RAPORT

Walka z cyberzagrożeniami przy użyciu sztucznej inteligencji

Wzmacnianie obrońców i rozbrajanie atakujących

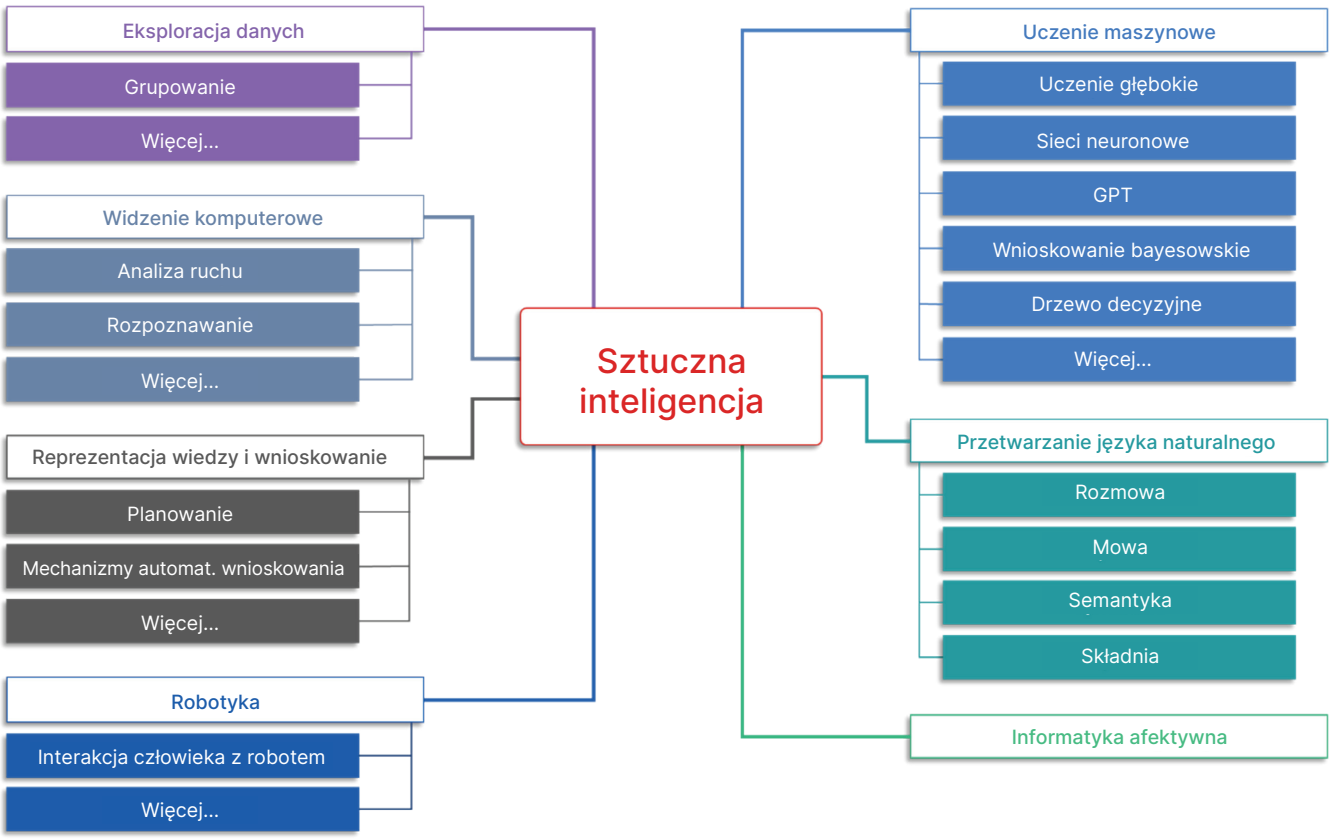
Streszczenie

Jeszcze parę lat temu z wielu dyskusji na temat sztucznej inteligencji (artificial intelligence, AI) wynikało prawdopodobnie więcej zamieszania niż pożytku. Obecnie jednak tempo innowacji w dziedzinie AI oraz jej pozytywny i negatywny wpływ na organizacje weszły w nową fazę. Z perspektywy cyberbezpieczeństwa pojawienie się istotnych i przełomowych narzędzi AI oraz ich wykorzystywanie przez cyberprzestępców zwiększa złożoność i pilność ochrony organizacji oraz ich infrastruktur cyfrowych przed nowymi cyberzagrożeniami opartymi na AI. Dobra wiadomość jest taka, że dostawcy rozwiązań cyberbezpieczeństwa od lat stosują różne technologie AI. Jeśli jednak chodzi o przyszłość pełną przestępców stosujących taktyki oparte na AI, kluczowe znaczenie dla liderów ds. bezpieczeństwa i IT oraz ich zespołów ma rozwijanie strategii bezpieczeństwa pod kątem przeciwdziałania tym nowym zaawansowanym zagrożeniom.

