DELL Technologies

NVIDIA

5 najważniejszych kwestii związanych z bezpieczeństwem generatywnej sztucznej inteligencji (GenAI)

Przyspiesz wdrożenie bezpiecznej i skalowalnej infrastruktury dzięki rozwiązaniu Dell AI Factory with NVIDIA

Transformacyjny potencjał generatywnej sztucznej inteligencji

Generatywna sztuczna inteligencja ma potencjał, aby zmienić zasady gry w sposób, który wizjonerzy dopiero zaczynają sobie wyobrażać.

76%

liderów IT i biznesu uważa, że generatywna sztuczna inteligencja przyniesie ich organizacji wartość transformacyjną.[1]

AI

Zaawansowane analizy i techniki oparte na logice, które interpretują zdarzenia oraz wspierają i automatyzują decyzje i działania.

Generatywna sztuczna inteligencja

Technologie i techniki, które wykorzystują duże ilości danych do generowania nowych treści na podstawie podpowiedzi języka naturalnego lub innych niezwiązanych z kodem i nietradycyjnych danych wejściowych.

Symulacje

- Cyfrowy bliźniak
- Dane syntetyczne
- Ramy projektowe
- Prognozowanie

Tworzenie treści

- Kodowanie
- Matematyka
- Pisanie/mowa
- Obraz/wideo
- Dźwięk

Wykrywanie treści

- Wyszukiwanie z wykorzystaniem języka naturalnego
- Analiza dużych zbiorów danych
- Zarządzanie wiedzą
- Spersonalizowana edukacja i szkolenia

Wygoda użytkowników

- Tłumaczenia w czasie rzeczywistym na ponad 70 języków
- Spersonalizowane interakcje z wykorzystaniem naturalnej mimiki twarzy i mowy ciała

1 Badanie Innovation Catalyst firmy Dell Technologies, luty 2024 r.

Większy potencjał, większe ryzyko

Dla liderów biznesowych kuszące jest szybkie działanie, z ominięciem implikacji związanych z danymi, zagrożeń dotyczących zgodności z przepisami, zarządzania i innych problemów. Jednakże GenAI to miecz obosieczny, jeśli chodzi o bezpieczeństwo.

Korzyści

- Ulepszone wykrywanie zagrożeń
- Wyższa wydajność operacyjna
- Szkolenie dotyczące spersonalizowanych usług zabezpieczeń

Wady

- Większe wyrafinowanie ataków
- Zaawansowana inżynieria społeczna
- Ukryte AI

33%

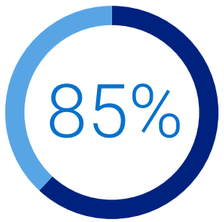
respondentów wymieniło cyberbezpieczeństwo jako główne ryzyko związane z generatywną sztuczną inteligencją, które ich organizacje starają się ograniczyć.[2]

2 Globalne badanie McKinsey na temat sztucznej inteligencji: The state of AI in early May 2024

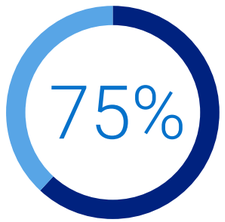
KWESTIA 1

Nowy krajobraz zagrożeń

Wraz z obietnicą GenAI pojawia się otrzeźwiająca rzeczywistość: atakujący tworzą nowe i bardziej skomplikowane ataki, które mogą omijać konwencjonalne zabezpieczenia, co utrudnia pracę zespołom ds. cyberbezpieczeństwa.



