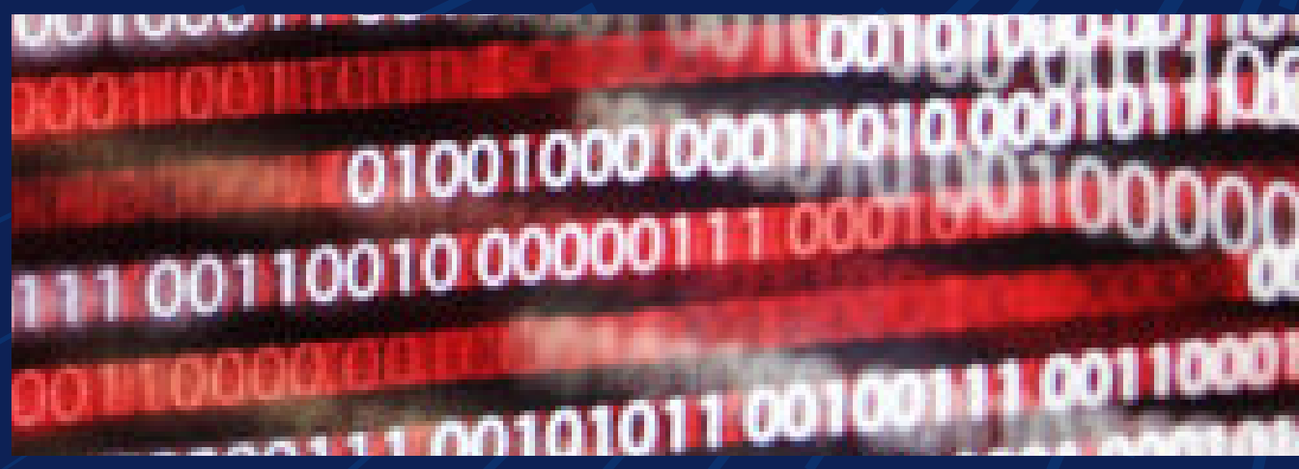
Mity dotyczące cyberbezpieczeństwa:

Demitologizacja mitów o bezpieczeństwie AI



Sztuczna inteligencja zmienia branże, ale jeśli chodzi o zabezpieczenie sztucznej inteligencji, wiele organizacji pada ofiarą mitów, które sprawiają, że wydaje się ona bardziej złożona, niż jest w rzeczywistości. Prawda? Ochrona systemów AI nie wymaga zaczynania od zera — zastosowanie istniejących zasad cyberbezpieczeństwa do unikalnych wyzwań AI jest bardzo skuteczne.

W Dell Technologies rozumiemy architekturę stojącą za AI i możemy Ci pomóc dostosować Twoje obecne rozwiązania do tego nowego modelu. Obalmy najczęstsze mity dotyczące bezpieczeństwa sztucznej inteligencji i odkryjmy prawdziwe wiadomości, które pomogą skutecznie zabezpieczyć systemy.

Mit 1: „Systemy sztucznej inteligencji są zbyt złożone, aby je zabezpieczyć”.

Prawda: To prawda, że sztuczna inteligencja stwarza nowe zagrożenia dla cyberbezpieczeństwa, takie jak szybkie wstrzyknięcia, manipulacje danymi i ujawnianie poufnych informacji, aby wymienić tylko kilka. Ponadto systemy Agentic AI mają również szerszą powierzchnię ataku, ponieważ można je wykorzystać do manipulowania wynikami lub eskalacji uprawnień.

Mimo to, chociaż ważne jest, aby znać te słabości i wdrożyć środki bezpieczeństwa w celu ochrony systemów AI przed tradycyjnymi i właściwymi dla AI zagrożeniami, ryzyko można kontrolować, a modele AI można zabezpieczyć. Ważne jest, aby pamiętać, że systemy sztucznej inteligencji wymagają znacznych ilości danych jako danych wejściowych i tworzą duże ilości danych jako danych wyjściowych. Z tego powodu jedną z kluczowych strategii bezpieczeństwa jest ochrona danych, a także:

- Zasady „zero trust”, takie jak zarządzanie tożsamością, dostęp oparty na rolach i ciągła weryfikacja.
- Regularne testy penetracyjne i zarządzanie lukami w zabezpieczeniach w celu identyfikacji słabych punktów.
- Rejestrowanie i kontrolowanie w celu weryfikacji danych wejściowych i wyjściowych

Silna obrona zaczyna się od dobrej higieny cybernetycznej, takiej jak poprawki systemowe, kontrola dostępu i zarządzanie lukami w zabezpieczeniach. Ważne jest dostosowanie tych praktyk do ryzyka związanego ze sztuczną inteligencją. Dzięki strategiom opartym na sztucznej inteligencji zintegrowanym z obecnym podejściem do bezpieczeństwa i odpowiednim narzędziom bezpieczeństwo oparte na sztucznej inteligencji staje się łatwe w zarządzaniu i skuteczne.

