou seulement certains dossiers.

3. **Configurations** : configurez la façon dont la source de données ingère les informations de vos fichiers et les fichiers qu'elle inclut dans les analyses :

◦ **Définir la source de données** : dans la section **Stratégie de partage**, définissez la façon dont le moteur GenAI divise le contenu de la source de données en blocs lorsque la source de données est intégrée à une base de connaissances. Vous pouvez choisir l'une des stratégies suivantes :

- **Chunking à plusieurs phrases** : organise les informations de votre source de données en blocs définis par des phrases. Vous pouvez choisir combien de phrases composent chaque morceau (jusqu'à 100).
- **Chunking basé sur le chevauchement** : organise les informations de votre source de données en blocs définis par des caractères qui peuvent chevaucher des blocs voisins. Vous pouvez choisir la taille de chaque bloc en caractères et la quantité de chaque bloc qui chevauche les blocs adjacents. Vous pouvez configurer une taille de bloc comprise entre 50 et 3000 caractères et un pourcentage de chevauchement compris entre 1 et 99 %.



Le choix d'un pourcentage de chevauchement élevé peut considérablement augmenter les besoins de stockage avec seulement de légères améliorations de la précision de récupération.

◦ **Conscient des autorisations** : Activer ou désactiver les réponses conscientes des autorisations :

- **Activé** : les utilisateurs du chatbot qui accèdent à cette base de connaissances n'obtiennent que les réponses aux requêtes des sources de données auxquelles ils ont accès.
- **Désactivé** : les utilisateurs du chatbot recevront des réponses en utilisant le contenu de toutes les sources de données intégrées.

◦ **Filtrage de fichiers** : configurez les fichiers inclus dans les analyses :

- Dans la section **prise en charge des types de fichiers**, choisissez soit d'inclure tous les types de fichiers, soit de sélectionner des types de fichiers individuels à inclure dans les analyses de sources de données.

Si vous incluez des images ou des fichiers PDF, NetApp Workload Factory for GenAI analyse le texte dans les images (y compris les images dans les documents PDF), ce qui entraîne un coût plus élevé.

Lors de l'inclusion de données texte à partir d'images, GenAI ne peut pas masquer les informations à caractère personnel (PII) de l'image car les données texte numérisées sont envoyées de votre environnement vers AWS. Cependant, une fois les données stockées, toutes les PII sont masquées dans la base de données GenAI.



Votre choix d'inclure des fichiers image dans les analyses est lié au modèle de chat de la base de connaissances. Si vous incluez des fichiers image dans les numérisations, le modèle de chat doit prendre en charge les images. Si des types de fichiers d'image sont sélectionnés ici, vous ne pouvez pas passer de la base de connaissances à un modèle de chat qui ne prend pas en charge les fichiers d'image.

◦ Dans la section **filtre de temps de modification de fichier**, choisissez d'activer ou de désactiver l'inclusion de fichiers en fonction de leur heure de modification. Si vous activez le filtrage de

l'heure de modification, sélectionnez une plage de dates dans la liste.



Si vous incluez des fichiers basés sur une plage de dates de modification, dès que la plage de dates n'est pas satisfaite (les fichiers n'ont pas été modifiés dans la plage de dates spécifiée), les fichiers seront exclus de l'analyse périodique et la source de données n'inclura pas ces fichiers.

4. Sélectionnez **Ajouter une source de données** pour ajouter cette source de données à votre base de connaissances.

Résultat

La source de données est intégrée à votre connecteur.

Synchronisez vos sources de données avec un connecteur

Les sources de données sont automatiquement synchronisées avec le connecteur associé une fois par jour, de sorte que les modifications apportées à la source de données soient répercutées dans Amazon Q Business. Si vous apportez des modifications à l'une de vos sources de données et que vous souhaitez synchroniser (analyser) les données immédiatement, vous pouvez effectuer une synchronisation à la demande.

