DELL Technologies

NVIDIA

De 5 viktigaste säkerhetsövervägandena för generativ AI (GenAI)

Påskynda implementeringen av en säker och skalbar infrastrukturbas med Dell AI Factory with NVIDIA

Den transformativa potentialen hos GenAI

GenAI har potential att förändra spelplanen på ett sätt som visionärer bara har börjat föreställa sig.

76 %

av IT- och företagsledare tror att GenAI kommer att leverera transformativt värde för deras organisation.[1]

AI

Avancerad analys och logikbaserad teknik för att tolka händelser, stödja och automatisera beslut och åtgärder.

GenAI

Tekniker som utnyttjar stora mängder data för att generera nytt innehåll från promptar med naturligt språk eller andra inmatningar som inte är kod och som inte är traditionella.

Simulering

- Digital tvilling
- Syntetiska data
- Utforma ramverk
- Förutsägelse

Skapa innehåll

- Kodning
- Matematik
- Text/tal
- Bild/video
- Ljud

Identifiering av innehåll

- Naturlig språksökning
- Analys av stora datauppsättningar
- Kunskapshantering
- Personligt anpassad utbildning och fortbildning

Användarupplevelse

- Realtidsöversättning på över 70 språk
- Personliga interaktioner med naturliga ansiktsuttryck och kroppsspråk

1 Dell Technologies Innovation Catalyst Study, februari 2024

© Dell Inc. Med ensamrätt.

Ökad potential, ökad risk

Det är frestande för företagsledare att vilja agera snabbt och kringgå konsekvenser som omfattar data, överensstämmelse, styrning och andra risker. Men GenAI är ett tveeggat svärd när det gäller säkerhet.

Fördelar

- Förbättrad hotdetektering
- Utökad operationell effektivitet
- Anpassad utbildning om säkerhetsmedvetenhet

Nackdelar

- Mer avancerade angrepp
- Avancerad socialteknik
- Skugg-AI

33 %

av de tillfrågade angav cybersäkerhet som den främsta GenAI-risken som deras organisationer arbetar för att minska.[2]

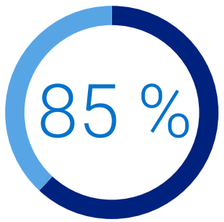
2 McKinsey Global Survey om AI: The state of AI in early, May 2024

© Dell Inc. Med ensamrätt.

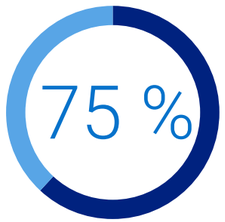
ÖVERVÄGANDE 1

Det nya hotlandskapet

Tillsammans med GenAI:s löfte följer en realistisk verklighet: Angripare skapar nya och mer invecklade attacker som kan kringgå konventionella försvar, och gör det svårt för cybersäkerhetsteam att hålla jämna steg.



av de tillfrågade anser att AI har gjort cybersäkerhetsattacker mer sofistikerade.[3]

