



Notas de lançamento

GenAI

NetApp

October 06, 2025

This PDF was generated from <https://docs.netapp.com/pt-br/workload-genai/whats-new.html> on October 06, 2025. Always check docs.netapp.com for the latest.

Índice

Notas de lançamento	1
Novidades do NetApp Workload Factory para GenAI	1
05 de outubro de 2025	1
03 de agosto de 2025	1
29 de junho de 2025	2
03 de junho de 2025	2
04 de maio de 2025	2
02 de março de 2025	3
02 de fevereiro de 2025	4
05 de janeiro de 2025	4
01 de dezembro de 2024	5
3 de novembro de 2024	5
29 de setembro de 2024	5
1 de setembro de 2024	6
4 de agosto de 2024	6
7 de julho de 2024	7

Notas de lançamento

Novidades do NetApp Workload Factory para GenAI

Saiba o que há de novo com o recurso de cargas de trabalho de IA generativa do Workload Factory.

05 de outubro de 2025

BlueXP workload factory agora NetApp Workload Factory

O BlueXP foi renomeado e redesenhado para refletir melhor o papel que ele tem no gerenciamento da sua infraestrutura de dados. Como resultado, a BlueXP workload factory foi renomeada para NetApp Workload Factory.

Supporte para adicionar fontes de dados NFS/SMB genéricas em conectores NetApp para Amazon Q Business

Usando a API do Workload Factory, agora você pode adicionar uma fonte de dados de um compartilhamento genérico NFSv3, NFSv4 ou SMB a um NetApp Connector para Amazon Q Business. Isso permite que você inclua arquivos armazenados em volumes hospedados por sistemas de arquivos diferentes do Amazon FSx for NetApp ONTAP.

["Crie um conector NetApp para o Amazon Q Business"](#)

["Adicionar fontes de dados a um conector"](#)

Configuração avançada de chat para bases de conhecimento

Agora você pode configurar definições avançadas de chat aplicáveis ao modelo de chat da base de conhecimento, como duração da resposta, temperatura, configurações de raciocínio e muito mais. Algumas dessas configurações, como configurações de tempo de modificação e atualidade, configurações de recuperação avançada e prompt do sistema, estão disponíveis somente usando a API do Workload Factory.

["Crie uma base de conhecimento do GenAI"](#)

A seleção do tipo de inferência agora é compatível com modelos de incorporação, bate-papo e reclassificação

Se o modelo de incorporação, bate-papo ou reclassificação escolhido tiver configurações de inferência, agora você pode selecionar um tipo de inferência. Isso permite que você ajuste melhor o desempenho do chatbot e os requisitos de recursos às suas necessidades.

["Crie uma base de conhecimento do GenAI"](#)

03 de agosto de 2025

Armazenamento seguro para resultados de dados estruturados

Se os resultados da consulta do chatbot contiverem dados estruturados, o GenAI poderá armazenar os resultados em um bucket do Amazon S3. Quando esses resultados são armazenados em um bucket S3, você pode baixá-los usando o link de download na sessão de bate-papo.

["Crie uma base de conhecimento do GenAI"](#)

Disponibilidade do servidor MCP

A NetApp agora fornece um servidor Model Context Protocol (MCP) com o NetApp Workload Factory para GenAI. Você pode instalar o servidor localmente para permitir que clientes MCP externos descubram e recuperem resultados de consultas de uma base de conhecimento do GenAI.

["Servidor NetApp Workload Factory GenAI MCP"](#)

29 de junho de 2025

Suporte para fontes de dados hospedadas em sistemas de arquivos NFS/SMB genéricos

Agora você pode adicionar uma fonte de dados de um compartilhamento SMB ou NFS genérico. Isso permite incluir arquivos armazenados em volumes hospedados por sistemas de arquivos diferentes do Amazon FSx para NetApp ONTAP.

["Adicionar fontes de dados a uma base de conhecimento"](#)

["Adicionar fontes de dados a um conector"](#)

03 de junho de 2025

Tracker disponível para operações de monitoramento e rastreamento

O recurso de monitoramento do Tracker agora está disponível no GenAI. Você pode usar o Tracker para monitorar e acompanhar o progresso e o status de operações pendentes, em andamento e concluídas, revisar detalhes de tarefas e subtarefas operacionais, diagnosticar problemas ou falhas, editar parâmetros de operações com falha e repetir operações com falha.

["Monitore as operações de carga de trabalho com o Tracker no NetApp Workload Factory"](#)

Escolha um modelo de reranker para uma base de conhecimento

Agora você pode aumentar a relevância dos resultados de consultas reclassificadas selecionando um modelo de reclassificação específico para usar com uma base de conhecimento. O GenAI suporta os modelos Cohere Rerank e Amazon Rerank.

