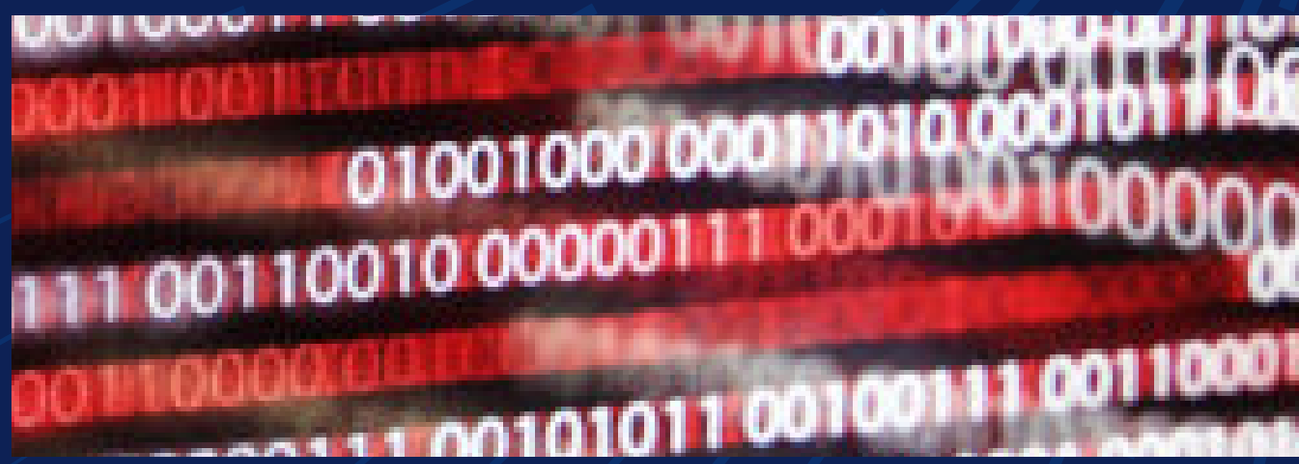
Krossa myterna om cybersäkerhet:

Avslöja AI-säkerhetsmyter



AI omvandlar många branscher, men när det gäller att skydda AI-lösningarna faller många organisationer offer för myter som gör att det här verkar mer komplicerat än vad det egentligen är. Hur förhåller det sig i verkligheten? Man behöver inte börja om från början för att skydda AI-system – att tillämpa befintliga cybersäkerhetsprinciper på AI:ns unika utmaningar räcker långt.

På Dell Technologies förstår vi arkitekturen bakom AI och kan hjälpa dig att anpassa dina nuvarande lösningar så att de passar det här nya ramverket. Låt oss gå igenom de vanligaste myterna kring AI-säkerhet och ta reda på sanningen som hjälper dig att skydda dina system ett effektivt sätt.

Myt 1: ”AI-system är för komplexa för att skyddas.”

I verkligheten: Det är sant att AI ger upphov till nya cybersäkerhetsrisker som promptinmatning, datamanipulering och avslöjande av känslig information, för att bara nämna några. Dessutom har agentbaserade AI-system även en bredare angreppsyta, eftersom de kan utnyttjas för att manipulera resultat eller eskalera behörigheter.

Även om det är viktigt att identifiera dessa sårbarheter och implementera säkerhetsåtgärder för att skydda AI-system från både traditionella och AI-specifika hot, kan riskerna hanteras och AI-modellerna kan skyddas. Det är viktigt att komma ihåg att AI-system kräver betydande mängder data som indata och skapar stora mängder data som utdata. Det gör att dataskyddet blir centralt, som en av de viktigaste säkerhetsstrategierna, tillsammans med följande:

- Nollförtroendeprinciper som identitetshantering, rollbaserad åtkomst och fortlöpande verifiering.
- Regelbundna intrångstester och sårbarhetshantering för att identifiera svagheter.
- Loggning och granskning för att validera in- och utdata

Myt 2: ”Inget av våra befintliga verktyg skyddar AI.”

