



## **Inizia subito**

### **GenAI**

NetApp  
October 06, 2025

This PDF was generated from <https://docs.netapp.com/it-it/workload-genai/connector/quick-start-connector.html> on October 06, 2025. Always check docs.netapp.com for the latest.

# Sommario

- Inizia subito ..... 1
  - Avvio rapido per i connettori GenAI ..... 1
- Requisiti del connettore GenAI ..... 1
  - Requisiti di base di GenAI ..... 2
  - Requisiti per NetApp Connector per Amazon Q Business ..... 2
- Identificare le origini dati da aggiungere a un connettore ..... 3
  - Numero massimo di origini dati ..... 3
  - Ubicazione delle origini dati ..... 3
  - Protocolli supportati ..... 3
  - Formati di file origine dati supportati ..... 4
- Implementare l'infrastruttura GenAI ..... 4
  - Dettagli dell'infrastruttura ..... 4
  - Implementare l'infrastruttura GenAI ..... 6

# Inizia subito

## Avvio rapido per i connettori GenAI

Inizia a creare un connettore NetApp per Amazon Q Business utilizzando i dati della tua organizzazione presenti sui file system Amazon FSx per NetApp ONTAP. Dopo aver creato un connettore, gli utenti finali possono accedere all'assistente aziendale Amazon Q per ottenere risposte alle loro domande incentrate sull'organizzazione.

1

### Accedi a Workload Factory

Avrai bisogno di ["creare un account con Workload Factory"](#) e accedi utilizzando uno dei ["esperienze di console"](#).

2

### Configura il tuo ambiente per soddisfare i requisiti Genai

Avrai bisogno delle credenziali AWS per implementare l'infrastruttura AWS, un file system FSX per ONTAP implementato e rilevato, l'elenco di origini dati che desideri integrare nel tuo connettore, l'accesso all'applicazione Amazon Q Business e altro ancora.

["Ulteriori informazioni sui requisiti GenAI"](#).

3

### Identificare il file system FSX per ONTAP che contiene le origini dati

Le origini dati che integrerai nel tuo connettore possono trovarsi in un singolo file system FSX per ONTAP o in più file system FSX per ONTAP. Se questi sistemi si trovano in VPC diversi, devono essere accessibili all'interno della stessa rete, oppure è necessario eseguire il peering dei VPC e utilizzare la stessa regione e lo stesso account AWS del motore di ai.

["Scopri come identificare le origini dati"](#).

4

### Implementare l'infrastruttura GenAI

Avvia la procedura guidata di implementazione dell'infrastruttura per implementare l'infrastruttura Genai nel tuo ambiente AWS. Questo processo implementa un'istanza EC2 per il motore NetApp Genai e un volume su un file system FSX for ONTAP in modo da contenere i database del motore ai di NetApp. Il volume viene utilizzato per memorizzare le informazioni sul connettore.

["Scopri come implementare l'infrastruttura GenAI"](#).

### Cosa succederà

Ora puoi creare un connettore per Amazon Q Business per fornire risposte incentrate sull'organizzazione agli utenti finali.

## Requisiti del connettore GenAI

Prima di creare un connettore NetApp per Amazon Q Business, assicurati che Workload

Factory e AWS siano configurati correttamente.

## Requisiti di base di GenAI

Genai ha requisiti generali che il tuo ambiente deve soddisfare prima di iniziare.

### Accesso e account Workload Factory

Avrai bisogno di ["creare un account con Workload Factory"](#) e accedi utilizzando uno dei ["esperienze di console"](#).

### Credenziali e autorizzazioni AWS

È necessario aggiungere le credenziali AWS a Workload Factory con autorizzazioni di lettura/scrittura, il che significa che utilizzerai Workload Factory in modalità *lettura/scrittura* per GenAI.

Al momento, le autorizzazioni in modalità *base e sola lettura* non sono supportate.