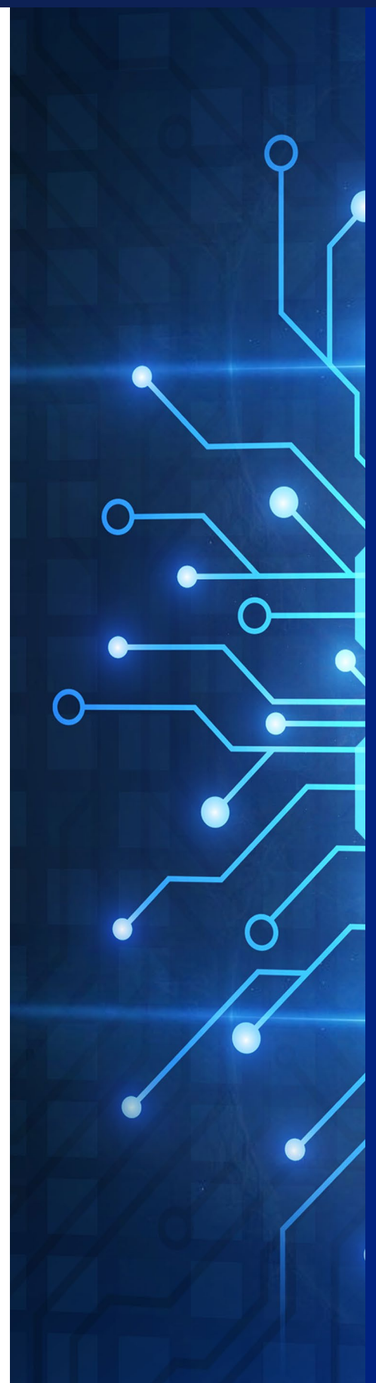
respondentów uważa, że sztuczna inteligencja sprawiła, że ataki cybernetyczne stały się bardziej wyrafinowane.[3]



specjalistów ds. bezpieczeństwa odnotowało wzrost liczby ataków w ciągu ostatnich 12 miesięcy.[4]

Aby chronić się przed tymi nowymi zagrożeniami, firmy muszą skupić się na minimalizacji powierzchni ataku, na przykład poprzez testy penetracyjne, monitorowanie i audyty.

3 Badanie „Human Risk in Cybersecurity Survey 2024", EY, maj 2024 r.
4 Raport Voice of SecOps „Generative AI and Cybersecurity: Bright Future or Business Battleground?" 2023



Nowe wektory ataków

Zaawansowane złośliwe oprogramowanie
Coraz bardziej zaawansowane złośliwe oprogramowanie, które wykorzystuje generatywną sztuczną inteligencję do „samoewolucji", stale zmieniając swój kod, aby nie było wykrywane przez istniejące zabezpieczenia, takie jak wykrywanie oparte na sygnaturach.

Wysoce spersonalizowane wiadomości e-mail i kampanie phishingowe
Coraz częstsze pojawianie się autentycznie wyglądających złośliwych wiadomości e-mail, które nie zawierają typowych oznak oszustwa.

Przekonujące dane typu deep fake
Kradzieże tożsamości, oszustwa finansowe i akcje dezinformacyjne są łatwiejsze dzięki możliwości naśladowania ludzkich działań, takich jak pisanie, mowa, obrazy lub wideo.

Zautomatyzowane rozeznanie
Gromadzenie informacji pod kątem znajdowania luk i słabości sieci lub systemu potencjalnego celu i ułatwiania bardziej ukierunkowanych ataków.

KWESTIA 2

Ryzyko związane z wdrożeniem

Organizacje, które chcą wykorzystać potencjalne korzyści płynące z generatywnej sztucznej inteligencji, potrzebują dużych ilości danych wysokiej jakości — danych wejściowych, które mogą być wykorzystane przez modele do uzyskania najlepszych wyników. Dane i ryzyko są jednak ze sobą nierozerwalnie związane. Przed wykorzystaniem jakichkolwiek informacji firmy muszą dokładnie ocenić i uwzględnić swoje niepowtarzalne wymagania, dane wejściowe i ryzyko.

Luki w zabezpieczeniach dużych modeli językowych (LLM)

Usługi generatywnej sztucznej inteligencji są podatne na ataki szybkich wstrzyknięć, w których osoby atakujące manipulują danymi wyjściowymi, aby ominąć zabezpieczenia lub uzyskać nieautoryzowany dostęp do plików, które mogły zostać użyte do udoskonalania modelu.