Niedawno przestępcy użyli wizerunków (wygenerowanych przy użyciu technologii deepfake) pewnego dyrektora finansowego i innych pracowników, aby podczas rozmowy wideo przekonać pracownika działu finansowego do zrobienia przelewu bankowego na kwotę 25,6 mln USD[1].

Innowacje wzmacniające obronę

W miarę jak organizacje kontynuują inicjatywy związane z cyfryzacją, nieuchronnie zwiększa się liczba wektorów ataków, na które są narażone. Bez względu na to, czy w inicjatywie chodzi o wdrożenie chmury, wyeliminowanie luk między technologią informatyczną (information technology, IT) a technologią operacyjną (operational technology, OT), zwiększenie liczby urządzeń z obszaru Internetu rzeczy (Internet-of-Things, IoT) połączonych z siecią, czy umożliwienie pracy hybrydowej, wiele organizacji coraz bardziej obciąża pracą zespoły ds. bezpieczeństwa i IT. Jakby tego było mało, teraz wykorzystywanie narzędzi AI przez przestępców dodatkowo pogarsza i tak już niełatwą sytuację.



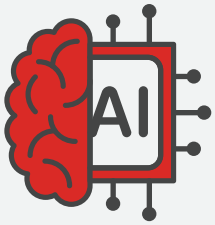
Ilustracja 1. Dziedziny i obszary zastosowania sztucznej inteligencji

Jak ci źli wykorzystują AI

Przestępcy wykorzystują zaawansowane możliwości AI w celu opracowywania oraz stosowania nowych, sprawniejszych i skuteczniejszych zagrożeń, w tym zagrożeń typu zero-day. Ze względu na użycie AI ataki mogą być przeprowadzane szybciej niż kiedykolwiek wcześniej. Gdy AI trafia w nieodpowiednie ręce, ma to różne skutki:

- Technologie AI, takie jak wstępnie przeszkolone transformatory generatywne (generative pretrained transformer, GPT), ułatwiają ataki, przez co mogą je przeprowadzać osoby, które wcześniej tego nie umiały. Obecnie osoba, która nie mówi po angielsku i przebywa w niemal dowolnym miejscu na świecie, może za pomocą technologii AI przygotować przekonujący atak phishingowy i socjotechniczny w języku angielskim (z użyciem poprawnej składni).
- AI może być wykorzystywana do pisania złośliwego kodu. Dzięki temu znacznie zmniejsza się nakład pracy związany z tworzeniem nowego złośliwego oprogramowania, a sam proces jego tworzenia jest dużo prostszy.
- Wykorzystywanie technologii deepfake przez przestępców już wzburzyło polityków i elektorat, a także umożliwiło popełnianie cyberprzestępstw na dużą skalę.
- AI może być wykorzystywana do szybszego wykrywania i wykorzystywania luk w zabezpieczeniach aplikacji, co zwiększa ryzyko w łańcuchach dostaw organizacji na całym świecie.

Obecnie złośliwe taktyki oparte na AI obejmują cały cykl życia ataku przedstawiony w bazie wiedzy MITRE ATT&CK. Zespół MITRE opracował bazę wiedzy o nazwie ATLAS (Adversarial Threat Landscape for Artificial-Intelligence Systems — Antagonistyczne zagrożenia dotyczące systemów sztucznej inteligencji), w której szczegółowo przedstawiono taktyki i techniki ataków opartych na AI[2].



Informatyk z Uniwersytetu Illinois w Urbanie i Champaign przetestował, jak można wykorzystać Open-AI Chat GPT-4 (wraz ze środowiskiem LangChain i przeglądarką z biblioteką Playwright) jako złośliwego agenta, który umożliwia wykrywanie luk w zabezpieczeniach witryn internetowych i ich atakowanie bez udziału człowieka. Efekt był zaskakujący: autorzy eksperymentu twierdzą, że narzędzie jest w stanie wykonać 38-krokową procedurę ataku wykorzystującego słowo kluczowe SQL Union[3].

