

차세대 AI 콘텐츠 제어 및 DLP AI 툴 적용 범위 확장 이해

목차

[소개](#)

[개요](#)

[DLP를 통해 ChatGPT에서 생성된 콘텐츠를 제어하려면 어떻게 해야 하나요?](#)

[AI가 생성한 콘텐츠를 제어하는 이유는 무엇일까요?](#)

[ChatGPT 응답에 DLP 검사를 적용하려면 어떻게 해야 하나요?](#)

[DLP의 Generative AI 애플리케이션 범주는 무엇입니까?](#)

[DLP 규칙을 전체 Generative AI 애플리케이션 카테고리에 적용할 수 있습니까?](#)

[관련 문서는 어디에서 찾을 수 있습니까?](#)

[이러한 흥미로운 Generative AI 보호 활용 사례에 대해 곧 소개될 Cisco Live Amsterdam에서 발표할 계획입니까?](#)

소개

이 문서에서는 새로운 제너레이션 AI 콘텐츠 제어 및 Umbrella용 DLP AI 툴 적용 범위 확장에 대해 설명합니다.

개요

제너레이티브 AI 콘텐츠 컨트롤 상용화를 발표하게 되어 기쁘게 생각합니다. 이 기능을 사용하면 ChatGPT에서 생성된 콘텐츠를 모니터링하고 필요한 경우 차단할 수 있습니다.

또한 Generative AI 툴에 대한 실시간 DLP 적용 범위를 확장했다는 사실을 공유하게 되어 기쁘게 생각합니다. 초기에는 ChatGPT로 제한되어 있었으나, 이제 최근에 출시된 Generative AI 애플리케이션 범주에서 70가지 AI 툴을 모두 지원합니다. 이렇게 크게 확장되면 AI Safe Usage 활용 사례의 범위를 넓혀 Generative AI 사용 보호를 위한 더욱 포괄적이고 강력한 솔루션을 제공할 수 있습니다.

DLP를 통해 ChatGPT에서 생성된 콘텐츠를 제어하려면 어떻게 해야 하나요?

DLP는 실시간 DLP 정책을 사용하여 ChatGPT 응답을 검사함으로써 조직이 생성된 콘텐츠를 제어하는 데 도움을 줄 수 있습니다. 이 릴리스에서는 모니터링하거나 차단하려는 생성된 콘텐츠 유형에 대해 ChatGPT 응답(즉, 인바운드 트래픽)을 스캔하도록 선택할 수 있습니다.

AI가 생성한 콘텐츠를 제어하는 이유는 무엇일까요?

AI가 생성한 콘텐츠를 사용하면 저작권 침해, 부정확한 정보, 잘못된 코드 등 다양한 이유로 조직에 위험이 따릅니다.

예를 들어, 저작권이 있거나 안전하지 않은 코드를 사용하지 못하도록 사용자가 AI가 생성한 소스

코드를 사용하지 못하도록 하는 한편, 부정확한 정보를 제출하기 위해 AI가 생성한 법원 인용문을 사용하지 못하도록 하는 사용자가 있을 수 있습니다.

ChatGPT 응답에 DLP 검사를 적용하려면 어떻게 해야 합니까?

일반적으로 실시간 DLP는 ChatGPT 프롬프트와 같은 아웃바운드 웹 트래픽을 검사하여 민감한 데이터 유출을 방지합니다. 이번 릴리스에서는 실시간 DLP가 스캔하는 트래픽 방향(즉, 인바운드 트래픽, 아웃바운드 트래픽 또는 둘 다)을 선택하여 인바운드 트래픽도 스캔하는 기능을 도입합니다. 이 기능은 현재 ChatGPT(챗봇 및 API)에서만 사용할 수 있습니다. 인바운드 트래픽을 스캔하도록 선택하면 ChatGPT 응답이 스캔됩니다.

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

2 Selected for Inclusion

REMOVE ALL

Applications Categories

OpenAI API / Generative AI, Inbound

OpenAI ChatGPT / Generative AI, Outbound

23281122679316

DLP의 Generative AI 애플리케이션 범주는 무엇입니까?

이 릴리스 이전에는 실시간 DLP 규칙의 대상 기준에 약 20개 애플리케이션의 유한한 선택 가능한 목록이 포함되었습니다. 이번 릴리스에서는 실시간 DLP를 통해 고객이 Generative AI를 비롯한 38가지 애플리케이션 범주 중 하나를 선택하거나, 그 안에서 분류된 ±4,600개의 사용 가능한 제어 가능한 앱 중 하나를 선택할 수 있습니다. 불과 몇 달 전 20여개의 앱으로 출시된 제너레이티브 AI 애플리케이션 카테고리는 현재 70여개의 앱이 올라왔으며, 최우선 AI 톨로 지속적으로 이 카테고리를 업데이트하기 위해 노력하고 있다.

DLP 규칙을 전체 Generative AI 애플리케이션 카테고리에 적용할 수 있습니까?

예, 실시간 DLP 규칙을 전체 카테고리 또는 그 안의 애플리케이션 하위 집합에 적용할 수 있습니다.

관련 문서는 어디에서 찾을 수 있습니까?

- ChatGPT 응답을 모니터링 또는 차단하기 위해 스캐닝 방향을 제어하는 방법을 알아보려면 다음을 수행합니다.
[데이터 손실 방지 정책에 실시간 규칙 추가](#)
- 차단된 chatGPT 프롬프트 또는 chatGPT 응답 여부를 확인하는 방법을 알아보려면 검사 방향 보고를 여기서 확인하십시오. [데이터 유출 방지 보고서](#)
- 실시간 DLP 정책 규칙에서 현재 사용 가능한 모든 애플리케이션 범주를 검토하려면 여기를 확인하십시오. [애플리케이션 범주](#)

이러한 흥미로운 Generative AI 보호 활용 사례에 대해 곧 소개될 Cisco Live Amsterdam에서 발표할 계획입니까?

예, 2월 6일 화요일 오후 3:00 - 오후 4:30(CET) Cisco Live Amsterdam에서 [Protecting Your Sensitive Data from Generative AI Usage](#)(민감한 데이터를 생성적 AI 사용으로부터 보호)라는 제목의 브레이크아웃 세션을 열 예정입니다.

자리를 맡아 주십시오!

이 번역에 관하여

Cisco는 전 세계 사용자에게 다양한 언어로 지원 콘텐츠를 제공하기 위해 기계 번역 기술과 수작업 번역을 병행하여 이 문서를 번역했습니다. 아무리 품질이 높은 기계 번역이라도 전문 번역가의 번역 결과물만큼 정확하지는 않습니다. Cisco Systems, Inc.는 이 같은 번역에 대해 어떠한 책임도 지지 않으며 항상 원본 영문 문서(링크 제공됨)를 참조할 것을 권장합니다.