["Crie uma base de conhecimento do GenAI"](#)

04 de maio de 2025

Suporte para NetApp Connector para Amazon Q Business

Esta versão do GenAI apresenta suporte ao NetApp Connector para Amazon Q Business, permitindo que você crie conectores para o Amazon Q Business. Aproveite de forma rápida e fácil o assistente de IA do Amazon Q Business com menos configuração inicial do que criar uma base de conhecimento do GenAI para a Amazon bedrock.

["Crie um conector NetApp para o Amazon Q Business"](#)

Suporte aprimorado ao modelo de chat

O GenAI agora suporta os seguintes modelos de bate-papo adicionais para bases de conhecimento:

- "Modelos Mistral AI"
- "Modelos de texto Amazon Titan"
- "Modelos Meta Llama"
- "Jamba 1,5 modelos"
- "Modelos Cohere Command"
- "Modelos Deepseek"

O GenAI suporta os modelos de cada provedor que a Amazon bedrock suporta: "[Modelos de base suportados na Amazon bedrock](#)"

["Crie uma base de conhecimento do GenAI"](#)

Terminologia de permissões atualizada

A interface do usuário e a documentação do Workload Factory agora usam "somente leitura" para se referir às permissões de leitura e "leitura/gravação" para se referir às permissões de automação.

02 de março de 2025

Aprimoramentos de chatbot incorporados

Agora você pode copiar perguntas e respostas diretamente para a área de transferência, ajustar o tamanho da janela de bate-papo e alterar seu título. Além disso, as respostas de bate-papo agora podem incluir tabelas, que também são copiáveis.

["Teste uma base de conhecimento do GenAI"](#)

Suporte a citações de resposta de chat

As respostas de bate-papo agora incluem citações que listam os arquivos e blocos de dados que foram usados para gerar a resposta.

["Teste uma base de conhecimento do GenAI"](#)

Suporte aprimorado ao tipo de arquivo

Esta versão do GenAI oferece suporte aprimorado a arquivos:

- Os modelos de chat apresentam suporte CSV melhorado. Isso permite respostas mais úteis ao consultar dados de arquivos CSV.
- O GenAI agora pode ingerir arquivos Apache Parquet a partir de fontes de dados.
- O GenAI agora suporta a introdução de arquivos DOCX do Microsoft Word que incluem imagens. As imagens incorporadas nos documentos DOCX são digitalizadas e os insights de texto das imagens incorporadas são incluídos nas respostas às consultas da base de conhecimento.

["Formatos de arquivo de origem de dados suportados"](#)

02 de fevereiro de 2025

Suporte para modelos de fundação Amazon Nova

O GenAI agora suporta os modelos de fundação Amazon Nova. São suportados Amazon Nova Micro, Amazon Nova Lite e Amazon Nova Pro.

["Requisitos do GenAI"](#)

Filtragem do tipo de arquivo para fontes de dados

O GenAI agora suporta a seleção de tipos de arquivo específicos para incluir na digitalização da fonte de dados quando você adiciona uma fonte de dados.

["Adicione fontes de dados à base de conhecimento"](#)

Filtragem de data de modificação de arquivo para fontes de dados

O GenAI agora suporta arquivos de filtragem para incluir na verificação da fonte de dados por data de modificação quando você adiciona uma fonte de dados. Você pode escolher um intervalo de datas de modificação para os arquivos incluídos.

["Adicione fontes de dados à base de conhecimento"](#)

Suporte para arquivos de imagem e suporte aprimorado para arquivos PDF

O GenAI agora oferece suporte ao aprimoramento de respostas a consultas de base de conhecimento com insights de imagens e descrições de gráficos, bem como texto de documentos, levando a respostas mais ricas e de maior qualidade. O GenAI agora pode digitalizar arquivos de imagem e imagens dentro de arquivos PDF (também conhecido como suporte a arquivos multimodais). Se você optar por digitalizar imagens ou arquivos PDF, o texto das imagens (incluindo imagens incorporadas em documentos PDF) é digitalizado na fonte de dados e os insights das digitalizações são incluídos nas respostas às consultas da base de conhecimento.

["Adicione fontes de dados à base de conhecimento"](#)

Pesquisa híbrida e suporte de reclassificação

O GenAI agora pode aumentar significativamente a relevância e a precisão dos resultados de pesquisa usando a pesquisa híbrida e reclassificando os resultados. A pesquisa híbrida combina os pontos fortes da pesquisa tradicional baseada em palavras-chave com técnicas avançadas de pesquisa semântica baseadas em vetores densos. Os resultados padrão de pesquisa de palavras-chave são aumentados com correspondências próximas e nuance linguística, aumentando a relevância. Em seguida, o GenAI refina esses resultados ainda mais usando modelos avançados de reclassificação, como cohore Rerank e Amazon Rerank, e retorna os resultados mais relevantes. Esta capacidade está disponível para bases de conhecimento recém-criadas.

