



版本資訊 GenAI

NetApp
October 06, 2025

目錄

- 版本資訊 1
 - NetApp Workload Factory for GenAI 的新功能 1
 - 2025年10月5日 1
 - 2025年8月3日 1
 - 2025 年 6 月 29 日 2
 - 2025 年 6 月 03 日 2
 - 2025 年 04 月 5 日 2
 - 2025年3月2日 3
 - 2025 年 02 月 2 日 3
 - 2025 年 1 月 05 日 4
 - 2024 年 12 月 01 日 5
 - 2024 年 11 月 3 日 5
 - 2024 年 9 月 29 日 5
 - 2024 年 9 月 1 日 6
 - 2024 年 8 月 4 日 6
 - 2024 年 7 月 7 日 6

版本資訊

NetApp Workload Factory for GenAI 的新功能

了解 Workload Factory 的生成式 AI 工作負載功能的新功能。

2025年10月5日

BlueXP workload factory現為**NetApp**工作負載工廠

BlueXP已重新命名並重新設計，以更好地反映其在管理資料基礎架構中的作用。因此，BlueXP workload factory已重新命名為NetApp工作負載工廠。

支援在適用於 **Amazon Q Business** 的**NetApp**連接器中新增通用 **NFS/SMB** 資料來源

使用 Workload Factory API，您現在可以將資料來源從通用 NFSv3、NFSv4 或 SMB 共用新增至NetApp Connector for Amazon Q Business。這使您能夠包含儲存在由Amazon FSx for NetApp ONTAP以外的檔案系統所託管的磁碟區上的檔案。

["為 Amazon Q Business 建立 NetApp 連接器"](#)

["新增資料來源"](#)

知識庫的高階聊天配置

現在您可以配置適用於知識庫聊天模型的高級聊天設置，例如回應長度、溫度、推理設定等。其中一些設定（例如最近度和修改時間設定、進階檢索設定和系統提示）僅可使用 Workload Factory API。

["建立 GenAI 知識庫"](#)

現支援嵌入、聊天和重新排名模型的推理類型選擇

如果您選擇的嵌入、聊天或重新排名模型具有推理設置，您現在可以選擇推理類型。這使您能夠更好地根據您的需求調整聊天機器人的效能和資源需求。

["建立 GenAI 知識庫"](#)

2025年8月3日

結構化資料結果的安全存儲

如果聊天機器人查詢結果包含結構化數據，GenAI 可以將結果儲存在 Amazon S3 儲存桶中。當這些結果儲存在 S3 儲存桶中時，您可以使用聊天會話中的下載連結下載它們。

["建立 GenAI 知識庫"](#)

MCP 伺服器可用性

NetApp現在為 GenAI 提供具有NetApp Workload Factory 的模型上下文協定 (MCP) 伺服器。您可以在本機安裝

伺服器，以允許外部 MCP 用戶端從 GenAI 知識庫中發現和檢索查詢結果。

["NetApp Workload Factory GenAI MCP 伺服器"](#)

2025 年 6 月 29 日

支援通用 **NFS/SMB** 檔案系統上託管的資料來源

現在您可以從通用 SMB 或 NFS 共用中新增資料來源。這使您能夠包含儲存在 Amazon FSx for NetApp ONTAP 以外的檔案系統所託管的磁碟區上的檔案。

["在知識庫中新增資料來源"](#)

["新增資料來源"](#)

2025 年 6 月 03 日

Tracker 可用於監控及追蹤作業

GenAI 現已提供 Tracker 監控功能。您可以使用 Tracker 監控和追蹤待處理、正在進行和已完成操作的進度 and 狀態，查看操作任務和子任務的詳細信息，診斷任何問題或故障，編輯失敗操作的參數，以及重試失敗的操作。

["使用NetApp Workload Factory 中的 Tracker 監控工作負載操作"](#)

為知識庫選擇重新排序模型

現在，您可以透過選擇與知識庫配合使用的特定重排序模型來提高重排序查詢結果的相關性。GenAI 支援 Cohere Rerank 和 Amazon Rerank 模型。

["建立 GenAI 知識庫"](#)

2025 年 04 月 5 日

支援 **NetApp Connector for Amazon Q Business**

此版本的 GenAI 引入了對 NetApp Connector for Amazon Q Business 的支持，使您能夠為 Amazon Q Business 建立連接器。與為 Amazon Bedrock 建立 GenAI 知識庫相比，較少的初始組態，可快速輕鬆地利用 Amazon Q Business AI 助理。

["為 Amazon Q Business 建立 NetApp 連接器"](#)

增強的聊天模式支援

GenAI 現在支援下列額外的聊天模式以供知識庫使用：

- ["Mistral AI 機型"](#)
- ["Amazon Titan 文字模型"](#)
- ["中繼 Llama 機型"](#)
- ["Jamba 1.5 機型"](#)

- ["Cohere Command 模型"](#)
- ["Deepseek 機型"](#)