Quando imposti le tue credenziali, selezionando i permessi come mostrato di seguito avrai accesso completo alla gestione dei file system FSX per ONTAP e alla distribuzione e gestione dell'istanza Genai EC2 e delle altre risorse AWS necessarie per la tua knowledge base e il tuo chatbot.

["Scopri come aggiungere le credenziali AWS a Workload Factory"](#)

## Requisiti per NetApp Connector per Amazon Q Business

Assicurati che il tuo ambiente soddisfi i seguenti requisiti specifici per Amazon Q Business Connector.

### Applicazione business Amazon Q

È necessario creare un'applicazione Amazon Q Business o utilizzarne una esistente.

- Verifica che l'applicazione esista in una delle tue regioni AWS.
- Assicurarsi di disporre di ["ha creato un indice"](#) per l'applicazione.
- Verificare che l'applicazione non sia in uno stato di errore.

### File system FSX per ONTAP

Ti serve almeno un file system FSX per ONTAP:

- Un file system verrà utilizzato (o creato, se non esiste) dal motore NetApp Genai per memorizzare le informazioni sul connettore.

Questo file system FSX per ONTAP deve usare FlexVol Volumes. I volumi FlexGroup non sono supportati.

- Uno o più file system conterranno le origini dati che verranno aggiunte al connettore.

Un unico file system FSX per ONTAP può essere utilizzato per entrambi questi scopi oppure è possibile utilizzare più file system FSX per ONTAP.

- Dovrai conoscere la regione, il VPC e la subnet di AWS in cui si trova il file system AWS FSX per ONTAP.
- Dovrai prendere in considerazione le coppie tag key/valore che desideri applicare alle risorse AWS che fanno parte di questa implementazione (opzionale).
- È necessario conoscere le informazioni sulla coppia di chiavi che consentono di connettersi in modo

sicuro all'istanza del motore NetApp ai.

["Scopri come implementare e gestire i file system FSX per ONTAP"](#)

## Identificare le origini dati da aggiungere a un connettore

Identificare o creare i documenti (origini dati) che risiedono nel file system FSX per ONTAP da integrare nel connettore. Queste fonti di dati consentono ad Amazon Q Business di fornire risposte accurate e personalizzate alle domande degli utenti in base ai dati rilevanti per l'organizzazione.

### Numero massimo di origini dati

Il numero massimo di origini dati supportate è 10.

### Ubicazione delle origini dati

Le origini dati possono essere memorizzate in un singolo volume o in una cartella all'interno di un volume, in una condivisione SMB o in un file system NFS su Amazon FSX per NetApp ONTAP. Le origini dei dati possono anche essere memorizzate su Amazon FSX per NetApp ONTAP Volumes che sono in una relazione di data Protection di NetApp SnapMirror.

Non è possibile selezionare singoli documenti all'interno di un volume o di una cartella, pertanto è necessario assicurarsi che ogni volume o cartella contenente origini dati non contenga documenti estranei che non devono essere integrati con la knowledge base.

Puoi aggiungere più origini dati in ogni connettore, ma tutte devono risiedere in FSX per ONTAP, accessibili dal tuo account AWS.

La dimensione massima del file per ciascuna origine dati è di 50 MB.

### Protocolli supportati

I connettori supportano i dati da volumi che utilizzano protocolli NFS o SMB/CIFS. Quando si selezionano i file archiviati utilizzando il protocollo SMB, è necessario immettere le informazioni di Active Directory in modo che il connettore possa accedere ai file su tali volumi. Sono inclusi il dominio Active Directory, l'indirizzo IP, il nome utente e la password.

Quando si archivia l'origine dati in una condivisione (file o directory) a cui si accede tramite SMB, i dati sono accessibili solo da utenti o gruppi di chatbot che dispongono delle autorizzazioni per accedere a tale condivisione. Quando questa funzionalità "in base alle autorizzazioni" è attivata, il sistema si confronta l'e-mail dell'utente in auth0 con gli utenti autorizzati a visualizzare o utilizzare i file nella condivisione SMB. Il chatbot fornirà le risposte in base alle autorizzazioni dell'utente per i file incorporati.