Trzeba jednak podkreślić, że nowy sprzęt może odegrać kluczową rolę w walce z cyberatakami. Na przykład nowoczesne komputery ze sztuczną inteligencją tworzą silną pierwszą linię obrony przed głównym wektorem ataku: ataku na punkty końcowe. Po zakończeniu obsługi systemu Windows 10 nieaktualne komputery stają się zagrożeniem. Ponadto system Windows 11 wymaga modułu TPM (Trusted Platform Module) w wersji 2.0, chipa zabezpieczającego, który pomaga w szyfrowaniu, bezpiecznym rozruchu i ochronie przed atakami na oprogramowanie wewnętrzne. Wiele starszych komputerów w ogóle nie ma modułu TPM lub obsługuje tylko starszą wersję. Firma Dell oferuje bezpieczne komercyjne komputery ze sztuczną inteligencją z wbudowanymi ulepszeniami.

To samo dotyczy infrastruktury sztucznej inteligencji, takiej jak serwery i pamięć masowa. Dell AI Factory obejmuje sprzęt zoptymalizowany pod kątem bezpieczeństwa sztucznej inteligencji i zawiera szereg wbudowanych funkcji zabezpieczeń, od bezpiecznego łańcucha dostaw po niezmienność danych po izolację i szyfrowanie.

Mit 2: „Żadne z moich istniejących narzędzi nie zabezpieczy sztucznej inteligencji”.

Prawda: Bezpieczeństwo sztucznej inteligencji nie polega na zaczynaniu od nowa — chodzi o inteligentniejsze wykorzystanie narzędzi, które już masz. Większość istniejących narzędzi cyberbezpieczeństwa można dostosować do skutecznego zabezpieczania systemów sztucznej inteligencji. W swojej istocie sztuczna inteligencja to kolejny typ obciążenia roboczego wspierającego rozwój Twojej firmy, choć ma unikatowe cechy. Podstawowe praktyki w zakresie cyberbezpieczeństwa, takie jak zarządzanie tożsamością, segmentacja i monitorowanie sieci, ochrona punktów końcowych i ochrona danych, pozostają niezbędne do ochrony środowisk sztucznej inteligencji. Kluczem jest dostosowanie tych praktyk do konkretnych wyzwań związanych ze sztuczną inteligencją, takich jak ochrona danych szkoleniowych, zabezpieczanie algorytmów i ograniczanie ryzyka, takiego jak dane wejściowe przeciwników.

Mit 3: „Bezpieczeństwo sztucznej inteligencji polega tylko na ochronie danych”.

Prawda: Bezpieczeństwo sztucznej inteligencji wykracza poza podstawową ochronę danych — obejmuje ochronę całego ekosystemu sztucznej inteligencji, w tym modeli, interfejsów API, danych wyjściowych, systemów i urządzeń. Wraz z coraz intensywniejszym wdrażaniem sztucznej inteligencji w krytycznych zastosowaniach ryzyko związane z jej niewłaściwym wykorzystaniem lub nadużyciem rośnie. Bez solidnych środków bezpieczeństwa modele sztucznej inteligencji mogą być manipulowane w celu generowania szkodliwych lub wprowadzających w błąd danych wyjściowych, interfejsy API mogą być wykorzystywane do uzyskania nieautoryzowanego dostępu do wrażliwych systemów, a dane wyjściowe mogą nieumyślnie ujawniać prywatne lub poufne informacje.

Kompleksowe zabezpieczenia sztucznej inteligencji wymagają wielowarstwowego podejścia. Obejmuje to ochronę modeli przed wrogimi atakami, które próbują manipulować danymi wejściowymi w celu oszukania systemów sztucznej inteligencji, zabezpieczenie interfejsów API za pomocą silnych metod uwierzytelniania, aby zapobiec nieautoryzowanemu użyciu, oraz **ciągłe monitorowanie danych wyjściowych** pod kątem nietypowych lub podejrzanych wzorców, które mogą sygnalizować atak lub awarię. Skuteczne zabezpieczenia sztucznej inteligencji nie tylko zapewniają integralność i niezawodność systemów AI, ale także budują zaufanie użytkowników i interesariuszy, ograniczając ryzyko złośliwego użycia lub niezamierzonych konsekwencji.