GenAI 支援 Amazon Bedrock 支援的每個供應商機型：["Amazon 基礎架構支援的基礎模型"](#)

["建立 GenAI 知識庫"](#)

更新的權限術語

Workload Factory 使用者介面和文件現在使用「只讀」來指涉讀取權限，使用「讀取/寫入」來指稱自動化權限。

2025年3月2日

嵌入式聊天機器人程式增強功能

您現在可以直接將問題和回應複製到剪貼簿，調整聊天視窗的大小，以及變更其標題。此外，聊天回應現在可以包含表格，這些表格也可以複製。

["測試 GenAI 知識庫"](#)

聊天回應引文支援

聊天回應現在包含的引文會列出用來產生回應的檔案和資料區塊。

["測試 GenAI 知識庫"](#)

增強的檔案類型支援

此版本的 GenAI 提供增強的檔案支援：

- 聊天模式提供更好的 CSV 支援。這可在從 CSV 檔案查詢資料時，提供更有用的回應。
- GenAI 現在可以從資料來源擷取 Apache Parquet 檔案。
- GenAI 現在支援擷取包含影像的 Microsoft Word DOCX 檔案。系統會掃描 DOCX 文件中內嵌的影像，並在回應知識庫查詢時，納入內嵌影像的文字洞見。

["支援的資料來源檔案格式"](#)

2025 年 02 月 2 日

支援 **Amazon Nova** 基礎模式

GenAI 現在支援 Amazon Nova 基礎模式。支援 Amazon Nova Micro，Amazon Nova Lite 和 Amazon Nova Pro。

["GenAI 需求"](#)

資料來源的檔案類型篩選

GenAI 現在支援在新增資料來源時，選取要納入資料來源掃描的特定檔案類型。

["將資料來源新增至知識庫"](#)

資料來源的檔案修改日期篩選

GenAI 現在支援篩選檔案，可在新增資料來源時，依修改日期納入資料來源掃描。您可以為包含的檔案選擇修改日期範圍。

["將資料來源新增至知識庫"](#)

支援影像檔案，並加強支援 PDF 檔案

GenAI 現在支援強化對知識庫查詢的回應，從影像和圖表說明以及文件文字中獲得深入見解，進而提供更豐富，更高品質的答案。GenAI 現在可以掃描 PDF 檔案中的影像檔案和影像（也稱為多模式檔案支援）。如果您選擇掃描影像或 PDF 檔案，則會將影像中的文字（包括內嵌於 PDF 文件中的影像）掃描至資料來源，而掃描結果中的洞見則會包含在知識庫查詢的回應中。

["將資料來源新增至知識庫"](#)

混合式搜尋與重新部署支援

GenAI 現在可以使用混合式搜尋並重新排列結果的順序，大幅提升搜尋結果的相關性和準確度。混合式搜尋結合了傳統關鍵字型搜尋的優勢，以及進階密集向量型語義搜尋技術。標準關鍵字搜尋結果會以接近的相符項目和語言差異加以擴充，以增強相關性。然後 GenAI 使用進階重新排名模型（例如 Cohere Rerank 和 Amazon Rerank）進一步精簡這些結果，並傳回最相關的結果。這項功能適用於新建立的知識庫。

["了解NetApp Workload Factory for GenAI"](#)

2025 年 1 月 05 日

自訂快照名稱

您現在可以為隨機操作快照提供快照名稱。

["利用快照保護知識庫"](#)

自訂 AI 引擎執行個體名稱

您現在可以在部署期間為 AI 引擎執行個體提供自訂名稱。

["部署 GenAI 基礎架構"](#)

重建毀損或遺失的 GenAI 基礎架構

如果您的 AI 引擎實例損壞或以某種方式被刪除，您可以讓 Workload Factory 為您重建它。重建完成後，工作負載工廠會自動將您的知識庫重新連接到基礎設施，以便隨時可以使用。

["疑難排解"](#)

2024 年 12 月 01 日

從快照複製知識庫

NetApp Workload Factory for GenAI 現在支援從快照複製知識庫。從而可以快速恢復知識庫，並利用現有資料來源創建新的知識庫，有助於資料復原和開發。

["複製知識庫"](#)

內部部署 **ONTAP** 叢集探索與複製

發現並將本機ONTAP叢集資料複製到 FSx for ONTAP檔案系統，以使用於豐富 AI 知識庫。所有本地發現和複製工作流程均可透過儲存清單中新的 **On-Premises ONTAP** 選單實現。

["探索內部部署 ONTAP 的叢集"](#)

2024 年 11 月 3 日

使用資料欄遮罩個人識別資訊

生成式人工智慧工作負載引入了由NetApp控制台分類提供支援的資料護欄功能。資料護欄功能可識別和封鎖個人識別資訊 (PII)，協助您保持合規性並加強敏感組織資料的安全性。

["建立 GenAI 知識庫"](#)

["了解NetApp控制台分類"](#)