I verkligheten: AI-säkerhet handlar inte om att börja om – det handlar om att arbeta smartare med de verktyg som redan finns. De flesta befintliga cybersäkerhetsverktyg kan anpassas för att skydda AI-system på ett effektivt sätt. AI är i grunden ännu en arbetsbelastning i din arsenal som driver verksamheten, även om den har unika egenskaper. Grundläggande cybersäkerhetspraxis, såsom identitetshantering, nätverkssegmentering och -övervakning, slutpunktsskydd och dataskydd, är fortfarande avgörande för att skydda AI-miljöer. Nyckeln är att anpassa dessa metoder för att hantera specifika AI-utmaningar, till exempel att skydda utbildningsdata, säkra algoritmer och minska risker som motstridiga maskininläsningsuppgifter.

Ett starkt försvar börjar med god cyberhygien – som systemkorrigeringar, åtkomstkontroll och hantering av sårbarheter. Det centrala är att anpassa dessa metoder för att hantera AI-specifika risker. Med AI-fokuserade strategier integrerade i den nuvarande säkerhetsstrategin och de rätta verktygen blir AI-säkerheten hanterbar och effektiv.

Det är dock viktigt att påpeka att uppdaterad hårdvara kan spela en avgörande roll i kampen mot cyberattacker. Moderna AI-datorer skapar till exempel en stark första försvarslinje mot en stor attackvektor, nämligen slutpunkter. När supporten för Windows 10 upphör blir föråldrade datorer en risk. Windows 11 kräver Trusted Platform Module (TPM) version 2.0, ett säkerhetschip som hjälper till med kryptering, säker start och skydd mot attacker med fast mjukvara. Många äldre datorer har antingen inte TPM alls eller bara stöd för en äldre version. Dell erbjuder säkra kommersiella AI-datorer där dessa förbättringar finns inbyggda.

Detsamma gäller för AI-infrastruktur som servrar och lagring. Dell AI Factory innehåller hårdvara som är optimerad för AI-säkerhet och innehåller ett antal inbyggda säkerhetsfunktioner som sträcker sig från en säker leverantörskedja till dataoföränderlighet, isolering och kryptering.

Myt 3: ”AI-säkerhet handlar bara om att skydda data.”

I verkligheten: AI-säkerhet sträcker sig bortom grundläggande dataskydd – det innebär att skydda hela AI-ekosystemet, inklusive modeller, API:er, utdata, system och enheter. I takt med att AI blir mer integrerat i kritiska program eskalerar de risker som är förknippade

med felaktig användning och utnyttjande. Utan robusta säkerhetsåtgärder kan AI-modeller manipuleras så att de genererar skadliga eller vilseledande utdata, API:er kan utnyttjas för att få obehörig åtkomst till känsliga system och utdata kan oavsiktligt exponera privat eller konfidentiell information.

Omfattande AI-säkerhet kräver ett flerskiktsperspektiv. Detta inbegriper att skydda modeller från angrepp som försöker manipulera indata för att lura AI-system, säkra API:er med starka autentiseringsmetoder för att förhindra obehörig användning och **kontinuerligt övervaka utdata** för att upptäcka ovanliga eller misstänkta mönster som kan tyda på en attack eller ett fel. Effektiv AI-säkerhet säkerställer inte bara integriteten och pålitligheten hos AI-system utan bygger även upp förtroendet hos användare och intressenter genom att minska riskerna för skadlig användning och oavsiktliga konsekvenser.

Myt 4: ”AI behöver inte mänsklig bevakning.”

I verkligheten: Styrning och mänsklig tillsyn är avgörande för att säkerställa att AI-system fungerar etiskt, förutsägbart och i linje med mänskliga värderingar. Avancerade AI-system, särskilt agentbaserad AI med autonoma beslutsfunktioner, medför unika utmaningar som kräver robusta skyddsåtgärder. Utan korrekt tillsyn kan dessa system avvika från avsedda mål eller uppvisa oavsiktliga beteenden som kan utgöra risker.

