DELL Technologies

NVIDIA

De 5 belangrijkste beveiligingsoverwegingen voor generatieve AI (GenAI)

Versnel de ingebruikname van een veilige en schaalbare infrastructuur met Dell AI Factory with NVIDIA

Het transformatieve potentieel van GenAI

GenAI heeft het potentieel om het speelveld te veranderen op manieren waarvan visionairs zich nog maar nauwelijks een voorstelling kunnen maken.

76%

van IT- en bedrijfsleiders is van mening dat GenAI hun organisatie transformatieve waarde zal opleveren[1]

AI

Geavanceerde analyse- en op logica gebaseerde technieken die gebeurtenissen interpreteren en beslissingen en acties ondersteunen en automatiseren.

GenAI

Technologieën en technieken die gebruikmaken van grote hoeveelheden data om nieuwe inhoud te genereren op basis van vragen in natuurlijke taal of andere niet-gecodeerde en niet-traditionele invoer.

Simulatie

- Digitale tweeling
- Synthetische data
- Frameworks ontwerpen
- Prognoses

Content creëren

- Codering
- Wiskunde
- Schrijven/spreken
- Afbeeldingen/video
- Audio

Content ontdekken

- Zoeken met natuurlijke taal
- Grote datasets analyseren
- Kennisbeheer
- Gepersonaliseerd onderwijs en gepersonaliseerde training

Gebruikerservaring

- Real-time vertalingen in meer dan 70 talen
- Gepersonaliseerde interacties met behulp van natuurlijke gezichtsuitdrukkingen en lichaamstaal

1 Dell Technologies Innovation Catalyst-onderzoek, februari 2024

Groter potentieel, groter risico

Het is verleidelijk voor bedrijfsleiders om snel te willen handelen en om implicaties rond data, compliance, governance en andere risico's te omzeilen. Maar GenAI heeft een keerzijde en dan hebben we het over beveiliging.

Voordelen

- Verbeterde bedreigingsdetectie
- Betere operationele efficiëntie
- Gepersonaliseerde beveiligingstrainingen

Nadelen

- Ingewikkeldere aanvallen
- Geavanceerde social-engineering
- Schaduw-AI

33%

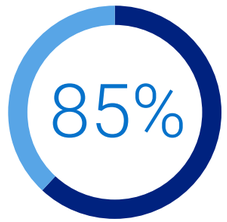
van de respondenten noemde cyberbeveiliging als het belangrijkste GenAI-risico dat hun organisaties proberen te beperken.[2]

2 McKinsey Global Survey over AI: De stand van zaken van AI in het vroege voorjaar van mei 2024

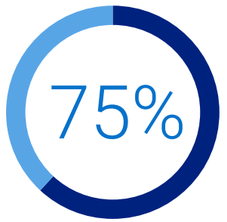
OVERWEGING 1

Het nieuwe dreigingslandschap

De belofte van GenAI gaat vergezeld van een ontnuchterende realiteit: aanvallers komen met nieuwe en ingewikkeldere aanvallen die conventionele verdedigingen kunnen omzeilen, waardoor cyberbeveiligingsteams de ontwikkelingen maar moeilijk kunnen bijbenen.



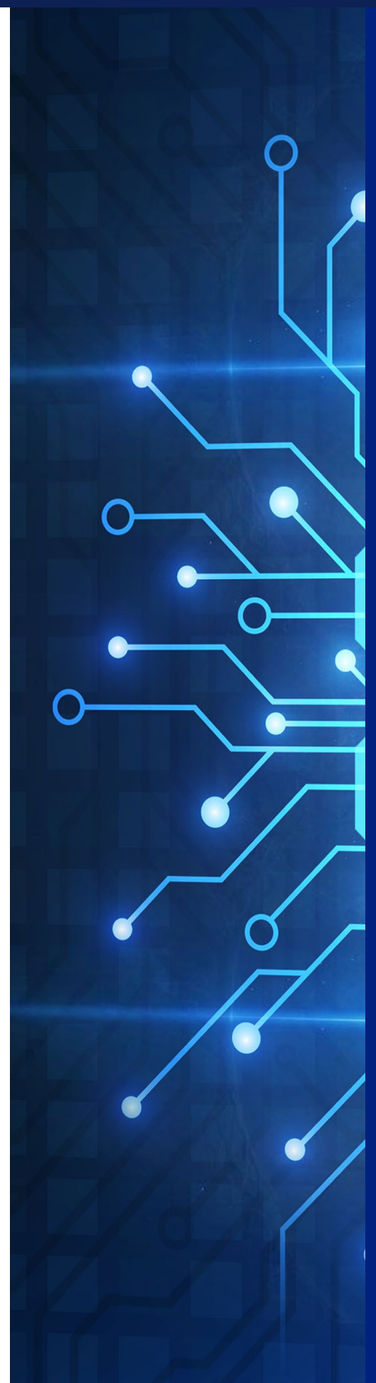
van de respondenten is van mening dat cyberaanvallen geavanceerder zijn geworden door AI.[3]