Większe wyzwania

Wyzwania wynikające ze stale ewoluujących zagrożeń są coraz większe przez to, że przestępcy korzystają z AI, co wywiera dodatkową presję na i tak już mocno obciążone pracą zespoły ds. bezpieczeństwa i IT. Obecnie ochrona rozbudowywanego środowiska sieciowego i zwiększającej się liczby wektorów ataku przed tymi nowymi zagrożeniami jest bardziej skomplikowana niż kiedykolwiek wcześniej. Pojawiają się wyzwania związane z następującymi kwestiami:

- Wgląd w poszczególne części środowiska
- Brak możliwości scentralizowanego i skoordynowanego stosowania oraz egzekwowania polityk
- Korzystanie z wielu narzędzi i konsol bezpieczeństwa powoduje, że monitorowanie, klasyfikowanie alertów oraz badanie incydentów i reagowanie na incydenty jest niezwykle czasochłonne
- Nieustające trudności z zatrudnieniem ekspertów ds. bezpieczeństwa i zatrzymaniem ich w firmie

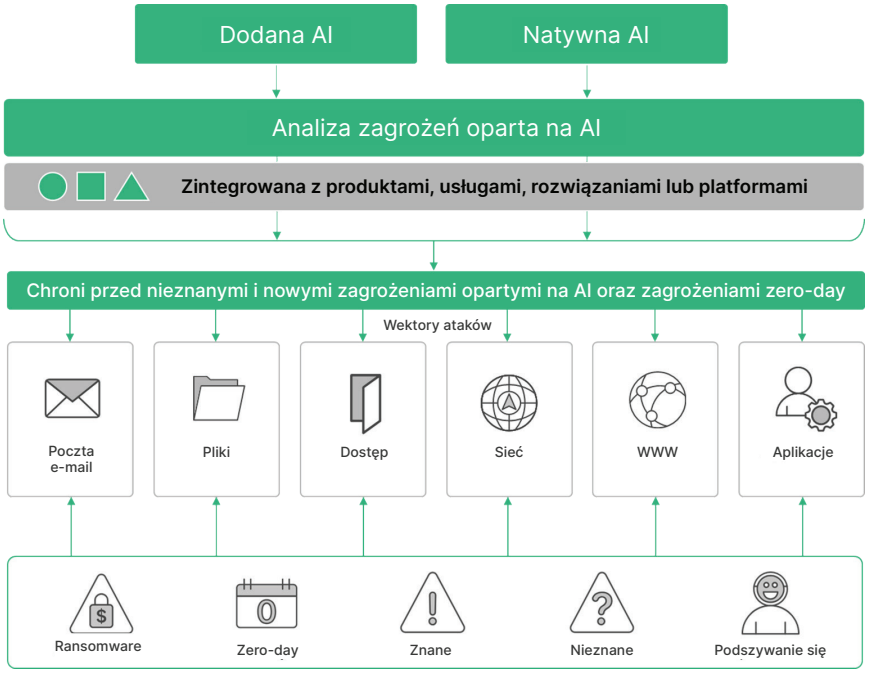
Aby organizacje mogły efektywnie radzić sobie z zagrożeniami wynikającymi z AI, będą musiały zmniejszyć złożoność i niezgodność oraz usprawnić działanie.

Jak ci dobrzy wykorzystują AI

Konwergencja w przypadku AI i cyberbezpieczeństwa nie polega tylko na ulepszeniu technologii. Wspólne jest też oczekiwanie na postęp (bardzo potrzebny i coraz pilniejszy), który musi się dokonać, aby organizacje mogły wzmocnić mechanizmy obronne pod kątem zupełnie nowych zagrożeń. Wielu dostawców rozwiązań cyberbezpieczeństwa od lat stosuje różne technologie AI. Na przykład Fortinet bada i wykorzystuje technologie AI od ponad 10 lat.