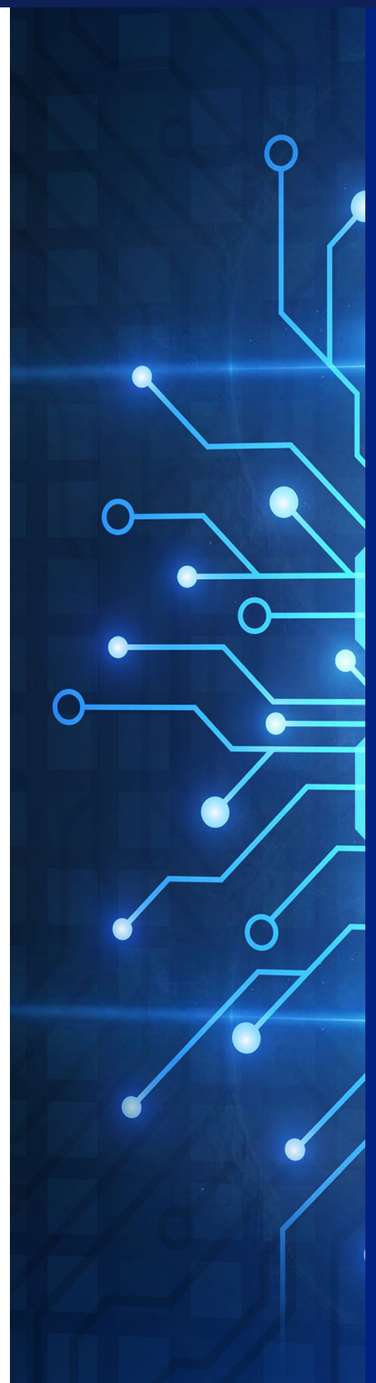


av säkerhetsproffs har sett en ökning av attacker under de senaste 12 månaderna.[4]

För att skydda mot dessa nya hot måste företag fokusera på att minimera attackytan genom exempelvis intrångstest, övervakning och granskning.

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

© Dell Inc. Med ensamrätt.



Nya attackvektorer

Avancerade skadliga program
Alltmer sofistikerade skadliga program som använder GenAI för "självutveckling" och ständigt ändrar sin kod för att inte upptäckas av befintlig säkerhet, som signaturbaserad detektering.

Mycket personliga nätfiskemejl och -kampanjer
Ökande frekvens av skadliga e-postmeddelanden som ser autentiska ut och saknar de vanliga tecknen på att det är ett bedrägeri.

Övertygande deepfake-data
Identitetsstöld, ekonomiska bedrägerier och felaktig information genomförs enklare tack vare förmågan att efterlikna mänskliga handlingar som att skriva, tala, bilder eller video.

Automatiserad spaning
Informationsinsamling som identifierar sårbarheter och svagheter i ett potentiellt måls nätverk eller system för att underlätta mer riktade attacker.

ÖVERVÄGANDE 2

Driftsättnings- och implementeringsrisker

Organisationer som vill dra nytta av de potentiella fördelarna med GenAI behöver stora mängder data av hög kvalitet – indata som modeller kan använda för att uppnå bästa möjliga resultat. Men data och risker går hand i hand. Innan företag drar nytta av någon information måste de noggrant utvärdera och ta hänsyn till sina unika krav, indata och risker.

Sårbarheter i stora språkmodeller (LLM)

GenAI-tjänster är sårbara för promptinmatningsattacker, där angripare manipulerar utdata för att kringgå säkerhetsskydd eller få obehörig åtkomst till filer som kan ha använts för att förfina modellen.

Dataförgiftning

Angripare kan avsiktligt mata en LLM med ändrade data under träningsfasen. Detta kan leda till att modellen är sårbar för attacker genom bakdörrar som byggts in i data. Ett exempel från verkligheten är att attackera och utnyttja skräppostfilter genom att träna dem på skräppostmeddelanden.

Regelkomplexitet

Tillsynsmyndigheter världen över skyndar sig att förstå, kontrollera och garantera säkerheten hos GenAI. Medan GenAI-modeller är föremål för aktuella datasuveränitetsregler som dikterar hur data lagras, bearbetas och används, definierar styrande organ fortfarande övervakning av IP och upphovsrättsskyddad information. Efterlevnad av regelverk kan vara kostsamt, men underlåtenhet att följa etablerade och nya regelverk kan leda till böter och andra påföljder.

© Dell Inc. Med ensamrätt.

ÖVERVÄGANDE 3

Skugg-AI

Många anställda använder redan offentliga text-, bild- och videogeneratorer som ChatGPT för att öka sina dagliga arbetsflöden. Men när dessa verktyg används utan rätt styrning utgör de ett kritiskt hot för organisationer som försöker skydda företagets immateriella egendom och data. Denna obehöriga användning av GenAI kallas skugg-AI.

Förlust av immateriell egendom

Företag hanterar redan nu förlust av immateriell egendom från anställda som delar känslig information i offentliga GenAI-verktyg.