van beveiligingsprofessionals zag een toename in het aantal aanvallen de afgelopen 12 maanden.[4]

Om zich tegen deze opkomende bedreigingen te beschermen, moeten bedrijven zich richten op het minimaliseren van het aanvalsoppervlak door bijvoorbeeld penetratietests, monitoring en audits uit te voeren.

3 2024 Human Risk in Cybersecurity Survey, EY, mei 2024
4 Voice of SecOps Report "Generative AI and Cybersecurity: Bright Future or Business Battleground?" 2023



Opkomende aanvalsvectoren

Geavanceerde malware
Steeds geavanceerdere malware die gebruikmaakt van GenAI om zichzelf te 'ontwikkelen', waarbij de code voortdurend wordt gewijzigd om onopgemerkt te blijven door bestaande beveiligingsmiddelen, zoals detectie op basis van handtekeningen.

Sterk gepersonaliseerde phishing e-mails en campagnes
Steeds vaker authentiek ogende, schadelijke e-mails zonder de gebruikelijke tekens die wijzen op fraude.

Overtuigende deep fake-data
Identiteitsdiefstal, financiële fraude en misleidende informatie worden steeds gemakkelijker doordat menselijke acties kunnen worden nagebootst, zoals schrijven, spraak, afbeeldingen of video.

Geautomatiseerde verkenning
Het verzamelen van informatie waarmee kwetsbaarheden en zwakke plekken in het netwerk of systeem van een potentieel doelwit worden geïdentificeerd om meer gerichte aanvallen mogelijk te maken.

OVERWEGING 2

Implementatie en implementatierisico's

Organisaties die willen profiteren van de potentiële voordelen van GenAI, hebben grote hoeveelheden hoogwaardige data nodig - input waarmee modellen de beste resultaten kunnen produceren. Maar data en risico's zijn onlosmakelijk met elkaar verbonden. Voordat bedrijven informatie gebruiken, moeten ze hun unieke vereisten, input en risico's zorgvuldig evalueren en verantwoorden.

Kwetsbaarheden in het Large Language Model (LLM)

GenAI-services zijn kwetsbaar voor prompt-injectieaanvallen, waarbij aanvallers uitvoer manipuleren om beveiligingsmaatregelen te omzeilen of ongeoorloofd toegang te krijgen tot bestanden die mogelijk zijn gebruikt bij het verfijnen van het model.

Datavergiftiging

Aanvallers kunnen tijdens de trainingsfase opzettelijk gewijzigde data doorgeven aan een LLM. Dit kan ertoe leiden dat het model kwetsbaar is voor aanvallen via achterdeuren die in de data zijn ingesloten. Een voorbeeld uit de praktijk is het aanvallen en exploiteren van spamfilters door ze te trainen op spammail.

Complexiteit van regelgeving

Regelgevende autoriteiten over de hele wereld proberen in een hoog tempo de veiligheid van GenAI te begrijpen, controleren en garanderen. Hoewel GenAI-modellen onderworpen zijn aan de huidige regels voor de datasoevereiniteit die bepalen hoe data worden opgeslagen, verwerkt en gebruikt, zijn bestuursorganen nog steeds bezig om het toezicht op intellectueel eigendom en auteursrechtelijk beschermde informatie te definiëren. Het naleven van regelgeving kan kostbaar zijn, maar het niet naleven van gevestigde en opkomende regelgeving kan leiden tot boetes en andere sancties.

OVERWEGING 3

Schaduw-AI

Veel werknemers gebruiken tegenwoordig al openbare tekst-, beeld- en videogeneratoren zoals ChatGPT om hun dagelijkse workflows uit te breiden. Wanneer deze tools echter worden gebruikt zonder passend governance, vormen ze een kritieke bedreiging voor organisaties die hun bedrijfsintellectuele eigendom en data willen beveiligen. Dit ongeoorloofde gebruik van GenAI staat bekend als 'schaduw-IT'.