Zatruwanie danych

Atakujący mogą celowo przekazywać zmienione dane do LLM podczas fazy trenowania. Może to prowadzić do tego, że model będzie podatny na ataki za pośrednictwem wirusów typu backdoor, osadzonych w danych. Rzeczywistym przykładem jest atakowanie i wykorzystywanie filtrów antyspamowych poprzez trenowanie ich na wiadomościach spamowych.

Złożoność przepisów

Organy regulacyjne na całym świecie ścigają się, aby poznawać, kontrolować i gwarantować bezpieczeństwo generatywnej sztucznej inteligencji. Podczas, gdy modele generatywnej sztucznej inteligencji podlegają obecnym zasadom suwerenności danych, które określają, w jaki sposób dane są przechowywane, przetwarzane i wykorzystywane, organy zarządzające nadal definiują nadzór nad własnością intelektualną i informacjami chronionymi prawem autorskim. Przestrzeganie regulacji może być kosztowne, ale nieprzestrzeganie ustalonych i pojawiających się regulacji może skutkować grzywnami i innymi karami.

KWESTIA 3

Shadow AI

Wielu pracowników korzysta już dziś z publicznych generatorów tekstu, obrazów i wideo, takich jak ChatGPT, aby usprawnić codzienną pracę. Jednak gdy narzędzia te są używane bez odpowiedniego zarządzania, stanowią krytyczne zagrożenie dla organizacji próbujących zabezpieczyć własność intelektualną i dane korporacyjne. Takie nieautoryzowane użycie GenAI jest znane jako Shadow AI.

Utrata własności intelektualnej

Już teraz firmy borykają się z utratą własności intelektualnej na skutek udostępniania przez pracowników danych wrażliwych w publicznych narzędziach generatywnej sztucznej inteligencji.

Wyciek danych z kodu źródłowego

Programiści próbujący zoptymalizować kod źródłowy za pomocą ChatGPT spowodowali wyciek danych.

Aby sprostać wyzwaniom związanym z Shadow AI, firmy powinny ustanowić radę lub komisję obejmującą całą firmę, która będzie miała uprawnienia do podejmowania decyzji dotyczących bezpiecznego zarządzania sztuczną inteligencją.

Gdzie znajdują się Twoje dane? Gdzie należy umieścić obciążenia robocze?

Sztuczna inteligencja działa najlepiej, gdy jest połączona z Twoimi danymi, bez względu na miejsce. Dzięki pełnej kontroli nad infrastrukturą i LLM nie ma ryzyka utraty adresu IP czy wycieku danych z kodu źródłowego.

Koszty

Wykorzystanie wdrożeń lokalnych może obniżyć całkowity koszt użytkowania nawet o 75% w ciągu 3 lat.[5]

Bezpieczeństwo i prywatność

Twórz bezpieczne środowiska sztucznej inteligencji/generatywnej sztucznej inteligencji w całej organizacji dzięki lokalnym przepływom pracy i operacjom. Sprawuj ścisłą kontrolę nad bezpieczeństwem danych i przestrzegaj regulacji dotyczących zgodności, szczególnie w branżach, które przetwarzają dane wrażliwe.

5 Na podstawie badania Enterprise Strategy Group zleconego przez firmę Dell, w ramach którego porównano lokalną infrastrukturę Dell z natywną infrastrukturą chmury publicznej jako usługi, kwiecień 2024 r. Analizowane modele pokazują, że LLM z 7 mld parametrów wykorzystujący RAG dla organizacji 5 tys. użytkowników jest do 38% bardziej opłacalny, podczas gdy LLM z 70 mld parametrów wykorzystujący RAG dla organizacji 50 tys. użytkowników jest do 75% bardziej opłacalny. Faktyczne wyniki mogą być inne. Podsumowanie ekonomiczne

KWESTIA 4

Kryteria oceny

W ciągu ostatniego roku społeczność AI w coraz większym stopniu koncentrowała się na trzech kluczowych kwestiach: odpowiedzialnym rozwoju i wdrażaniu, ocenie wpływu i ograniczaniu ryzyka. Oceniając modele GenAI, firmy muszą brać pod uwagę kilka ważnych zastrzeżeń:

Brak spójnych wymogów dotyczących sprawozdawczości

Czołowi programiści testują swoje modele przede wszystkim w odniesieniu do różnych testów porównawczych odpowiedzialnej sztucznej inteligencji. Ze względu na ten znaczny brak standaryzacji w raportowaniu trudno jest metodycznie porównywać ryzyko i ograniczenia najlepszych modeli sztucznej inteligencji.