Analiza zagrożeń oparta na AI

Najważniejszą funkcją sztucznej inteligencji (AI) w obszarze cyberbezpieczeństwa jest udoskonalenie analizy zagrożeń. Technologie AI mają kluczowe znaczenie z perspektywy zbierania danych, analiz, korelacji i wreszcie przekształcania tych danych w przydatne informacje. Tego typu analizę zagrożeń opartą na AI można wykorzystywać w ramach integracji w celu kontrolowania szerokiej gamy wektorów ataków i różnorodnych zagrożeń — bez względu na to, czy atakujący korzystają z AI. Istotne jest, jak dostawca rozwiązania stosuje AI, jak szeroki jest zakres jego źródeł danych i jaką ilością danych dysponuje. Im lepszy wgląd w dane ma dostawca rozwiązania, tym więcej mogą się nauczyć modele AI.

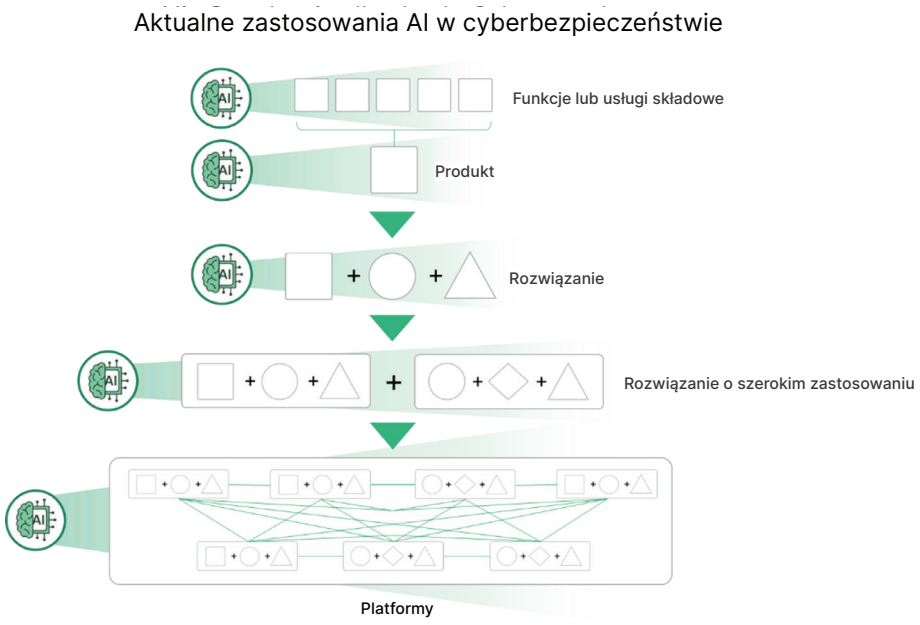


Ilustracja 2. Funkcje z dodaną AI i funkcje z natywną AI wykorzystywane do formułowania informacji o zagrożeniach

Poznanie właściwości analizy zagrożeń, na której bazuje podstawowa infrastruktura zabezpieczeń organizacji, to dobry punkt wyjścia do lepszego zapoznania się ze sposobami korzystania z technologii AI przez dostawcę rozwiązania. Jednym z kluczowych obszarów może być infrastruktura zapory, czyli główna linia obrony.

Dzisiaj zapory nowej generacji (next-generation firewall, NGFW) oferują mnóstwo możliwości wykraczających poza tradycyjne funkcje zapór. Na przykład zapora może mieć wbudowane funkcje zapobiegania włamaniom, funkcje ochrony przed złośliwym oprogramowaniem, w tym oprogramowanie antywirusowe i środowisko testowe (sandboxing), a także funkcje zabezpieczeń WWW, takie jak filtrowanie DNS i adresów URL. Porozmawiaj z dostawcą o tym, jak wykorzystuje AI w celu zwiększenia możliwości zapory, z uwzględnieniem szerokiej gamy jej funkcji i znaczenia każdej z nich.

Jeśli dostawca nie jest w stanie udzielić informacji o sposobie stosowania AI, lepiej wcześniej zareagować i poszukać dostawców, którzy niewątpliwie wykorzystują najnowsze technologie do zwiększania skuteczności swoich rozwiązań.