Ad esempio, se avete integrato file 10 (origini dati) nel connettore e 2 dei file sono file di risorse umane che contengono informazioni riservate, solo gli utenti di chatbot che sono autenticati per accedere a questi file 2 riceveranno risposte dal chatbot che includono dati da questi file.



Quando si aggiungono origini dati a un Amazon Q Business Connector, ai file di origine dati vengono applicate solo le autorizzazioni utente. Le autorizzazioni del gruppo non vengono applicate.



Se un file nell'origine dati non contiene testo (ad esempio, un'immagine senza testo), Amazon Q Business non lo indicizza ma registra una voce in Amazon CloudWatch Logs rilevando l'assenza di testo.

## Formati di file origine dati supportati

I seguenti formati di file di origine dati sono attualmente supportati da NetApp Connector per Amazon Q Business.

Formato file	Interno
File di valori separati da virgola	.csv
JSON e JSONP	json
Valutazione	md
Microsoft Word	.docx
Testo normale	.txt
Formato documento portatile	in formato .pdf
Microsoft PowerPoint	.ppt o .pptx
Hypertext Markup Language	html
Extensible Markup Language	xml
XSLT	xslt
Microsoft Excel	.xls
Formato RTF	rtf

## Implementare l'infrastruttura GenAI

Prima di creare knowledge base, connettori e applicazioni FSX per ONTAP per la tua organizzazione, devi implementare l'infrastruttura Genai per il framework RAG nel tuo ambiente. I componenti principali dell'infrastruttura sono il servizio Amazon Bedrock, un'istanza di macchina virtuale per il motore NetApp Genai e un file system FSX per ONTAP.

L'infrastruttura distribuita può supportare più knowledge base, chatbot e connettori, quindi in genere è necessario eseguire questa attività una sola volta.

### Dettagli dell'infrastruttura

L'implementazione Genai deve trovarsi in una regione AWS con Amazon Bedrock abilitato. ["Consente di visualizzare l'elenco delle regioni supportate"](#)

L'infrastruttura è costituita dai seguenti componenti.

#### Servizio Amazon Bedrock

Amazon Bedrock è un servizio completamente gestito che ti consente di utilizzare i modelli di base (FMS) di aziende di intelligenza artificiale leader attraverso una singola API. Fornisce inoltre le funzionalità

necessarie per creare applicazioni ai generative sicure.

["Scopri di più su Amazon Bedrock"](#)

### **Amazon Q Business**

Amazon Q si basa su Amazon Bedrock per fornire un assistente ai generativo completamente gestito che puoi utilizzare per rispondere alle domande e generare contenuti sulla base delle informazioni provenienti dalle tue origini dati.

["Scopri di più su Amazon Q Business"](#)

### **Macchina virtuale per il motore NetApp Genai**

Il motore NetApp Genai viene implementato durante questo processo. Fornisce la potenza di elaborazione per acquisire i dati dalle origini dati e poi scriverli nel database vettoriale.