La synchronisation est incrémentielle. Amazon Q Business ne traite donc que les objets de vos sources de données qui ont été ajoutés, modifiés ou supprimés depuis la dernière synchronisation.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans le menu Bases de connaissances et connecteurs, sélectionnez le connecteur que vous souhaitez synchroniser.
4. Sélectionnez **...** et sélectionnez **gérer le connecteur**.
5. Sélectionnez le menu **actions** et sélectionnez **Rechercher maintenant**.

Vous verrez un message indiquant que vos sources de données sont en cours d'analyse et un message final lorsque l'analyse est terminée.

Résultat

Le connecteur est synchronisé avec les sources de données jointes et Amazon Q Business commencera à utiliser les informations les plus récentes de vos sources de données.

Mettre en pause ou reprendre une synchronisation planifiée

Si vous souhaitez interrompre ou reprendre la prochaine synchronisation (analyse) des sources de données, vous pouvez le faire à tout moment. Vous devrez peut-être interrompre la prochaine synchronisation programmée si vous allez apporter des modifications à une source de données et que vous ne souhaitez pas que la synchronisation se produise pendant la fenêtre de modification.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des "[expériences de la console](#)".
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.

3. Sur la page d'inventaire des connecteurs, sélectionnez le connecteur pour lequel vous souhaitez interrompre ou reprendre les numérisations.
4. Sélectionnez **...** et sélectionnez **gérer le connecteur**.
5. Sélectionnez le menu **actions** et sélectionnez **Scan > Pause scan planifié** ou **Scan > reprendre scan planifié**.

Un message vous indique que la prochaine analyse programmée a été interrompue ou reprise.

Supprimer un connecteur

Si vous n'avez plus besoin d'un connecteur, vous pouvez le supprimer. Lorsque vous supprimez un connecteur, il est supprimé de Workload Factory et le volume qui contient le connecteur est supprimé. La suppression d'un connecteur n'est pas réversible.

Lorsque vous supprimez un connecteur, vous devez également dissocier le connecteur de tous les agents auxquels il est associé pour supprimer entièrement toutes les ressources associées au connecteur.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des ["expériences de la console"](#) .
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire bases de connaissances et connecteurs, sélectionnez le connecteur à supprimer.
4. Sélectionnez **...** et sélectionnez **gérer le connecteur**.
5. Sélectionnez le menu **actions** et sélectionnez **Supprimer le connecteur**.
6. Dans la boîte de dialogue Supprimer le connecteur, confirmez que vous souhaitez le supprimer et sélectionnez **Supprimer**.

Résultat

Le connecteur est supprimé de Workload Factory et son volume associé est supprimé.

Gérez les sources de données GenAI

Après avoir créé une base de connaissances ou un connecteur à l'aide de sources de données sur votre système de fichiers FSX pour ONTAP, vous pouvez afficher les détails de la source de données, mettre à jour ou modifier le contenu de la source de données, modifier les paramètres de la source de données ou supprimer la source de données.

Afficher des informations sur une source de données

Vous pouvez afficher des informations sur le contenu d'une source de données et afficher son état d'intégration avec la base de connaissances ou le connecteur. Étant donné que les sources de données sont associées à une base de connaissances ou à un connecteur, vous devez d'abord choisir la base de connaissances ou le connecteur avant de pouvoir afficher les détails de la source de données.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des ["expériences de la console"](#) .
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.

3. Sélectionnez la base de connaissances ou le connecteur où réside la source de données, puis **...** sélectionnez **gérer la base de connaissances** ou **gérer le connecteur**.

La partie inférieure de la page répertorie les sources de données associées.

4. Développez chaque ligne en sélectionnant **▼** pour afficher des informations détaillées sur chaque source de données, telles que le système de fichiers FSX pour ONTAP, le volume et le chemin d'accès à la source de données.

Il répertorie également les informations d'intégration et indique si cette source de données est actuellement intégrée dans la base de connaissances ou le connecteur.

Modifier les paramètres de la source de données

Vous pouvez modifier les informations relatives à une source de données que vous avez intégrée à une base de connaissances ou à un connecteur. La plupart des informations sont corrigées après l'ajout d'une source de données, mais vous pouvez apporter des modifications à certaines configurations (telles que la définition de la suppression ou la reconnaissance des autorisations).