Ilustracja 3. Zastosowania AI i analiz zagrożeń (od komponentów po platformę)

Obecnie rozwiązania z dodaną AI mogą pomóc uzyskać lepsze efekty, co jest korzystne zarówno dla dostawców rozwiązań, jak i dla klientów:

- **Zapory:** zapory NGFW obejmują funkcje bezpieczeństwa, które często wykorzystują różne modele AI. Takie funkcje jak zapobieganie włamaniom, oprogramowanie antywirusowe, zabezpieczenia WWW i wbudowane środowisko testowe (sandboxing), mogą używać technologii AI do ulepszania poszczególnych funkcji zintegrowanych z zaporą. W przypadku połączenia hybrydowych zapór w topologii mesh (hybrid mesh firewall, HMF) z zaporami NGFW organizacje mogą czerpać podwójne korzyści z opartej na AI ochrony przed zagrożeniami, a także z ulepszeń w zakresie zapewnianego przez zapory wglądu oraz scentralizowanego zarządzania politykami i zaporami.
- **Skanowanie aplikacji:** mimo że przestępcy mogą wykorzystywać AI do tworzenia złośliwych agentów, rozwiązania do skanowania aplikacji i testy penetracyjne mogą używać tych samych funkcji do szybszego znajdowania i eliminowania luk w zabezpieczeniach — zarówno na etapie programistycznym, jak i w środowisku produkcyjnym.
- **Wykrywanie zagrożeń i reagowanie na urządzeniach końcowych (endpoint detection and response, EDR):** rozwiązanie EDR używa sieci neuronowych do rozpoznawania wzorców, aby sensownie interpretować dane zdarzeń rejestrowane na urządzeniu końcowym, w tym informacje o działaniach, procesach, zmianach w rejestrze i próbach dostępu do pamięci operacyjnej.
- **Zarządzanie informacjami i zdarzeniami dotyczącymi bezpieczeństwa (security information and event management, SIEM):** rozwiązanie SIEM wykorzystuje modele nadzorowanego i nienadzorowanego uczenia maszynowego (machine learning, ML) do przeprowadzania zaawansowanej regresji liniowej, w tym regresji wektorów nośnych, regresji procesów Gaussa i regresji drzew decyzyjnych. Ponadto używa ML do uruchamiania różnych algorytmów grupowania. Analiza ta pomaga SIEM precyzyjnie identyfikować zagrożenia i luki w zabezpieczeniach, a jednocześnie pozwala zminimalizować liczbę fałszywych alarmów. Rozwiązania SIEM wykorzystują też technologię GPT i przetwarzania języka naturalnego (natural language processing, NLP), aby pracownikom operacyjnego centrum bezpieczeństwa zapewniać wspomagane środowisko z większą ilością informacji. Analitycy mogą wysyłać zapytania bezpośrednio do silnika AI oraz uzyskiwać wgląd w informacje o zagrożeniach i wskazówki dotyczące odpowiedniego reagowania na incydenty.
- **Analiza obrazu:** widzenie komputerowe, rozpoznawanie obrazów i technologia sieci neuronowych są wykorzystywane razem na potrzeby analizy obrazu. Ponadto może być używany algorytm najbliższego sąsiada. Przychodzące obrazy osadzone w e-mailach lub pobierane z Internetu mogą być skanowane w celu ustalenia, czy wiąże się z nimi jakiekolwiek ryzyko lub zagrożenie. Do takich obrazów zaliczają się kody QR, zdjęcia i filmy pornograficzne, a także zdjęcia i filmy przedstawiające przemoc, przejawy ekstremizmu, broń, alkohol lub narkotyki.
- **Testy penetracyjne:** OpenAI Chat GPT4 można wykorzystywać do wykonywania zaawansowanych testów penetracyjnych, a na różnych filmach online pokazano, jak za pomocą dużego modelu językowego Chat GPT4 można w kilka minut napisać skrypty Python i Bash na potrzeby testów penetracyjnych.

Potencjał AI nie kończy się na formułowaniu i stosowaniu informacji o zagrożeniach przy użyciu AI. Na przykład w rozwiązaniach Fortinet technologie AI są wykorzystywane do ulepszania platformy Fortinet Security Fabric, tak aby była jeszcze bardziej proaktywna, zunifikowana i inteligentna.