Materiały chronione prawem autorskim w danych wyjściowych

Wyniki popularnych dużych modeli językowych mogą zawierać materiały chronione prawem autorskim, potencjalnie naruszając prawo i narażając firmy korzystające z tych materiałów na ryzyko kar.

Luki w zabezpieczeniach są coraz bardziej złożone

Naukowcy znajdują mniej oczywiste strategie, które powodują, że duże modele językowe wykazują szkodliwe zachowania, takie jak proszenie modeli o niekończące się powtarzanie losowych słów.

Programistom brakuje przejrzystości

W wielu przypadkach twórcy sztucznej inteligencji nie są otwarci na swoje dane i metodologie szkoleniowe. Utrudnia to dążenia do lepszego zrozumienia solidności i bezpieczeństwa systemów sztucznej inteligencji.

KWESTIA 5

Korzyści w zakresie bezpieczeństwa

Obok zagrożeń dla bezpieczeństwa w związku z GenAI istnieją też potencjalne korzyści. Generatywna sztuczna inteligencja staje się kluczowym sprzymierzeńcem w dziedzinie cyberbezpieczeństwa, otwierając nowe możliwości ochrony.

Teraz można rozpocząć tworzenie skalowalnych operacji bezpieczeństwa dzięki szybszemu dostępowi do bogatszych informacji i automatycznemu wykrywaniu zagrożeń, zapewniając wydajność i uzupełniając zespoły ds. bezpieczeństwa, w których brakuje personelu.

Wykrywanie zagrożeń i reagowanie na nie

Analizując dane historyczne oraz identyfikując wzorce i anomalie, generatywna sztuczna inteligencja może rozpoznawać nowe i ewoluujące zagrożenia w czasie rzeczywistym. Może stale monitorować ruch sieciowy, dzienniki systemowe i zachowanie użytkowników oraz szybko identyfikować nietypowe działania, które mogą oznaczać zagrożenia bezpieczeństwa.

Rezultatem jest w pełni adaptacyjne wykrywanie zagrożeń, umożliwiające szybką reakcję na zmieniające się wektory ataków i zapewniające proaktywny mechanizm obrony przed pojawiającymi się zagrożeniami cybernetycznymi.

Symulacja zagrożeń i trening

Dzięki GenAI firmy mogą symulować szeroki zakres zagrożeń cyberbezpieczeństwa i scenariuszy ataków w kontrolowanym środowisku. W rezultacie zespoły są lepiej przygotowane do identyfikowania zagrożeń cybernetycznych, reagowania na nie i łagodzenia ich, gdy czas ma kluczowe znaczenie.

Dogłębna analiza i podsumowanie

Generatywna sztuczna inteligencja umożliwia zespołom badanie danych z różnych źródeł lub modułów, umożliwiając im szybsze i dokładniejsze przeprowadzanie tradycyjnie czasochłonnej, żmudnej analizy danych. Zespoły mogą również tworzyć podsumowania incydentów i ocen zagrożeń w języku naturalnym, poprawiając wydajność i zwiększając wydajność zespołu.

Szkolenie dotyczące spersonalizowanych usług zabezpieczeń

Łącząc konwersacyjną sztuczną inteligencję z generatywną sztuczną inteligencją i włączając awatara sztucznej inteligencji do interfejsu użytkownika, organizacje mogą dostarczać spersonalizowane interakcje (dostępne na dużą skalę 24 godziny na dobę, 7 dni w tygodniu) przy użyciu naturalnej mimiki twarzy i mowy ciała. Można to wykorzystać do szkoleń i edukacji w zakresie bezpieczeństwa, zapewniając bardziej naturalne, dostosowane i interaktywne środowisko uczenia się, automatyczne oceny i nie tylko.