Dataläckage av källkod

Utvecklare som försöker optimera källkod genom att använda ChatGPT har orsakat dataläckage.

För att hantera utmaningarna med skugg-AI bör företag implementera ett företagsomfattande råd eller en styrelse med behörighet att fatta beslut som rör säker AI-styrning.

© Dell Inc. Med ensamrätt.

Var finns dina data? Var ska arbetsbelastningar placeras?

AI fungerar bäst när den kombineras med dina data, var de än finns. Med fullständig kontroll över infrastruktur och stora språkmodeller finns det ingen risk för IP-förlust eller dataläckage från källkod.

Kostnader

Att dra nytta av implementeringar på plats kan sänka den totala ägandekostnaden med upp till 75 % under 3 år.[5]

Säkerhet och sekretess

Skapa säkra AI-/GenAI-miljöer i hela organisationen med arbetsflöden och drift på plats. Utöva strikt kontroll över datasäkerhet och regelöverensstämmelse följs, särskilt för branscher som hanterar känsliga data.

5 Baserat på Enterprise Strategy Groups ekonomiska sammanfattning på uppdrag av Dell, där Dells lokala infrastruktur jämfördes med inbyggd offentlig molninfrastruktur som en tjänst, april, 2024. Analyserade modeller visar att en LLM med 7B md-parametrar som använder RAG för en organisation med 5 000 användare är upp till 38 % mer kostnadseffektiv, medan en LLM med 70 md-parameter som använder RAG för en organisation med 50 000 användare är upp till 75 % mer kostnadseffektiv. De faktiska resultaten kan variera. Ekonomisk sammanfattning

ÖVERVÄGANDE 4

Utvärderingskriterier

Under det senaste året har AI-communityn alltmer fokuserat på tre viktiga frågor: ansvarsfull utveckling och distribution, bedömning av påverkan och riskreducering. När företag utvärderar GenAI-modeller måste de ta hänsyn till några viktiga förbehåll:

Inga konsekventa rapporteringskrav

Ledande utvecklare testar främst sina modeller mot olika ansvarsfulla AI-prestandatester. På grund av denna betydande brist på standardisering i rapporteringen är det svårt att metodiskt jämföra riskerna och begränsningarna hos de främsta AI-modellerna.

Upphovsrättsskyddat material i utdata

Utdata från populära LLM:er kan innehålla upphovsrättsskyddat material, vilket kan bryta mot lagen och utsätta företag som använder materialet för risk för påföljder.

Sårbarheter blir allt mer komplexa

Forskare hittar mindre uppenbara strategier som får LLM att visa skadligt beteende, till exempel att be modeller att i det oändliga upprepa slumpmässiga ord.

Utvecklarna saknar transparens

I många fall är AI-utvecklare inte tillmötesgående när det gäller deras träningsdata och metoder. Detta hindrar arbetet med att ytterligare förstå hur robusta och säkra AI-system är.

© Dell Inc. Med ensamrätt.

ÖVERVÄGANDE 5

Säkerhetsfördelar

Förutom säkerhetsrisker med GenAI finns det också potentiella säkerhetsfördelar. GenAI håller på att bli ett viktigt verktyg inom cybersäkerhet som öppnar upp för nya skyddsmöjligheter.

Du kan nu börja bygga upp skalbara säkerhetsåtgärder med snabbare åtkomst till större insikter och automatisk hotdetektering – vilket ger effektivitet och kompletterar underbemannade säkerhetsteam.

Hotdetektering och svar

Genom att analysera historiska data och identifiera mönster och avvikelser kan GenAI känna igen nya och föränderliga hot i realtid. Det kan kontinuerligt övervaka nätverkstrafik, systemloggar och användarbeteende för att snabbt identifiera oregelbundna aktiviteter som kan tyda på säkerhetshot.