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des ["expériences de la console"](#).
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire des bases de connaissances, sélectionnez la base de connaissances où réside la source de données, puis **...** sélectionnez et sélectionnez **gérer la base de connaissances**.

La partie inférieure de la page répertorie les sources de données qui font partie de cette base de connaissances.

4. Dans la ligne de la source de données que vous souhaitez modifier, sélectionnez **...** et sélectionnez **Modifier la source de données**.
5. Dans la page Modifier la source de données, sélectionnez **▼** pour développer la ligne pour la définition de bloc.
6. Mettez à jour les paramètres de la stratégie et de la configuration de chunking, ainsi que la prise en compte des autorisations (pour les volumes SMB) et sélectionnez **Enregistrer**.

Résultat

Les paramètres de la source de données sont mis à jour et le système ai synchronise la source de données afin qu'elle soit réindexée à la base de connaissances.

Mettre à jour le contenu d'une source de données existante

Vous pouvez modifier le contenu d'une source de données à tout moment pour ajouter ou mettre à jour vos données organisationnelles. Si cette source de données est activement utilisée dans une base de connaissances, vous devez synchroniser la source de données afin qu'elle soit réindexée dans la base de connaissances. La synchronisation est incrémentielle. Amazon Bedrock ne traite donc que les objets de votre volume FSX pour ONTAP qui ont été ajoutés, modifiés ou supprimés depuis la dernière synchronisation.

Les sources de données sont synchronisées automatiquement une fois par jour avec la base de connaissances. Les modifications apportées aux sources de données sont donc répercutées dans le chatbot. Si vous apportez des modifications à une source de données et que vous souhaitez synchroniser les données immédiatement, vous pouvez ["effectuez une synchronisation à la demande"](#).

Supprimer une source de données

Si vous n'avez plus besoin d'une source de données pour faire partie de votre base de connaissances, vous pouvez la supprimer.

Étapes

1. Connectez-vous à Workload Factory en utilisant l'un des ["expériences de la console"](#) .
2. Dans la mosaïque workloads d'IA, sélectionnez **déployer et gérer**.
3. Dans la page d'inventaire de la base de connaissances, sélectionnez la base de connaissances où réside la source de données, puis **...** sélectionnez et sélectionnez **gérer la base de connaissances**.

La partie inférieure de la page répertorie les sources de données qui font partie de cette base de connaissances.

4. Sur la ligne de la source de données à supprimer, sélectionnez **...** et sélectionnez **Supprimer la source de données**.
5. Dans la boîte de dialogue Supprimer la source de données, confirmez que vous souhaitez la supprimer et sélectionnez **confirmer**.

Résultat

La source de données est supprimée de la base de connaissances et le système ai supprime de la base de connaissances les informations indexées sur cette source de données. Les informations issues de cette source de données ne seront plus disponibles pour les chatbots qui utilisent la base de connaissances.

Surveillez les opérations de charge de travail avec Tracker dans NetApp Workload Factory

Surveillez et suivez l'exécution des opérations de charge de travail et surveillez la progression des tâches avec Tracker dans NetApp Workload Factory.

Description de la tâche

NetApp Workload Factory fournit Tracker, une fonctionnalité de surveillance, qui vous permet de surveiller et de suivre la progression et l'état des opérations de charge de travail, d'examiner les détails des tâches et sous-tâches d'opération et de diagnostiquer les problèmes ou les pannes.

Plusieurs actions sont disponibles dans Tracker. Vous pouvez filtrer les travaux par période (24 heures, 7 jours, 14 jours ou 30 jours), charge de travail, état et utilisateur, rechercher des travaux à l'aide de la fonction de recherche et télécharger la table des travaux sous forme de fichier CSV. Vous pouvez actualiser le Tracker à tout moment et réessayer rapidement une opération ayant échoué ou modifier les paramètres d'une opération ayant échoué et réessayer l'opération.