Cyberbezpieczeństwo z natywną AI

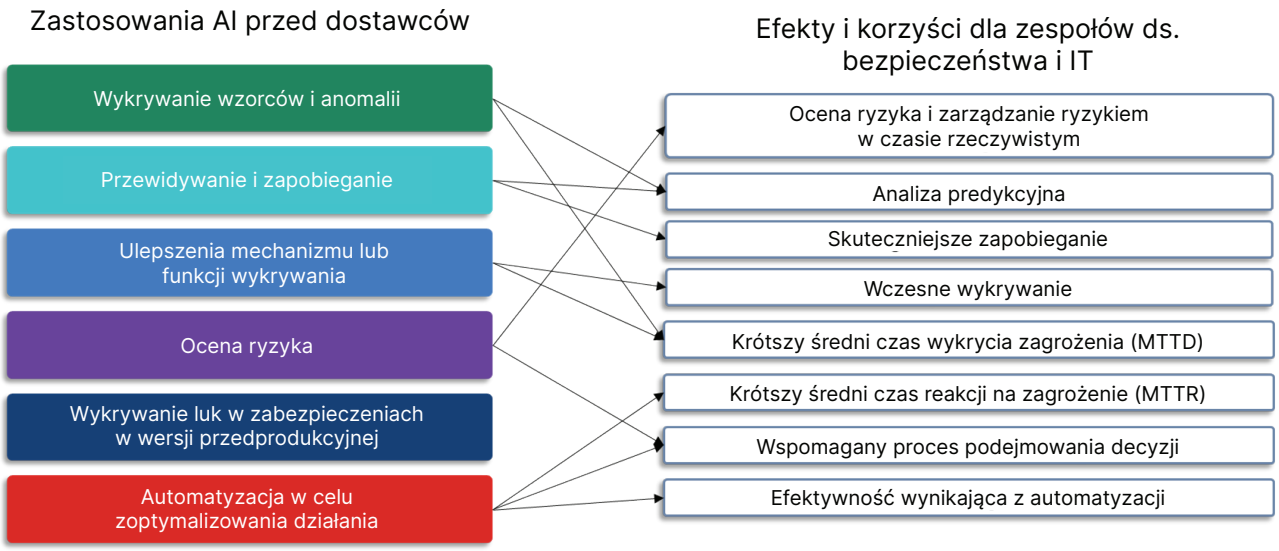
W związku z pojawieniem się rozwiązań cyberbezpieczeństwa, które już w punkcie początkowym wykorzystują możliwości AI, jest stosowany termin „cyberbezpieczeństwo z natywną AI". Mimo że nie ma branżowej definicji tego terminu, narzędzia cyberbezpieczeństwa z natywną AI działają z szybkością maszyny. Na przykład w przypadku analizy potencjalnych zagrożeń wydanie werdyktu i podjęcie zalecanych działań odbywa się z szybkością maszyny. Szybsze wykonywanie działań jest korzystne zarówno ze względów cyberbezpieczeństwa, jak i ze względów biznesowych. Zasadniczo narzędzia cyberbezpieczeństwa z natywną AI mają następujące cechy:

- Wykorzystanie modeli AI tworzonych w konkretnym celu
- Osadzenie AI jako rdzenia lub podstawy rozwiązania
- Ciągłe uczenie się i przystosowywanie do nowych zagrożeń
- Wykonywanie działań z szybkością maszyny

Zastosowania i zalecenia

Zespoły ds. bezpieczeństwa i IT spoza branży cyberbezpieczeństwa muszą wiedzieć o tym, że dostawcy rozwiązań cyberbezpieczeństwa stosują AI. W szczególności zespoły te potrzebują informacji o typie używanej AI, o sposobie jej używania i — co najważniejsze — o bezpośrednich korzyściach, jakie AI przynosi organizacji.

Ilustracja 4 przedstawia zastosowania technologii AI w rozwiązaniach cyberbezpieczeństwa oraz procesy pomocnicze i różne korzyści. Lista z ilustracji 4 może być przydatna podczas zadawania dostawcom pytań o to, jak AI poprawia stan bezpieczeństwa ich klientów (przede wszystkim w kontekście zagrożeń opartych na AI).



Ilustracja 4. Zastosowania AI przez dostawców i wynikające z nich korzyści dla klientów

Przy uwzględnianiu AI w strategiach bezpieczeństwa warto wziąć pod uwagę poniższe zalecenia.