### **File system FSX per ONTAP**

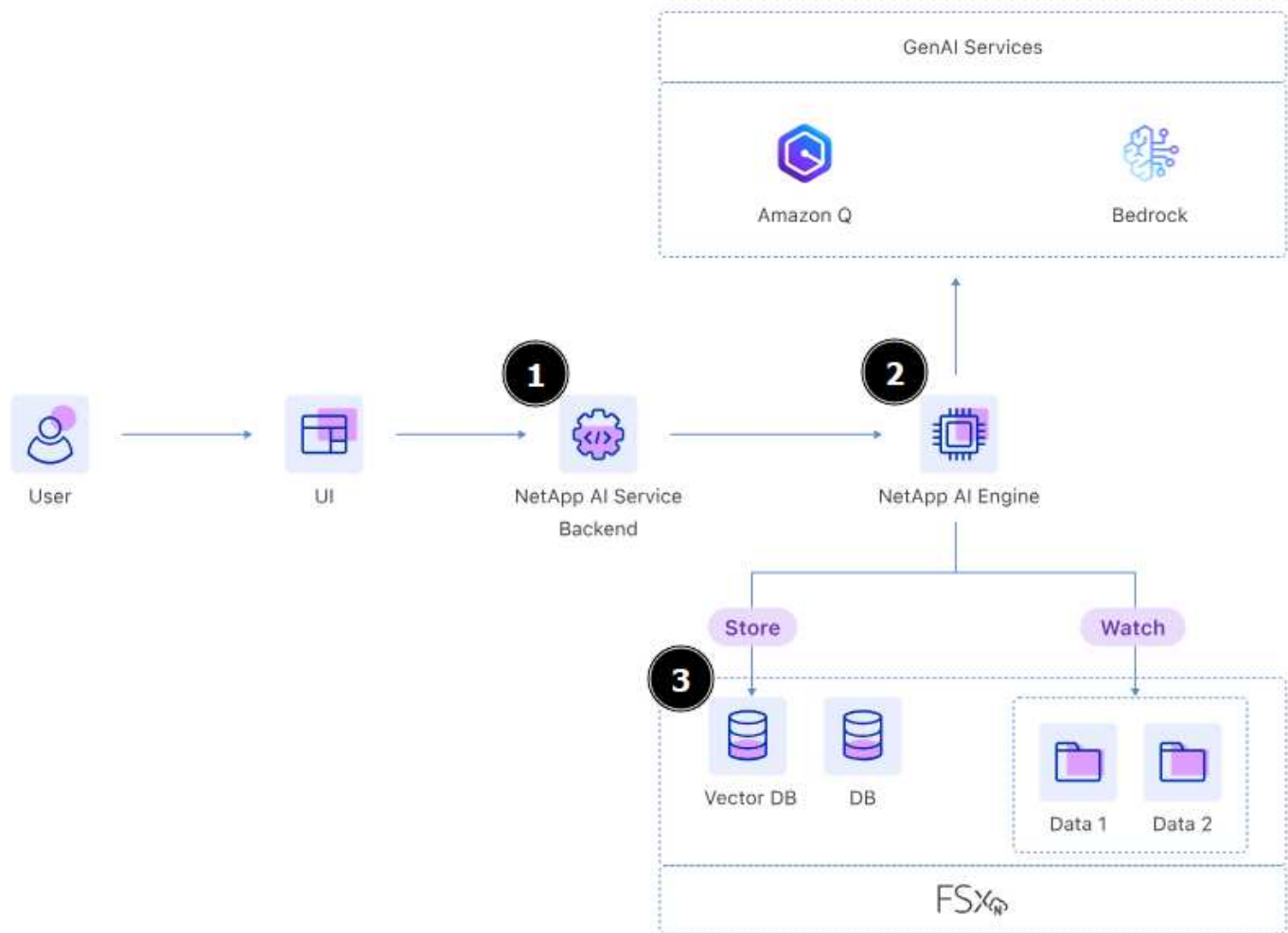
Il file system FSX per ONTAP offre lo storage per il tuo sistema GenAI.

Viene distribuito un singolo volume contenente il database vettoriale che memorizza i dati generati dal modello di base in base alle origini dati.

Le origini dati che integrerai nella tua knowledge base possono risiedere nello stesso file system FSX per ONTAP o su un sistema diverso.

Il motore NetApp Genai monitora e interagisce con entrambi questi volumi.

L'immagine seguente mostra l'infrastruttura GenAI. I componenti numerati 1, 2 e 3 vengono attivati durante questa procedura. Gli altri elementi devono essere presenti prima di iniziare la distribuzione.



## Implementare l'infrastruttura GenAI

Dovrai inserire le tue credenziali AWS e selezionare il file system FSX per ONTAP per implementare l'infrastruttura di recupero e generazione aumentata (RAG).

