

Nieuwe Generative AI Content Control en uitbreiding van DLP AI-tools Coverage begrijpen

Inhoud

[Inleiding](#)

[Overzicht](#)

[Hoe kan DLP helpen bij het beheren van door ChatGPT gegenereerde inhoud?](#)

[Waarom AI-gegenereerde content beheren?](#)

[Hoe kan ik DLP-scanning toepassen op ChatGPT-reacties?](#)

[Wat is de categorie Generatieve AI-toepassingen in DLP?](#)

[Kan een DLP-regel worden toegepast op de volledige categorie Generative AI-toepassingen?](#)

[Waar kan ik relevante documentatie vinden?](#)

[Zijn we van plan om in de komende Cisco Live Amsterdam een aankondiging te doen over deze spannende Generative AI-bescherming use cases?](#)

Inleiding

Dit document beschrijft de nieuwe generatieve AI Content Control en de uitbreiding van de DLP AI-tools dekking voor Umbrella.

Overzicht

We zijn verheugd om de algemene beschikbaarheid van Generative AI Content Control aan te kondigen.

Met deze functie kunt u inhoud die door ChatGPT wordt gegenereerd, controleren en, indien nodig, blokkeren.

We zijn ook verheugd om te delen dat we de reikwijdte van onze Real-Time DLP-dekking voor Generatieve AI-tools hebben uitgebreid. Aanvankelijk beperkt tot ChatGPT, ondersteunen we nu alle 70 AI-tools in onze onlangs vrijgegeven categorie Generatieve AI-toepassingen. Deze aanzienlijke uitbreiding stelt u in staat om de AI Safe Usage use case te verbreden en biedt een uitgebreidere en robuustere oplossing voor de bescherming van Generative AI-gebruik.

Hoe kan DLP helpen bij het beheren van door ChatGPT gegenereerde inhoud?

DLP kan organisaties helpen bij het beheren van gegenereerde inhoud door ChatGPT-antwoorden te scannen met behulp van Real-Time DLP-beleid. Met deze release kunt u ervoor kiezen om ChatGPT-antwoorden (dat wil zeggen, inkomend verkeer) te scannen op elk type gegenereerde inhoud dat u wilt controleren of blokkeren.

Waarom AI-gegenereerde content beheren?

Het gebruik van AI-gegenereerde inhoud brengt om verschillende redenen risico's met zich mee voor organisaties, waaronder schending van het auteursrecht, onjuiste informatie, foutieve code, enzovoort.

U wilt bijvoorbeeld voorkomen dat gebruikers AI-gegenereerde broncode gebruiken om het gebruik van auteursrechtelijk beschermde of onveilige code te voorkomen, terwijl anderen het gebruik van AI-gegenereerde rechtbankcitataten willen voorkomen uit angst voor het indienen van onjuiste informatie.

Hoe kan ik DLP-scanning toepassen op ChatGPT-reacties?

Over het algemeen scant real-time DLP uitgaand webverkeer, zoals chatgpt-prompts, om gevoelige datalekken te voorkomen. Met deze release introduceren we de mogelijkheid om ook inkomend verkeer te scannen door de richting van het verkeer te kiezen die Real-Time DLP scant, dat wil zeggen inkomend verkeer, uitgaand verkeer of beide. Deze mogelijkheid is momenteel alleen beschikbaar voor ChatGPT (zowel chatbot als API). Kiezen om inkomende verkeer scant ChatGPT reacties.

Destinations

Manage destination lists and vetted applications for this rule.

- ☐ All Destinations
Selecting All Destinations will scan the traffic to any application or website the user is browsing to.
- ☒ Select Destinations Lists and Applications for Inclusion
Scans selected destination lists and vetted applications.

< Destinations / Application Categories / Generative AI

Direction

☒ OpenAI API (Vetted)

Inbound

☒ OpenAI ChatGPT (Vetted)

Outbound

Outbound

Inbound

Outbound & Inbound

3 Selected for Inclusion

REMOVE ALL

Applications Categories

OpenAI API / Generative AI, Inbound

OpenAI ChatGPT / Generative AI, Outbound

23281122679316

Wat is de categorie Generatieve AI-toepassingen in DLP?

Vóór deze release bevatten de bestemmingscriteria in de Real-Time DLP-regels een eindige selecteerbare lijst van ongeveer 20 toepassingen. Met deze release stelt Real-Time DLP klanten in staat om een van onze 38 toepassingscategorieën te kiezen, inclusief Generative AI, of een van de ±4.600 beschikbare controleerbare apps die erin zijn gecategoriseerd. De categorie Generatieve AI-applicaties, die slechts een paar maanden geleden werd gelanceerd met 20 apps, heeft nu 70 apps en we zijn vastbesloten om deze categorie voortdurend bij te werken met top-of-mind AI-tools.

Kan een DLP-regel worden toegepast op de volledige categorie Generative AI-toepassingen?

Ja, een Realtime DLP-regel kan worden toegepast op een hele categorie of op een subset van toepassingen binnen deze categorie.

Waar kan ik gerelateerde documentatie vinden?

- Om te leren hoe u de scanrichting kunt regelen om ChatGPT-antwoorden te controleren of te blokkeren, controleert u: [Voeg een Realtime Regel toe aan het Beleid ter voorkoming van gegevensverlies](#)
- Om te leren hoe u kunt controleren of het een chatGPT-prompt of chatGPT-reactie was die werd geblokkeerd, controleert u hier de scanrichting die rapporteert: [Rapport over preventie van gegevensverlies](#)
- Als u alle toepassingscategorieën wilt bekijken die nu beschikbaar zijn in real-time DLP-beleidsregels, controleert u hier: [Toepassingscategorieën](#)

Zijn we van plan om in de komende Cisco Live Amsterdam een aankondiging te doen over deze spannende Generative AI-bescherming use cases?

Ja, we gaan een breakout-sessie houden met de titel [Protecting Your Sensitive Data from Generative AI Usage](#) in Cisco Live Amsterdam, op dinsdag 6 februari, 15:00 - 16:30 uur CET.

Bewaar alsjeblieft je stoel!

Over deze vertaling

Cisco heeft dit document vertaald via een combinatie van machine- en menselijke technologie om onze gebruikers wereldwijd ondersteuningscontent te bieden in hun eigen taal. Houd er rekening mee dat zelfs de beste machinevertaling niet net zo nauwkeurig is als die van een professionele vertaler. Cisco Systems, Inc. is niet aansprakelijk voor de nauwkeurigheid van deze vertalingen en raadt aan altijd het oorspronkelijke Engelstalige document ([link](#)) te raadplegen.