Traktuj AI priorytetowo

Skieruj uwagę swojego zespołu na AI. Oceń aktualną wiedzę zespołu na temat technologii i reguł AI. Wyznacz wdrożenie AI jako cel strategiczny w głównych obszarach zabezpieczeń i infrastruktury IT oraz traktuj priorytetowo sprawdzanie tych obszarów. Dostawcy mogą już w pewnym zakresie używać AI w swoich rozwiązaniach, więc utwórz kwestionariusz i proces do oceniania dostawców na podstawie integracji AI oraz wiedzy specjalistycznej na temat AI.

Zdobywaj wiedzę

Liderzy ds. IT i bezpieczeństwa powinni zdobywać wiedzę na temat AI oraz możliwości jej wykorzystywania do usprawniania pracy i ulepszania kupowanych przez siebie rozwiązań. Ponadto powinni przekazywać te informacje swoim zespołom. Posiadanie tej wiedzy jest ważne, ponieważ pomaga ona lepiej zrozumieć różne kwestie, gdy organizacja eksperymentuje z używaniem technologii AI lub wdraża ją na własne potrzeby. Dzięki temu zespoły będą lepiej przygotowane do zadawania właściwych pytań dotyczących tych zastosowań. Istnieje wiele zasobów online, przy użyciu których można bezpłatnie uzyskać informacje o AI. Gdy liderzy i ich zespoły osiągną pewien poziom wiedzy, warto rozważyć uczestnictwo w płatnych szkoleniach online, a nawet w szkoleniach organizowanych przez renomowaną organizację zajmującą się cyberbezpieczeństwem, taką jak SANS.

Bądź na bieżąco

Na bieżąco śledź nowe rozwiązania AI dotyczące cyberbezpieczeństwa. Szybko pojawiają się kolejne innowacje, więc łatwo zostać w tyle.

Sprawdź infrastrukturę zabezpieczeń

Oceń, jak możesz używać technologii AI w swojej infrastrukturze zabezpieczeń. Najpierw zastanów się nad tym, jakie korzyści organizacja może osiągnąć dzięki zastosowaniu technologii AI w przypadku podstawowego elementu zabezpieczeń lub głównej linii obrony. Następnie zajmij się mechanizmami kontroli, którymi dysponujesz. Pod wieloma względami kolejność sprawdzania elementów infrastruktury może być zgodna z istniejącymi priorytetami zagrożeń określonymi dla całego środowiska (w którym priorytetowo sprawdzane są obszary największego ryzyka).

Zadawaj pytania

Pytaj dostawców rozwiązań cyberbezpieczeństwa o to, jak korzystają z AI. Dowiedz się, jak działają używane przez nich technologie i jak są stosowane. Przede wszystkim jednak poznaj korzyści, jakie dzięki AI odnoszą klienci. Poniżej znajduje się kilka pytań, od których możesz zacząć:

- Jaki wgląd w zagrożenia i powiązane źródła danych wykorzystuje Państwa firma do formułowania informacji o zagrożeniach, na których bazuje Państwa produkt, usługa lub rozwiązanie?
- Jak jest wykorzystywana sztuczna inteligencja do formułowania informacji o zagrożeniach?
- Jakie doświadczenie ma Państwa firma w wykorzystywaniu technologii sztucznej inteligencji w Państwa produktach, usługach i rozwiązaniach?

- Jakie konkretnie technologie sztucznej inteligencji zostały zastosowane w produkcie, usłudze lub rozwiązaniu i jak są one stosowane?
- Czy mogliby Państwo wskazać źródła danych używane w produkcie, usłudze lub rozwiązaniu do szkolenia technologii sztucznej inteligencji?
- Jak są szkolone i ponownie szkolone modele sztucznej inteligencji?
- Czy w jakimkolwiek zakresie jest możliwa interakcja ze sztuczną inteligencją?
- Jak zastosowana sztuczna inteligencja sprzyja osiąganiu poniższych celów?
  - Mniejsze ryzyko
  - Skuteczniejsze zapobieganie
  - Krótszy średni czas wykrycia zagrożenia (MTTD)
  - Mniej fałszywych alertów
  - Łatwiejsze klasyfikowanie alertów i badanie incydentów
  - Krótszy średni czas reakcji na zagrożenie (MTTR)
  - Ułatwienie rutynowych zadań analitykom w operacyjnym centrum bezpieczeństwa

Dodaj AI do kryteriów dotyczących dostawcy

Uwzględnij wykorzystywanie AI jako kryterium w zapytaniu ofertowym. Użyj powyższych i innych pytań, aby uzyskać informacje o wykorzystywaniu AI przez dostawców rozwiązań cyberbezpieczeństwa. Nie tylko pomoże to sprawdzić, jak dany dostawca używa technologii AI w imieniu klientów, ale także pozwoli porównać odpowiedzi różnych dostawców w celu ustalenia, czy konkretne zastosowanie AI może przynieść organizacji rzeczywiste korzyści.