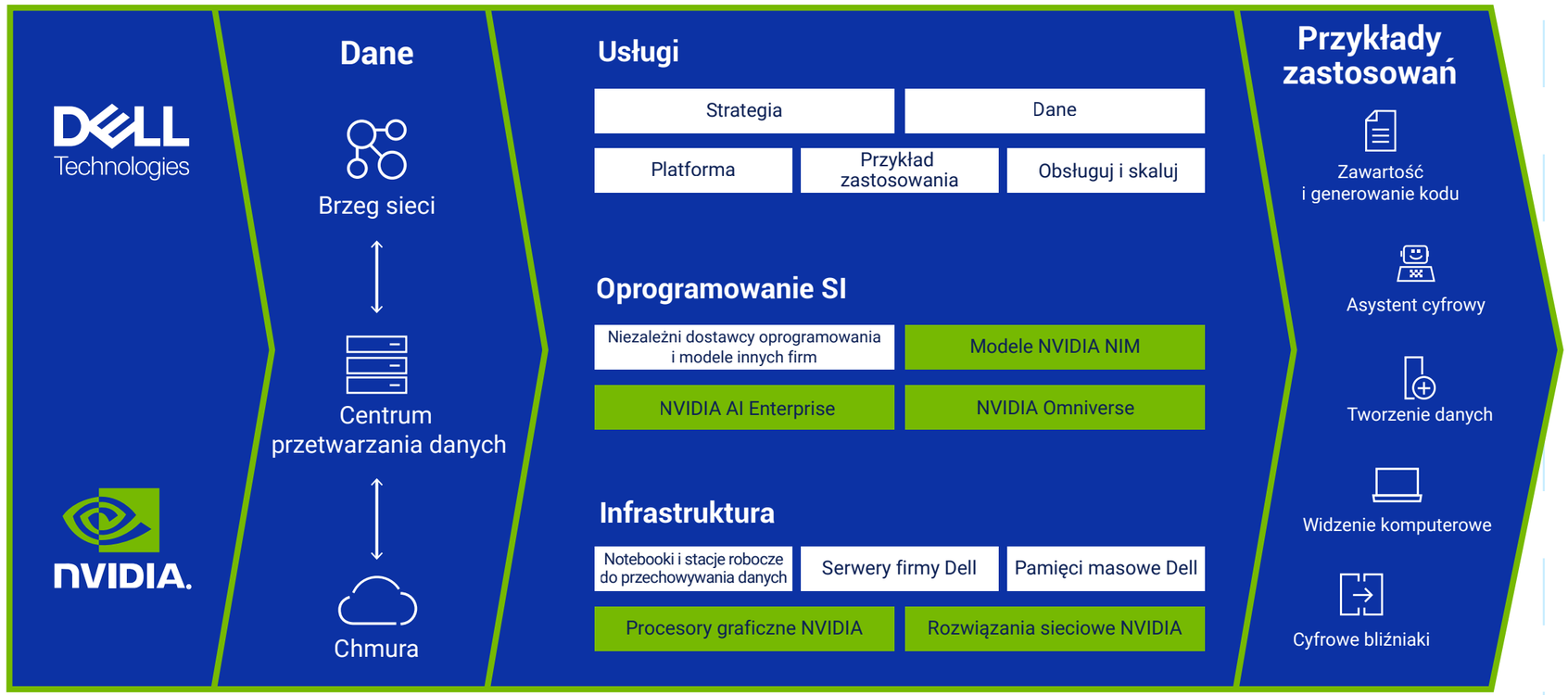
Dell AI Factory with NVIDIA

Przyspiesz proces wdrażania sztucznej inteligencji i bezpiecznie przekształcaj dane w szczegółowe informacje, korzystając z pierwszego kompleksowego, gotowego do użycia rozwiązania sztucznej inteligencji. Rozwiązanie Dell AI Factory with NVIDIA jest odpowiedzią na złożone potrzeby przedsiębiorstw, które chcą wykorzystywać sztuczną inteligencję i generatywną sztuczną inteligencję. Dzięki wiodącej infrastrukturze i usługom oraz oprogramowaniu do sztucznej inteligencji NVIDIA można skrócić czas do uzyskania wartości w projektach, upraszczając proces programowania i wdrażania.