### Prima di iniziare

Prima di iniziare la procedura, accertarsi che l'ambiente in uso soddisfi i requisiti delle knowledge base o dei connettori, a seconda di quale scelta si scelga.

- ["Requisiti della Knowledge base"](#)
- ["Requisiti del connettore"](#)

### Fasi

1. Accedi a Workload Factory utilizzando uno dei ["esperienze di console"](#).
2. Nella sezione carichi di lavoro ai, selezionare **Deploy & Manage**.
3. Esaminare il diagramma dell'infrastruttura e selezionare **Avanti**.
4. Completare gli elementi nella sezione **Impostazioni AWS**:
  - a. **Credenziali AWS**: Selezionare o aggiungere le credenziali AWS che forniscono le autorizzazioni per implementare le risorse AWS.
  - b. **Posizione**: Selezionare una regione AWS, un VPC e una subnet.



L'implementazione Genai deve trovarsi in una regione AWS con Amazon Bedrock abilitato. ["Consente di visualizzare l'elenco delle regioni supportate"](#)

5. Completare gli elementi nella sezione **Impostazioni infrastruttura**:

- a. **Tag**: inserisci le coppie chiave/valore dei tag che desideri applicare a tutte le risorse AWS che fanno parte di questa distribuzione. Questi tag sono visibili nella AWS Management Console e nell'area delle informazioni sull'infrastruttura all'interno di Workload Factory e possono aiutarti a tenere traccia delle risorse di Workload Factory.

6. Completare la sezione **connettività**:

- a. **Coppia di chiavi**: Selezionare una coppia di chiavi che consente di connettersi in modo sicuro all'istanza del motore NetApp Genai.

7. Completare la sezione **motore ai**:

- a. **Nome istanza**: facoltativamente, seleziona **Definisci nome istanza** e inserisci un nome personalizzato per l'istanza del motore AI. Il nome dell'istanza appare nella AWS Management Console e nell'area delle informazioni sull'infrastruttura all'interno di Workload Factory e può aiutarti a tenere traccia delle risorse di Workload Factory.

8. Selezionare **Deploy** per iniziare la distribuzione.



Se la distribuzione non riesce con un errore di credenziali, è possibile ottenere ulteriori dettagli di errore selezionando i collegamenti ipertestuali all'interno del messaggio di errore. È possibile visualizzare un elenco di autorizzazioni mancanti o bloccate, nonché un elenco di autorizzazioni necessarie al carico di lavoro GenAI per poter implementare l'infrastruttura GenAI.

## Risultato

Workload Factory inizia a distribuire l'infrastruttura del chatbot. Questo processo può richiedere fino a 10 minuti.

Durante il processo di distribuzione, vengono impostati i seguenti elementi:

- La rete viene impostata insieme agli endpoint privati.
- Vengono creati il ruolo IAM, il profilo di istanza e il gruppo di protezione.
- Viene implementata l'istanza della macchina virtuale per il motore GenAI.
- Amazon Bedrock è configurato per inviare i registri ad Amazon CloudWatch Logs, utilizzando un gruppo di registri con il prefisso `/aws/bedrock/`.
- Il motore GenAI è configurato per inviare log ad Amazon CloudWatch Logs, utilizzando un gruppo di log con il nome `/netapp/wlmai/<tenancyAccountId>/randomId`, Dove `<tenancyAccountId>` è il ["ID account console NetApp"](#) per l'utente corrente.

## Informazioni sul copyright

Copyright © 2025 NetApp, Inc. Tutti i diritti riservati. Stampato negli Stati Uniti d'America. Nessuna porzione di questo documento soggetta a copyright può essere riprodotta in qualsiasi formato o mezzo (grafico, elettronico o meccanico, inclusi fotocopie, registrazione, nastri o storage in un sistema elettronico) senza previo consenso scritto da parte del detentore del copyright.