Podsumowanie

Liderzy oraz specjaliści ds. bezpieczeństwa i IT muszą się zapoznać z różnymi technologiami AI. Ważne jest aktywne zdobywanie wiedzy o AI i wykorzystywanie jej potencjału w zakresie cyberbezpieczeństwa. Wiele rozwiązań bezpieczeństwa pochodzi od dostawców używających AI, więc warto poznać różne zastosowania, zanim narzędzia oparte na AI zaczną być szeroko stosowane w organizacji. Przed podjęciem decyzji o zakupie zadawaj dostawcom konkretne pytania o to, jak wykorzystują AI w swoich rozwiązaniach. Wiedza o tym, jak AI jest używana w tych rozwiązaniach do działań związanych z bezpieczeństwem i IT, może pomóc w walce z zaawansowanymi zagrożeniami opartymi na AI.

[1] Finance worker pays out $25 million after video call with deepfake 'chief financial officer' (Pracownik działu finansowego wypłacił 25 mln USD po rozmowie wideo z „dyrektorem finansowym" wygenerowanym za pomocą technologii deepfake), Heather Chen i Kathleen Magramo, CNN, 4 lutego 2024 r.

[2] MITRE ATLAS, Adversarial Threat Landscape for Artificial-Intelligence Systems.

[3] LLM Agents can Autonomously Hack Websites (Agenci LLM mogą anonimowo hakować witryny internetowe), Richard Fang i in., 6 lutego 2024 r.

FORTINET

www.fortinet.com

Copyright © 2024 Fortinet, Inc. Wszelkie prawa zastrzeżone. Fortinet®, FortiGate®, FortiCare® i FortiGuard® oraz określone inne znaki są zastrzeżonymi znakami towarowymi firmy Fortinet, Inc. Inne nazwy produktów Fortinet występujące w niniejszym dokumencie mogą być również zastrzeżonymi i/lub uznaniowymi znakami towarowymi firmy Fortinet. Pozostałe nazwy produktów i firm mogą być znakami towarowymi odpowiednich podmiotów. Wartości wydajności i inne wskaźniki zawarte w niniejszym dokumencie zostały uzyskane w wewnętrznych testach laboratoryjnych w idealnych warunkach. Rzeczywista wydajność i inne wyniki mogą być różne od podanych wartości. Na wyniki wydajności mogą mieć wpływ zmienne sieciowe, różne środowiska sieciowe i inne warunki. Żadne z przedstawionych tutaj treści nie stanowią wiążącego zobowiązania ze strony firmy Fortinet. Ponadto firma Fortinet wyłącza wszelkie gwarancje — wyraźne i dorozumiane — z wyjątkiem sytuacji, gdy firma Fortinet zawrze z nabywcą wiążącą umowę pisemną podpisaną przez radcę prawnego firmy Fortinet, która wyraźnie gwarantuje, że wskazany produkt będzie działać zgodnie z jasno określonymi wskaźnikami wydajności. W takim przypadku wiążące dla firmy Fortinet będą tylko konkretne wskaźniki wydajności wyraźnie ujęte w takiej wiążącej umowie pisemnej. W celu zapewnienia całkowitej jasności każda tego rodzaju gwarancja będzie ograniczona do wydajności uzyskanej w takich samych idealnych warunkach jak te panujące podczas wewnętrznych testów laboratoryjnych prowadzonych przez firmę Fortinet. Firma Fortinet wyłącza w całości wszelkie zobowiązania, oświadczenia i gwarancje wynikające z niniejszego dokumentu — zarówno wyraźne, jak i dorozumiane. Firma Fortinet zastrzega sobie prawo do zmiany, modyfikacji, przeniesienia lub wprowadzenia innych poprawek w niniejszej publikacji bez powiadomienia, przy czym zastosowanie ma najbardziej aktualna wersja tej publikacji.

16 sierpnia 2024 7:26

2604482-0-0-EN