Verlies van intellectueel eigendom

Bedrijven hebben nu al te maken met het verlies van intellectueel eigendom doordat werknemers gevoelige informatie in openbare GenAI-tools invoeren.

Lekken van broncodedata

Ontwikkelaars die de broncode proberen te optimaliseren met behulp van ChatGPT veroorzaken datalekken.

Om de uitdagingen van schaduw-AI aan te pakken, moeten bedrijven een bedrijfsbrede raad instellen met de bevoegdheid om beslissingen te nemen met betrekking tot veilige AI-governance.

Waar bevinden uw data zich? Waar moeten workloads worden geplaatst?

AI werkt het beste wanneer het is gekoppeld aan uw data, waar deze zich ook bevinden. Met volledige controle over uw infrastructuur en LLM's, is er geen risico op verlies van intellectueel eigendom of lekken van broncodedata.

Kosten

Door gebruik te maken van on-premise implementaties kan de TCO in 3 jaar tijd met wel 75% worden verlaagd.[5]

Beveiliging en privacy

Zorg voor een veilige AI/GenAI-omgeving in de gehele organisatie met on-premise workflows en activiteiten. Oefen strikte controle uit over databeveiliging en naleving van de regelgeving, met name in sectoren die gevoelige data verwerken.

5 Gebaseerd op onderzoek van de Enterprise Strategy Group in opdracht van Dell, waarbij on-premise Dell infrastructuur wordt vergeleken met native public-cloudinfrastructuur as a service, april 2024. Geanalyseerde modellen laten zien dat een 7B-parameter LLM die RAG gebruikt voor een organisatie met 5000 gebruikers tot 38% rendabeler is, terwijl een 70B-parameter LLM die RAG gebruikt voor een organisatie met 50.000 gebruikers tot 75% rendabeler is. De werkelijke resultaten kunnen variëren. Economisch overzicht

OVERWEGING 4

Evaluatiecriteria

In het afgelopen jaar heeft de AI-gemeenschap zich steeds meer gericht op drie belangrijke kwesties: verantwoorde ontwikkeling en implementatie, beoordeling van de impact en beperking van de risico's. Wanneer bedrijven GenAI-modellen evalueren, moeten ze rekening houden met enkele belangrijke kanttekeningen:

Geen consistente rapportagevereisten

Toonaangevende ontwikkelaars testen hun modellen voornamelijk aan de hand van verschillende verantwoorde AI-benchmarks. Vanwege dit aanzienlijke gebrek aan standaardisatie in de rapportage is het moeilijk om de risico's en beperkingen van de beste AI-modellen methodisch te vergelijken.

Auteursrechtelijk beschermd materiaal in uitvoer

De uitvoer van populaire LLM's kan auteursrechtelijk beschermd materiaal bevatten, wat mogelijk in strijd is met de wet en waardoor bedrijven die het materiaal gebruiken het risico lopen op boetes.

Kwetsbaarheden worden steeds complexer

Onderzoekers ontdekken minder voor de hand liggende strategieën die ervoor zorgen dat LLM's schadelijk gedrag vertonen, zoals het vragen aan modellen om willekeurige woorden eindeloos te herhalen.

Ontwikkelaars zijn niet transparant

In veel gevallen zijn AI-ontwikkelaars niet openhartig over hun trainingsdata en methodologieën. Dit belemmert inspanningen om meer inzicht te krijgen in de robuustheid en veiligheid van AI-systemen.

OVERWEGING 5

Voordelen op het gebied van beveiliging

Naast de beveiligingsrisico's van GenAI zijn er ook potentiële beveiligingsvoordelen. GenAI wordt een belangrijke bondgenoot binnen cyberbeveiliging en opent nieuwe manieren van bescherming.

U kunt nu schaalbare beveiligingsactiviteiten ontwerpen die sneller toegang geven tot volwaardige inzichten en automatisch dreigingen detecteren, waardoor efficiëntie wordt geleverd en onderbezette beveiligingsteams worden ontlast.

Bedreigingsdetectie en reactie

Door historische data te analyseren en patronen en afwijkingen te identificeren, kan GenAI nieuwe en zich ontwikkelende bedreigingen in realtime herkennen. GenAI kan continu netwerkverkeer, systeemlogboeken en gebruikersgedrag bewaken en onmiddellijk onregelmatige activiteiten identificeren die op beveiligingsrisico's kunnen wijzen.