["Saiba mais sobre o NetApp Workload Factory para GenAI"](#)

05 de janeiro de 2025

Nome do instantâneo personalizado

Agora você pode fornecer um nome de instantâneo para um instantâneo ad-hoc.

["Proteja uma base de conhecimento com snapshots"](#)

Nome de instância personalizado do mecanismo de AI

Agora você pode dar um nome personalizado à instância do mecanismo de AI durante a implantação.

["Implantar a infraestrutura do GenAI"](#)

Reconstruir a infraestrutura do GenAI corrompida ou ausente

Se a instância do seu mecanismo de IA for corrompida ou excluída de alguma forma, você pode deixar o Workload Factory reconstruí-la para você. O Workload Factory reconecta automaticamente suas bases de conhecimento à infraestrutura após a reconstrução ser concluída, para que elas estejam prontas para uso.

["Solução de problemas"](#)

01 de dezembro de 2024

Clonar uma base de conhecimento a partir de um instantâneo

O NetApp Workload Factory para GenAI agora oferece suporte à clonagem de uma base de conhecimento a partir de um snapshot. Isso permite a recuperação rápida de bases de conhecimento e a criação de novas bases de conhecimento com fontes de dados existentes, além de ajudar na recuperação e no desenvolvimento de dados.

["Clonar uma base de conhecimento"](#)

Detecção e replicação de clusters ONTAP no local

Descubra e replique dados de cluster ONTAP locais em um sistema de arquivos FSx for ONTAP para que eles possam ser usados para enriquecer bases de conhecimento de IA. Todos os fluxos de trabalho de descoberta e replicação no local são possíveis no novo menu **On-Premises ONTAP** no inventário de armazenamento.

["Descubra um cluster ONTAP no local"](#)

3 de novembro de 2024

Mascarar informações pessoais identificáveis com guardrails de dados

A carga de trabalho de IA generativa apresenta o recurso de proteção de dados, alimentado pela classificação do NetApp Console. O recurso de proteção de dados identifica e mascara Informações Pessoais Identificáveis (PII), ajudando você a manter a conformidade e fortalecer a segurança dos seus dados organizacionais confidenciais.

["Crie uma base de conhecimento do GenAI"](#)

["Saiba mais sobre a classificação do NetApp Console"](#)

29 de setembro de 2024

Supor te a snapshot e restauração para volumes da base de conhecimento

Agora, você pode proteger seus dados generativos de workloads de AI usando uma cópia pontual de uma base de conhecimento. Isso permite que você proteja seus dados contra perda accidental ou alterações de

teste nas configurações da base de conhecimento. Você pode restaurar a versão anterior do volume da base de conhecimento a qualquer momento.

["Tire um instantâneo de um volume da base de conhecimento"](#)

["Restaurar um snapshot de um volume da base de conhecimento"](#)

Pausar digitalizações programadas

Agora você pode pausar digitalizações de origem de dados agendadas. Por padrão, workloads de AI generativo varre cada fonte de dados diariamente para ingerir novos dados em cada base de conhecimento. Se você não quiser que as alterações mais recentes sejam ingeridas (durante o teste ou durante a restauração de um instantâneo, por exemplo), você pode pausar as verificações agendadas e retomá-las a qualquer momento.

["Gerenciar bases de conhecimento"](#)

Volumes de proteção de dados agora compatíveis com bases de conhecimento

Ao selecionar um volume da base de conhecimento, agora você pode escolher um volume de proteção de dados que faça parte de uma relação de replicação do NetApp SnapMirror. Isso permite armazenar bases de conhecimento em volumes que já estão protegidos pela replicação do SnapMirror.

["Identifique as fontes de dados a serem integradas em sua base de conhecimento"](#)

1 de setembro de 2024

Estratégias adicionais de divisão

Agora, as cargas de trabalho generativas de IA oferecem suporte a volumes de várias frases e conjuntos baseados em sobreposição para fontes de dados.

Volume dedicado para cada base de conhecimento

Agora, cria um volume dedicado do Amazon FSX for NetApp ONTAP para cada nova base de conhecimento, permitindo políticas de snapshot individuais para cada base de conhecimento e proteção aprimorada contra falhas e envenenamento de dados.

4 de agosto de 2024

Integração do Amazon CloudWatch Logs

As cargas de trabalho generativas de IA agora estão integradas ao Amazon CloudWatch Logs, permitindo que você monitore arquivos de log de cargas de trabalho generativas de IA.

Exemplo de aplicativo chatbot

O aplicativo de exemplo NetApp Workload Factory GenAI permite que você teste a autenticação e a recuperação da sua base de conhecimento publicada do NetApp Workload Factory interagindo diretamente com ela em um aplicativo de chatbot baseado na Web.