Il software derivato dal materiale sottoposto a copyright di NetApp è soggetto alla seguente licenza e dichiarazione di non responsabilità:

IL PRESENTE SOFTWARE VIENE FORNITO DA NETAPP "COSÌ COM'È" E SENZA QUALSIVOGLIA TIPO DI GARANZIA IMPLICITA O ESPRESSA FRA CUI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, GARANZIE IMPLICITE DI COMMERCIALIZZABILITÀ E IDONEITÀ PER UNO SCOPO SPECIFICO, CHE VENGONO DECLINATE DAL PRESENTE DOCUMENTO. NETAPP NON VERRÀ CONSIDERATA RESPONSABILE IN ALCUN CASO PER QUALSIVOGLIA DANNO DIRETTO, INDIRETTO, ACCIDENTALE, SPECIALE, ESEMPLARE E CONSEGUENZIALE (COMPRESI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, PROCUREMENT O SOSTITUZIONE DI MERCI O SERVIZI, IMPOSSIBILITÀ DI UTILIZZO O PERDITA DI DATI O PROFITTI OPPURE INTERRUZIONE DELL'ATTIVITÀ AZIENDALE) CAUSATO IN QUALSIVOGLIA MODO O IN RELAZIONE A QUALUNQUE TEORIA DI RESPONSABILITÀ, SIA ESSA CONTRATTUALE, RIGOROSA O DOVUTA A INSOLVENZA (COMPRESA LA NEGLIGENZA O ALTRO) INSORTA IN QUALSIASI MODO ATTRAVERSO L'UTILIZZO DEL PRESENTE SOFTWARE ANCHE IN PRESENZA DI UN PREAVVISO CIRCA L'EVENTUALITÀ DI QUESTO TIPO DI DANNI.

NetApp si riserva il diritto di modificare in qualsiasi momento qualunque prodotto descritto nel presente documento senza fornire alcun preavviso. NetApp non si assume alcuna responsabilità circa l'utilizzo dei prodotti o materiali descritti nel presente documento, con l'eccezione di quanto concordato espressamente e per iscritto da NetApp. L'utilizzo o l'acquisto del presente prodotto non comporta il rilascio di una licenza nell'ambito di un qualche diritto di brevetto, marchio commerciale o altro diritto di proprietà intellettuale di NetApp.

Il prodotto descritto in questa guida può essere protetto da uno o più brevetti degli Stati Uniti, esteri o in attesa di approvazione.

LEGENDA PER I DIRITTI SOTTOPOSTI A LIMITAZIONE: l'utilizzo, la duplicazione o la divulgazione da parte degli enti governativi sono soggetti alle limitazioni indicate nel sottoparagrafo (b)(3) della clausola Rights in Technical Data and Computer Software del DFARS 252.227-7013 (FEB 2014) e FAR 52.227-19 (DIC 2007).

I dati contenuti nel presente documento riguardano un articolo commerciale (secondo la definizione data in FAR 2.101) e sono di proprietà di NetApp, Inc. Tutti i dati tecnici e il software NetApp forniti secondo i termini del presente Contratto sono articoli aventi natura commerciale, sviluppati con finanziamenti esclusivamente privati. Il governo statunitense ha una licenza irrevocabile limitata, non esclusiva, non trasferibile, non cedibile, mondiale, per l'utilizzo dei Dati esclusivamente in connessione con e a supporto di un contratto governativo statunitense in base al quale i Dati sono distribuiti. Con la sola esclusione di quanto indicato nel presente documento, i Dati non possono essere utilizzati, divulgati, riprodotti, modificati, visualizzati o mostrati senza la previa approvazione scritta di NetApp, Inc. I diritti di licenza del governo degli Stati Uniti per il Dipartimento della Difesa sono limitati ai diritti identificati nella clausola DFARS 252.227-7015(b) (FEB 2014).

## Informazioni sul marchio commerciale

NETAPP, il logo NETAPP e i marchi elencati alla pagina <http://www.netapp.com/TM> sono marchi di NetApp, Inc. Gli altri nomi di aziende e prodotti potrebbero essere marchi dei rispettivi proprietari.