Het resultaat is een uiterst adaptieve bedreigingsdetectie, waardoor snel kan worden gereageerd op veranderende aanvalsvectoren en een proactief verdedigingsmechanisme wordt geboden tegen opkomende cyberbedreigingen.

Bedreigingssimulatie en -training

Met GenAI kunnen bedrijven een breed scala aan cyberbeveiligingsbedreigingen en aanvalsscenario's in een gecontroleerde omgeving simuleren. Hierdoor zijn teams beter voorbereid om cyberbedreigingen te identificeren, erop te reageren en deze in te dammen wanneer de tijd dringt.

Diepgaande analyses en samenvattingen

GenAI stelt teams in staat om data uit verschillende bronnen of modules te onderzoeken, waardoor ze traditioneel tijdrovende, vervelende data-analyses sneller en nauwkeuriger kunnen uitvoeren. Teams kunnen ook samenvattingen in natuurlijke taal laten maken van incidenten en dreigingsanalyses, waardoor de efficiëntie wordt verbeterd en de prestaties van het team worden verhoogd.

Gepersonaliseerde beveiligingstrainingen

Door conversationele AI toe te voegen aan GenAI en een AI-avatar in de gebruikersinterface op te nemen, kunnen organisaties gepersonaliseerde interacties leveren (24/7 bij elke schaalgrootte beschikbaar) met behulp van natuurlijke gezichtsuitdrukkingen en lichaamstaal. Deze aanpak kan worden ingezet voor beveiligingstrainingen en educatieve doeleinden, wat zorgt voor een meer natuurlijke, aangepaste en interactieve leerervaring, geautomatiseerde beoordelingen en meer.