Resultatet är starkt adaptiv hotdetektering, vilket möjliggör snabba svar på föränderliga attackvektorer och ger en proaktiv försvarsmekanism mot nya cyberhot.

Hotsimulering och träning

Med GenAI kan företag simulera ett brett utbud av cybersäkerhetshot och attackscenarier i en kontrollerad miljö. Det innebär att teamen är bättre förberedda på att identifiera, reagera på och mildra cyberhot när tiden är avgörande.

Djupgående analys och sammanfattning

GenAI möjliggör för team att undersöka data från olika källor eller moduler, så att de kan utföra traditionellt tidskrävande, tråkiga dataanalyser snabbare och med större noggrannhet. Teamen kan också skapa sammanfattningar på naturligt språk av incidenter och hotbedömningar, vilket förbättrar effektiviteten och ökar teamets resultat.

Anpassad utbildning om säkerhetsmedvetenhet

Genom att kombinera konversationsbaserad AI med GenAI och integrera en AI-avatar i användargränssnittet kan organisationer leverera anpassade interaktioner (tillgängliga dygnet runt och i stor skala) med naturliga ansiktsuttryck och kroppsspråk. Detta kan användas för säkerhetsutbildning och fortbildning, för en mer naturlig, anpassad och interaktiv inlärningsupplevelse, automatiserade bedömningar med mera.

Dell AI Factory with NVIDIA

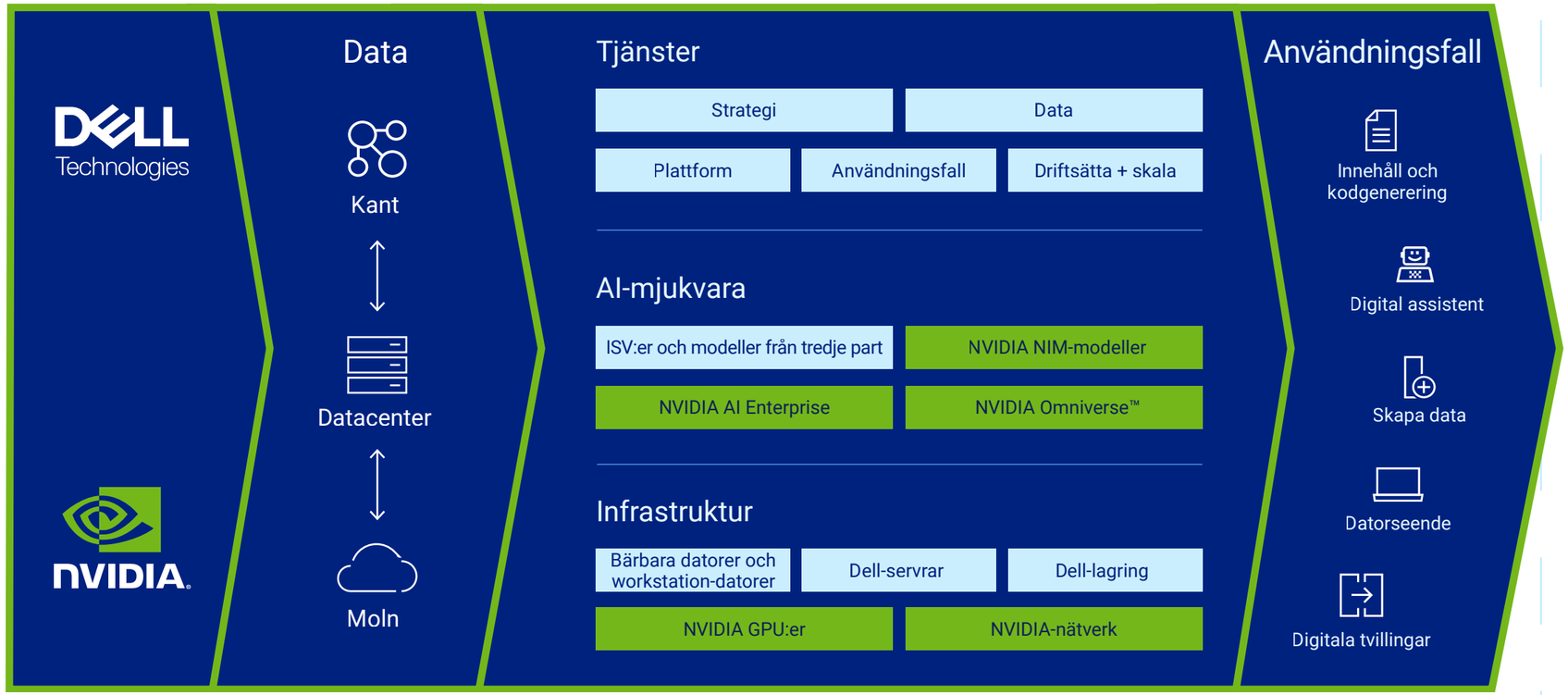
Rivstarta din AI-resa och omvandla dina data till insikter på ett säkert sätt med branschens första heltäckande och nyckelfärdiga AI-lösning. Dell AI Factory with NVIDIA tillgodoser de komplexa behoven hos företag som vill dra nytta av AI och GenAI. Med ledande infrastruktur och tjänster tillsammans med NVIDIA AI-mjukvara kan du öka tid till värde för dina projekt genom att förenkla både utveckling och distribution.