- Zmniejsz ryzyko naruszenia bezpieczeństwa dzięki infrastrukturze z zabezpieczeniami wewnętrznymi, w tym zaufanymi źródłami i innymi kluczowymi funkcjami.
- Chroń swoje dane przed wyciekiem, który mógłby spowodować utratę własności intelektualnej, dzięki lokalnemu rozwiązaniu AI, które kontrolujesz.
- Spełniaj ścisłe wymagania dotyczące zgodności z przepisami i suwerenności danych, wprowadzając sztuczną inteligencję do danych z bezpiecznym dostępem.
- Chroń prywatność interesariuszy, kontrolując, kto i gdzie ma dostęp do Twoich danych.

Dell AI Factory with NVIDIA

PIERWSZE W BRANŻY KOMPLEKSOWE ROZWIĄZANIE SZTUCZNEJ INTELIGENCJI DLA PRZEDSIĘBIORSTW



Dane stymulują fabrykę sztucznej inteligencji i przykłady zastosowań

Najcenniejsze dane są przechowywane lokalnie i na brzegu. Firma Dell Technologies pomaga wdrażać sztuczną inteligencję w zakresie przetwarzania najcenniejszych danych i jest liderem pod względem ich przechowywania i ochrony oraz zarządzania nimi.

Od zastosowania do wyników

Fabryka sztucznej inteligencji generuje wyniki biznesowe w oparciu o priorytetowe przykłady zastosowań. Dell Technologies upraszcza wdrażanie najważniejszych przykładów zastosowań sztucznej inteligencji dzięki sprawdzonym rozwiązaniom i usługom dopasowanym do potrzeb.

Nie pozwól, by zagrożenia bezpieczeństwa stłumiły innowacje

Pomożemy Ci poruszać się po świecie sztucznej inteligencji i generatywnej sztucznej inteligencji, a w konsekwencji czerpać z tego korzyści.

PLANOWANIE STRATEGII

Bezpłatne warsztaty Accelerator Workshop poświęcone generatywnej sztucznej inteligencji

- Rozpocznij swoją podróż do opracowania zwycięskiej strategii
- Rozwiązywanie problemów i luk, ustalanie priorytetów celów i identyfikowanie możliwości
- Uzyskaj ocenę gotowości, aby zdobyć szczegółowe informacje na temat wymagań dotyczących infrastruktury, modeli sztucznej inteligencji, integracji operacyjnych i nie tylko

PRZYGOTOWANIE TECHNICZNE

Gotowe do użycia mobilne laboratorium

Rozpocznij swoją drogę do sukcesu. Obejmuje mobilną stację roboczą Dell Precision 5690/7780 z procesorami graficznymi NVIDIA i dwa dni konsultacji ułatwiających rozpoczęcie pracy.

- Przenośne środowisko piaskownicy do testowania i demonstracji generatywnej sztucznej inteligencji
- Wstępnie zweryfikowano za pomocą platformy NVIDIA AI Workbench, gotowej dla programistów
- Przykład pierwszego zastosowania chatbota, zaimplementowanego Twoimi danymi
- Opłacalne podejście o niskim ryzyku do eksperymentowania i rozwijania umiejętności generatywnej sztucznej inteligencji

MOBILNA STACJA ROBOCZA DELL PRECISION 5690/7780 Z PROCESORAMI GRAFICZNYMI NVIDIA

ZACZNIJ JUŻ DZIŚ

DELL Technologies AI Factory WITH NVIDIA