Mit 4: „Sztuczna inteligencja nie wymaga ludzkiego nadzoru”.

Prawda: Zarządzanie i nadzór człowieka mają kluczowe znaczenie dla gwarancji, że systemy sztucznej inteligencji działają etycznie, przewidywalnie i zgodnie z ludzkimi wartościami. Zaawansowane systemy sztucznej inteligencji, w szczególności sztuczna inteligencja z autonomicznymi możliwościami podejmowania decyzji, wprowadzają wyjątkowe wyzwania, które wymagają solidnych zabezpieczeń. Bez odpowiedniego nadzoru systemy te mogą odbiegać od zamierzonych celów lub wykazywać niezamierzone zachowania, które mogą stwarzać ryzyko.

Aby rozwiązać ten problem, konieczne jest ustalenie jasnych granic, wdrożenie warstwowych mechanizmów kontroli i zapewnienie ciągłego zaangażowania człowieka w kluczowe procesy podejmowania decyzji. Regularne audyty, przejrzystość operacji sztucznej inteligencji i dokładne testy mogą jeszcze bardziej zwiększyć odpowiedzialność i zaufanie, pomagając zapobiegać niewłaściwemu wykorzystaniu i promować odpowiedzialne wdrażanie technologii sztucznej inteligencji.

Najlepsze praktyki dotyczące wzmacniania zabezpieczeń sztucznej inteligencji

Aby wyeliminować luki w zabezpieczeniach właściwe dla sztucznej inteligencji, organizacje muszą przyjąć proaktywne i strategiczne podejście. Oto 10 najlepszych praktyk w zakresie zabezpieczania systemów sztucznej inteligencji:

Warstwowa architektura zabezpieczeń:
Segmentacja, zapory sieciowe i silne uwierzytelnianie chronią infrastrukturę, oprogramowanie i dane w każdej warstwie.

Zabezpiecz łańcuch dostaw:
Wdrożenie solidnego programu zarządzania dostawcami. Kontroluj dostawców i komponenty innych firm, weryfikuj integralność oraz korzystaj z podpisanego kodu, aby zapobiegać podatnościom w cyklu rozwoju systemów AI.

Chroń dane szkoleniowe i modele:
Zabezpiecz się przed zatrutymi danymi, złośliwymi wejściami i innymi zagrożeniami, monitorując integralność danych i stosując solidne narzędzia walidacyjne.

Dokręcanie kontroli dostępu:
Egzekwuj zasady ograniczonego dostępu, wdrażaj kontrolę dostępu opartą na rolach (RBAC), regularnie zmieniaj poświadczenia i monitoruj uprawnienia, aby zapobiec nieautoryzowanemu dostępowi.

Bezpieczne interfejsy API:
Używaj silnych protokołów uwierzytelniania (takich jak OAuth 2.0), wymuszaj szyfrowanie HTTPS i regularnie aktualizuj interfejsy API, aby wyeliminować potencjalne luki w zabezpieczeniach.

Monitoruj i weryfikuj dane wyjściowe sztucznej inteligencji:
Korzystaj z wykrywania anomalii, rejestrowania i alertów, aby monitorować nietypowe wzorce lub szkodliwe zachowania w danych wyjściowych sztucznej inteligencji.

Plan odporności:
Regularne tworzenie kopii zapasowych danych i testowanie planów odtwarzania po awarii, aby zminimalizować przestoje i zapewnić szybkie odzyskiwanie danych w przypadku naruszenia.

Implementacja solidnego szyfrowania:
Szyfrowanie wrażliwych danych w spoczynku i w ruchu przy użyciu silnych algorytmów oraz bezpieczne zarządzanie kluczami szyfrowania i ich regularne zmienianie.

Przeprowadzanie regularnych audytów bezpieczeństwa i testów penetracyjnych:
Często oceniaj systemy pod kątem luk w zabezpieczeniach i przeprowadzaj testy penetracyjne w celu wykrycia zagrożeń, zanim zostaną wykorzystane.

Szkolenie personelu w zakresie najlepszych praktyk w obszarze bezpieczeństwa sztucznej inteligencji:
Regularnie organizuj szkolenia dla zespołu w zakresie bezpiecznego rozwoju, rozpoznawania zagrożeń i utrzymywania silnych praktyk bezpieczeństwa w celu zapobiegania naruszeniom.