2024 年 9 月 29 日

快照與還原支援知識庫磁碟區

您現在可以取得知識庫的時間點複本、以保護您的泛用 AI 工作負載資料。這可讓您保護資料、避免意外遺失或測試知識庫設定的變更。您可以隨時還原舊版的知識庫磁碟區。

["拍攝知識庫磁碟區的快照"](#)

["還原知識庫磁碟區的快照"](#)

暫停排程掃描

您現在可以暫停排程的資料來源掃描。根據預設、泛化 AI 工作負載會每天掃描每個資料來源、以便將新資料擷取到每個知識庫中。如果您不想擷取最新的變更（例如在測試期間或還原快照時）、您可以暫停排程的掃描、並隨時繼續進行。

["管理知識庫"](#)

知識庫現在支援資料保護磁碟區

選取知識庫磁碟區時、您現在可以選擇屬於 NetApp SnapMirror 複製關係一部分的資料保護磁碟區。這可讓您將知識庫儲存在已受 SnapMirror 複製保護的磁碟區上。

"識別要整合到知識庫中的資料來源"

2024 年 9 月 1 日

其他的分塊策略

泛用 AI 工作負載現在支援資料來源的多重句子區塊和重疊區塊。

每個知識庫的專用磁碟區

現在、泛用 AI 工作負載會為每個新知識庫建立專屬的 Amazon FSX for NetApp ONTAP Volume、為每個知識庫啟用個別的快照原則、並針對故障和資料中毒提供更好的保護。

2024 年 8 月 4 日

Amazon CloudWatch 記錄整合

現在、泛用 AI 工作負載已與 Amazon CloudWatch 記錄整合、可讓您監控泛用 AI 工作負載記錄檔。

例如：**chatbot** 應用程式

NetApp Workload Factory GenAI 範例應用程式可讓您透過在基於 Web 的聊天機器人應用程式中直接與已發佈的 NetApp Workload Factory 知識庫進行互動來測試其驗證和檢索。

2024 年 7 月 7 日

GenAI 工作負載工廠的初始版本

初始版本包含開發知識庫的功能、可藉由內嵌組織資料來自訂。使用者可透過聊天機器人應用程式存取知識庫。這項功能可確保針對組織特定問題做出準確且相關的回應、提高所有使用者的滿意度與生產力。

版權資訊

Copyright © 2025 NetApp, Inc. 版權所有。台灣印製。非經版權所有人事先書面同意，不得將本受版權保護文件的任何部分以任何形式或任何方法（圖形、電子或機械）重製，包括影印、錄影、錄音或儲存至電子檢索系統中。

由 NetApp 版權資料衍伸之軟體必須遵守下列授權和免責聲明：

此軟體以 NETAPP「原樣」提供，不含任何明示或暗示的擔保，包括但不限於有關適售性或特定目的適用性之擔保，特此聲明。於任何情況下，就任何已造成或基於任何理論上責任之直接性、間接性、附隨性、特殊性、懲罰性或衍生性損害（包括但不限於替代商品或服務之採購；使用、資料或利潤上的損失；或企業營運中斷），無論是在使用此軟體時以任何方式所產生的契約、嚴格責任或侵權行為（包括疏忽或其他）等方面，NetApp 概不負責，即使已被告知有前述損害存在之可能性亦然。

NetApp 保留隨時變本文所述之任何產品的權利，恕不另行通知。NetApp 不承擔因使用本文所述之產品而產生的責任或義務，除非明確經過 NetApp 書面同意。使用或購買此產品並不會在依據任何專利權、商標權或任何其他 NetApp 智慧財產權的情況下轉讓授權。

本手冊所述之產品受到一項（含）以上的美國專利、國外專利或申請中專利所保障。

有限權利說明：政府機關的使用、複製或公開揭露須受 DFARS 252.227-7013（2014 年 2 月）和 FAR 52.227-19（2007 年 12 月）中的「技術資料權利 - 非商業項目」條款 (b)(3) 小段所述之限制。

此處所含屬於商業產品和 / 或商業服務（如 FAR 2.101 所定義）的資料均為 NetApp, Inc. 所有。根據本協議提供的所有 NetApp 技術資料和電腦軟體皆屬於商業性質，並且完全由私人出資開發。美國政府對於該資料具有非專屬、非轉讓、非轉授權、全球性、有限且不可撤銷的使用權限，僅限於美國政府為傳輸此資料所訂合約所允許之範圍，並基於履行該合約之目的方可使用。除非本文另有規定，否則未經 NetApp Inc. 事前書面許可，不得逕行使用、揭露、重製、修改、履行或展示該資料。美國政府授予國防部之許可權利，僅適用於 DFARS 條款 252.227-7015(b)（2014 年 2 月）所述權利。

商標資訊

NETAPP、NETAPP 標誌及 <http://www.netapp.com/TM> 所列之標章均為 NetApp, Inc. 的商標。文中所涉及的所有其他公司或產品名稱，均為其各自所有者的商標，不得侵犯。