7 de julho de 2024

Lançamento inicial do Workload Factory para GenAI

A versão inicial inclui a capacidade de desenvolver uma base de conhecimento que é personalizada incorporando os dados da sua organização. A base de conhecimento pode ser acessada por um aplicativo de chatbot para seus usuários. Essa capacidade garante respostas precisas e relevantes a perguntas específicas da organização, aumentando a satisfação e a produtividade de todos os seus usuários.

Informações sobre direitos autorais

Copyright © 2025 NetApp, Inc. Todos os direitos reservados. Impresso nos EUA. Nenhuma parte deste documento protegida por direitos autorais pode ser reproduzida de qualquer forma ou por qualquer meio — gráfico, eletrônico ou mecânico, incluindo fotocópia, gravação, gravação em fita ou storage em um sistema de recuperação eletrônica — sem permissão prévia, por escrito, do proprietário dos direitos autorais.

O software derivado do material da NetApp protegido por direitos autorais está sujeito à seguinte licença e isenção de responsabilidade:

ESTE SOFTWARE É FORNECIDO PELA NETAPP "NO PRESENTE ESTADO" E SEM QUAISQUER GARANTIAS EXPRESSAS OU IMPLÍCITAS, INCLUINDO, SEM LIMITAÇÕES, GARANTIAS IMPLÍCITAS DE COMERCIALIZAÇÃO E ADEQUAÇÃO A UM DETERMINADO PROPÓSITO, CONFORME A ISENÇÃO DE RESPONSABILIDADE DESTE DOCUMENTO. EM HIPÓTESE ALGUMA A NETAPP SERÁ RESPONSÁVEL POR QUALQUER DANO DIRETO, INDIRETO, INCIDENTAL, ESPECIAL, EXEMPLAR OU CONSEQUENCIAL (INCLUINDO, SEM LIMITAÇÕES, AQUISIÇÃO DE PRODUTOS OU SERVIÇOS SOBRESSALENTES; PERDA DE USO, DADOS OU LUCROS; OU INTERRUPÇÃO DOS NEGÓCIOS), INDEPENDENTEMENTE DA CAUSA E DO PRINCÍPIO DE RESPONSABILIDADE, SEJA EM CONTRATO, POR RESPONSABILIDADE OBJETIVA OU PREJUÍZO (INCLUINDO NEGLIGÊNCIA OU DE OUTRO MODO), RESULTANTE DO USO DESTE SOFTWARE, MESMO SE ADVERTIDA DA RESPONSABILIDADE DE TAL DANO.

A NetApp reserva-se o direito de alterar quaisquer produtos descritos neste documento, a qualquer momento e sem aviso. A NetApp não assume nenhuma responsabilidade nem obrigação decorrentes do uso dos produtos descritos neste documento, exceto conforme expressamente acordado por escrito pela NetApp. O uso ou a compra deste produto não representam uma licença sob quaisquer direitos de patente, direitos de marca comercial ou quaisquer outros direitos de propriedade intelectual da NetApp.

O produto descrito neste manual pode estar protegido por uma ou mais patentes dos EUA, patentes estrangeiras ou pedidos pendentes.

LEGENDA DE DIREITOS LIMITADOS: o uso, a duplicação ou a divulgação pelo governo estão sujeitos a restrições conforme estabelecido no subparágrafo (b)(3) dos Direitos em Dados Técnicos - Itens Não Comerciais no DFARS 252.227-7013 (fevereiro de 2014) e no FAR 52.227- 19 (dezembro de 2007).

Os dados aqui contidos pertencem a um produto comercial e/ou serviço comercial (conforme definido no FAR 2.101) e são de propriedade da NetApp, Inc. Todos os dados técnicos e software de computador da NetApp fornecidos sob este Contrato são de natureza comercial e desenvolvidos exclusivamente com despesas privadas. O Governo dos EUA tem uma licença mundial limitada, irrevogável, não exclusiva, intransferível e não sublicenciável para usar os Dados que estão relacionados apenas com o suporte e para cumprir os contratos governamentais desse país que determinam o fornecimento de tais Dados. Salvo disposição em contrário no presente documento, não é permitido usar, divulgar, reproduzir, modificar, executar ou exibir os dados sem a aprovação prévia por escrito da NetApp, Inc. Os direitos de licença pertencentes ao governo dos Estados Unidos para o Departamento de Defesa estão limitados aos direitos identificados na cláusula 252.227-7015(b) (fevereiro de 2014) do DFARS.

Informações sobre marcas comerciais

NETAPP, o logotipo NETAPP e as marcas listadas em <http://www.netapp.com/TM> são marcas comerciais da NetApp, Inc. Outros nomes de produtos e empresas podem ser marcas comerciais de seus respectivos proprietários.