DELL Technologies

Propozycja wartości firmy Dell: praktyczne rozwiązania w zakresie bezpieczeństwa AI.

Bezpieczeństwo sztucznej inteligencji może wydawać się skomplikowane, ale nie jest tak trudne, jak się wydaje. Prawda? Zabezpieczenie sztucznej inteligencji nie różni się tak bardzo od zabezpieczania istniejących obciążeń roboczych — chodzi o zrozumienie architektury i zastosowanie właściwych strategii. To właśnie rola firmy Dell Technologies.

Upraszczamy zagadnienia związane z bezpieczeństwem AI, wykorzystując istniejące rozwiązania i płynnie integrując je z architekturami skoncentrowanymi na sztucznej inteligencji.

Podejmujemy wyzwania takie jak szybkie wprowadzanie danych, nadużywanie interfejsu API i wrogie ataki, bez konieczności całkowitego przeglądu infrastruktury.

Wiedza ekspercka firmy Dell pozwala na obalanie mitów dotyczących bezpieczeństwa sztucznej inteligencji i pokazaniu, jak naprawdę jest to osiągalne. Niezależnie od tego, czy dopiero zaczynasz swoją przygodę ze sztuczną inteligencją, czy chcesz zwiększyć swoją ochronę, pomożemy Ci chronić inwestycje, zabezpieczyć systemy i zbudować odporną cyfrową przyszłość — bezpiecznie i skutecznie. Uprośćmy wspólnie bezpieczeństwo sztucznej inteligencji.

Produkty i rozwiązania firmy Dell, które mogą pomóc

Polecane rozwiązanie firmy Dell	Opis
Dell AI Factory	Dell AI Factory zabezpiecza obciążenia robocze związane ze sztuczną inteligencją za pośrednictwem bezpiecznego łańcucha dostaw, zapewniając zaufaną infrastrukturę od rozwoju po wdrożenie. Dzięki takim funkcjom jak niezmienność danych, izolacja i szyfrowanie, chroni wrażliwe modele i zestawy danych, chroni przed zagrożeniami cybernetycznymi oraz umożliwia skalowalne, wydajne i bezproblemowe operacje sztucznej inteligencji w dynamicznych środowiskach opartych na danych.
Cyberodporność	PowerProtect zabezpiecza obciążenia robocze związane ze sztuczną inteligencją dzięki zaawansowanym funkcjom, takim jak niezmienność i izolacja, zapewniając integralność danych i ochronę przed zagrożeniami cybernetycznymi. Oferuje kompleksowe szyfrowanie i wykrywanie anomalii, jednocześnie umożliwiając szybkie odzyskiwanie w celu zminimalizowania przestojów.
Dell Trusted Workspace (zabezpieczenia punktów końcowych)	Połączenie wbudowanych i opcjonalnych funkcji dodatkowych zaprojektowanych z myślą o zabezpieczeniu komercyjnych komputerów ze sztuczną inteligencją i uruchomionych na nich obciążeń roboczych związanych ze sztuczną inteligencją. Wykorzystanie bezpiecznych praktyk łańcucha dostaw i wbudowanych funkcji takich jak SafeBIOS i SafeID z modułem TPM. Opcjonalne dodatki obejmują Secured Component Verification, SafeID with ControlVault oraz oprogramowanie partnerskie CrowdStrike i Absolute, które maksymalizują bezpieczeństwo przestrzeni roboczej.
Usługi doradztwa w zakresie bezpieczeństwa AI	Zestaw usług, które mogą pomóc w opracowaniu i wdrożeniu kompleksowej strategii bezpieczeństwa AI. Oferta obejmuje usługi doradcze, AI VCISO i planowanie bezpieczeństwa danych.
Zarządzane operacje zabezpieczeń dla sztucznej inteligencji	Zapewnia szczegółowy wgląd w stos, aby szybko wykrywać zagrożenia i reagować na nie. Funkcje obejmują zarządzane wykrywanie i reagowanie, zarządzaną ochronę sztucznej inteligencji, testy penetracyjne sztucznej inteligencji oraz usługi reagowania na incydenty i odzyskiwania danych.
Integracja oprogramowania zabezpieczającego	Projektowanie, instalowanie i konfigurowanie narzędzi bezpieczeństwa, które chronią zarządzanie dostępem, aplikacje, sieci, chmury itd.

Więcej informacji o tym, jak radzić sobie z największymi wyzwaniami związanymi z cyberbezpieczeństwem, znajdziesz na stronie **dell.com/cybersecuritymonth**