Tracker prend en charge deux niveaux de surveillance en fonction de l'opération. Chaque tâche, telle que le déploiement du système de fichiers, affiche la description de la tâche, l'état, l'heure de début, la durée de la tâche, l'utilisateur, la région, la ressource proxy, l'ID de la tâche et toutes les sous-tâches associées. Vous pouvez afficher les réponses de l'API pour comprendre ce qui s'est passé pendant l'opération.

Niveaux de tâche de suivi avec des exemples

- Niveau 1 (tâche) : assure le suivi du déploiement du système de fichiers.
- Niveau 2 (sous-tâche) : effectue le suivi des sous-tâches liées au déploiement du système de fichiers.

État de l'opération

L'état d'opération dans le Tracker est le suivant : *in Progress*, *Success* et *failed*.

Fréquence de fonctionnement

La fréquence des opérations est basée sur le type de travail et le planning de travail.

Conservation des événements

Les événements sont conservés dans l'interface utilisateur pendant 30 jours.

Suivez et surveillez les opérations

Suivez et surveillez les opérations dans la console Workload Factory avec Tracker.

Étapes

1. Connectez-vous à l'aide de l'un des "[expériences de la console](#)".
2. Dans le menu de charge de travail, sélectionnez **Administration** puis **Tracker**.
3. Dans Tracker, utilisez les filtres ou effectuez une recherche pour affiner les résultats des emplois. Vous pouvez également télécharger un rapport sur les emplois.

Afficher une demande d'API

Affichez la demande API dans la zone de code pour une tâche dans Tracker.

Étapes

1. Dans Tracker, sélectionnez une tâche.
2. Sélectionnez le menu Actions, puis sélectionnez **Afficher la demande API**.

Réessayez une opération ayant échoué

Réessayez une opération ayant échoué dans le Tracker. Vous pouvez également copier le message d'erreur d'une opération ayant échoué.



Vous pouvez réessayer une opération ayant échoué jusqu'à 10 fois.

Étapes

1. Dans Tracker, sélectionnez une opération ayant échoué.
2. Sélectionnez le menu Actions, puis sélectionnez **Réessayer**.

Résultat

L'opération est lancée à nouveau.

Modifiez et réessayez une opération ayant échoué

Modifiez les paramètres de l'opération ayant échoué et relancez l'opération en dehors du Tracker.

Étapes

1. Dans Tracker, sélectionnez une opération ayant échoué.
2. Sélectionnez le menu Actions, puis sélectionnez **Modifier et réessayer**.

Vous êtes redirigé vers la page d'opération où vous pouvez modifier les paramètres et réessayer l'opération.

Résultat

L'opération est lancée à nouveau. Accédez à Tracker pour afficher l'état de l'opération.

Informations sur le copyright

Copyright © 2025 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUISSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ), QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTEUELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp décline toute responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (b)(3) de la clause Rights in Technical Data-Noncommercial Items du DFARS 252.227-7013 (février 2014) et du FAR 52.227-19 (décembre 2007).

Les données contenues dans les présentes se rapportent à un produit et/ou service commercial (tel que défini par la clause FAR 2.101). Il s'agit de données propriétaires de NetApp, Inc. Toutes les données techniques et tous les logiciels fournis par NetApp en vertu du présent Accord sont à caractère commercial et ont été exclusivement développés à l'aide de fonds privés. Le gouvernement des États-Unis dispose d'une licence limitée irrévocable, non exclusive, non cessible, non transférable et mondiale. Cette licence lui permet d'utiliser uniquement les données relatives au contrat du gouvernement des États-Unis d'après lequel les données lui ont été fournies ou celles qui sont nécessaires à son exécution. Sauf dispositions contraires énoncées dans les présentes, l'utilisation, la divulgation, la reproduction, la modification, l'exécution, l'affichage des données sont interdits sans avoir obtenu le consentement écrit préalable de NetApp, Inc. Les droits de licences du Département de la Défense du gouvernement des États-Unis se limitent aux droits identifiés par la clause 252.227-7015(b) du DFARS (février 2014).

Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques citées sur le site <http://www.netapp.com/TM> sont des marques déposées ou des marques commerciales de NetApp, Inc. Les autres noms de marques et de produits sont des marques commerciales de leurs propriétaires respectifs.