- Minska risken för kompromettering med infrastruktur som har inneboende säkerhet, inklusive förtroenderot och andra nyckelfunktioner.
- Skydda dina data från läckage som kan leda till förlust av immateriell egendom med en AI-lösning på plats som du kontrollerar.
- Uppfyll strikta krav på överensstämmelse och datasuveränitet genom att använda AI för dina data med säker åtkomst.
- Skydda dina intressenters integritet genom att kontrollera var och vem som har åtkomst till dina data.

© Dell Inc. Med ensamrätt.

Dell AI Factory with NVIDIA

BRANSCHENS FÖRSTA HELTÄCKANDE AI-LÖSNING FÖR FÖRETAG



Data ger kraft åt AI-fabriken och dina användningsfall

Dina mest värdefulla data finns på plats och i kanten.
Dell Technologies hjälper dig att använda AI på dessa värdefulla data och är ledande inom lagring, skydd och hantering av dem.

Användningsfall till resultat

AI-fabriken producerar affärsresultat som drivs av dina högst prioriterade användningsfall. Dell Technologies förenklar driftsättningen av de viktigaste användningsfallen för AI med validerade lösningar och skräddarsydda tjänster.

© Dell Inc. Med ensamrätt.

Låt inte säkerhetsrisker kväva innovationen

Låt oss hjälpa dig att navigera genom AI- och GenAI-världen så att du kan inhösta belöningarna.

STRATEGISK PLANERING

Kostnadsfri Accelerator Workshop för GenAI

- Inled din resa mot att utveckla en vinnande strategi
- Hantera utmaningar och luckor, prioritera mål och identifiera möjligheter
- Få en beredskapsbedömning för en närmare titt på infrastrukturkrav, AI-modeller, driftsintegreringar med mera

TEKNISK FÖRBEREDELSE

Mobilt labb som är färdigt att använda

Kickstarta din resa mot framgång. Inkluderar en Dell Precision mobil workstation 5690/7780 med NVIDIA GPU:er och två dagars konsulttjänster för att hjälpa dig att komma igång.

- Bärbar sandlådemiljö för GenAI-testning och demonstration
- Förvaliderad med NVIDIA AI Workbench-plattformen redo för utvecklare
- Inledande användningsfall för chattrobot implementerat med dina data
- Kostnadseffektiv metod med låg risk för att experimentera och bygga GenAI-kompetens

DELL MOBILE PRECISION WORKSTATION 5690/7780 MED NVIDIA GPU:ER

KOM IGÅNG IDAG

DELL Technologies AI Factory WITH NVIDIA

© Dell Inc. Med ensamrätt.