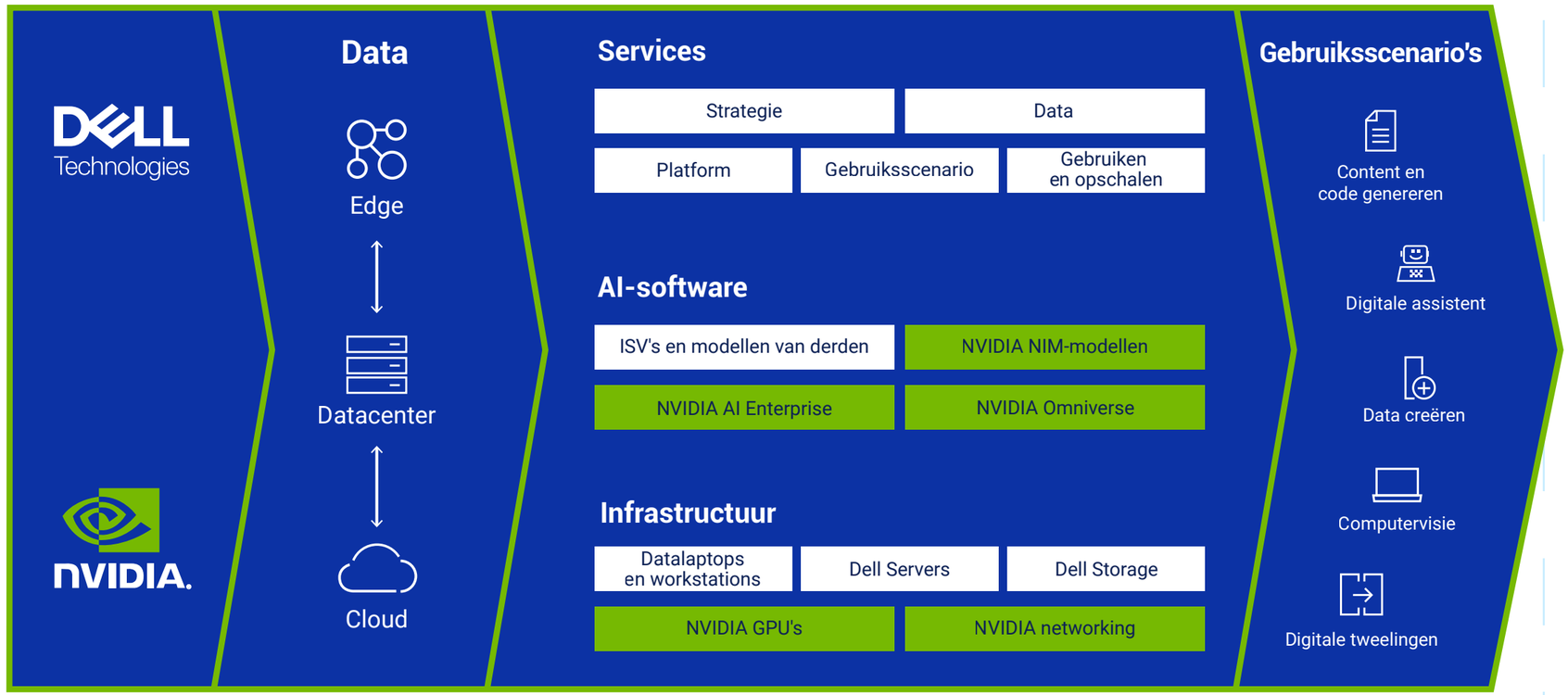
Dell AI Factory with NVIDIA

Versnel uw AI-traject en zet uw data veilig om in inzichten met de eerste uitgebreide, kant-en-klare AI-oplossing in de sector. Dell AI Factory with NVIDIA richt zich op de complexe behoeften van ondernemingen die gebruik willen maken van AI en GenAI. Met toonaangevende infrastructuur en services in combinatie met NVIDIA AI-software kunt u de time-to-value voor uw projecten versnellen door de ontwikkeling en implementatie te vereenvoudigen.

- Verminder het risico op compromissen met infrastructuur die is uitgerust met intrinsieke beveiliging, waaronder ***root of trust*** en andere belangrijke functies.
- Bescherm uw data tegen lekken die kunnen leiden tot verlies van intellectueel eigendom met een on-premise AI-oplossing die u zelf beheert.
- Voldoe aan strikte vereisten voor wat betreft compliance en datasoevereiniteit door AI met beveiligde toegang tot uw data in te voeren.
- Bescherm de privacy van uw belanghebbenden door te bepalen waar en wie toegang tot uw data hebben.

Dell AI Factory with NVIDIA

DE EERSTE END-TO-END AI-OPLOSSING VOOR BEDRIJVEN



Data is de motor achter de AI-factory en uw gebruiksscenario's

De meest waardevolle data bevinden zich on-premise en aan de edge. Met Dell Technologies kunt u AI toepassen op uw meest waardevolle data. Dell Technologies is toonaangevend op het gebied van het opslaan, beschermen en beheren van die data.

Van gebruiksscenario naar resultaten

De AI-factory levert bedrijfsresultaten die mogelijk worden gemaakt door uw belangrijkste gebruiksscenario's. Dell Technologies vereenvoudigt de implementatie van uw belangrijkste AI-gebruiksscenario's met gevalideerde oplossingen en services op maat.

Laat beveiligingsrisico's u niet stoppen om te blijven innoveren

Wij helpen u uw weg te vinden in de wereld van AI en GenAI, zodat u er de vruchten van kunt plukken.

STRATEGIEPLANNING

Gratis Accelerator Workshop voor GenAI

- Begin uw AI-traject met het ontwikkelen van een winnende strategie
- Pak uitdagingen en hiaten aan, stel prioriteiten aan doelstellingen en identificeer kansen
- Ontvang een gereedheidsbeoordeling met een uitvoerige beschrijving van de infrastructuurvereisten, AI-modellen, operationele integraties en meer

TECHNISCHE VOORBEREIDING

Een kant-en-klaar mobiel lab

Maak een vliegende start en begin uw reis naar succes vandaag nog. Inclusief een Dell Mobile Precision Workstation 5690/7780 met NVIDIA GPU's en twee dagen consultingservices om u op weg te helpen.

- Draagbare sandbox-omgeving voor het testen en demonstreren van GenAI
- Vooraf gevalideerd met NVIDIA AI Workbench-platform, klaar voor ontwikkelaars
- Eerste gebruiksscenario voor chatbots geïmplementeerd met uw data
- Kostenbesparende aanpak met weinig risico om te experimenteren en GenAI-vaardigheden te ontwikkelen

DELL MOBILE PRECISION WORKSTATION 5690/7780 MET NVIDIA GPU'S

GA VANDAAG NOG AAN DE SLAG

DELL Technologies AI Factory WITH NVIDIA