För att hantera detta är det viktigt att ställa upp tydliga gränser, implementera kontrollmekanismer i lager och säkerställa en kontinuerlig mänsklig inblandning i avgörande beslutsprocesser. Regelbundna revisioner, transparens i AI-driften och noggranna tester kan ytterligare förbättra ansvarsskyldigheten och förtroendet, vilket bidrar till att förhindra missbruk och främja en ansvarsfull distribution av AI-teknik.

Bästa praxis för att stärka AI-säkerheten

För att täppa till AI-specifika säkerhetsluckor måste organisationen anta en proaktiv och strategisk strategi. Här är 10 beprövade metoder som tillämpas för att skydda AI-system:

Skiktad säkerhetsarkitektur:
Använd segmentering, brandväggar och stark autentisering för att skydda infrastruktur, mjukvara och data i varje lager.

Säkra leverantörskedjan:
Implementera ett effektivt leverantörshanteringsprogram. Granska leverantörer och komponenter från tredje part, validera integriteten och förlita dig på signerad kod för att förhindra sårbarheter i AI-utvecklingslivscykeln.

Skydda utbildningsdata och -modeller:
Skydda dig mot förgiftade data, motstridiga maskininlärningsuppgifter och andra hot genom att övervaka dataintegriteten och tillämpa robusta valideringsverktyg.

Förstärk åtkomstkontrollerna:
Tillämpa principer om minsta behörighet, implementera rollbaserad åtkomstkontroll (RBAC), ändra inloggningsuppgifter regelbundet och granska behörigheter för att förhindra obehörig åtkomst.

Säkra API:er:
Använd starka autentiseringsprotokoll (som OAuth 2.0), tillämpa HTTPS-kryptering och uppdatera regelbundet API:erna för att åtgärda potentiella sårbarheter.

Övervaka och validera AI-utdata:
Använd detektering, loggning och varningar för att övervaka ovanliga mönster och skadligt beteende i AI-utdata.

Planera för motståndskraft:
Säkerhetskopiera data regelbundet och testa katastrofåterställningsplaner för att minimera driftavbrotten och säkerställa snabb återställning vid intrång.

Implementera robust kryptering:
Kryptera känsliga data i vila och under överföring med hjälp av starka algoritmer och hantera och rotera krypteringsnycklar regelbundet och på ett säkert sätt.

Genomför regelbundna säkerhetsrevisioner och intrångstester:
Utvärderar ofta systemen för att hitta sårbarheter och använder intrångstestning för att upptäcka risker innan de kan utnyttjas.

Utbilda personalen i beprövade rutiner för AI-säkerhet:
Fortbilda ditt team regelbundet i säker utveckling, hotigenkänning och kraftfulla säkerhetsmetoder för att förhindra intrång.

DELL Technologies

Dells värdeerbjudande: praktiska AI-säkerhetslösningar.

AI-säkerhet kan verka komplicerat, men så är det inte. Hur förhåller det sig i verkligheten? Att skydda AI skiljer sig inte så mycket från att säkra befintliga arbetsbelastningar – det handlar om att förstå arkitekturen och tillämpa rätt strategier. Det är där Dell Technologies kommer in i bilden.

Vi avmystifierar AI-säkerhet genom att använda dina nuvarande lösningar och smidigt integrera dem i AI-inriktade arkitekturer. Vi hanterar utmaningar som promptinmatning, API-missbruk och angrepp utan att det krävs en fullständig infrastrukturöversyn.

Dells expertis handlar om att avslöja myterna kring AI-säkerhet och visa att det faktiskt går att uppnå. Oavsett om du precis har påbörjat din AI-resa eller vill förbättra det befintliga försvaret hjälper vi dig att skydda dina investeringar, säkra systemen och bygga en motståndskraftig digital framtid – tryggt och effektivt. Låt oss förenkla AI-säkerheten tillsammans.

Dell-produkter och -lösningar som kan hjälpa dig

Dells utvalda lösning	Beskrivning
Dell AI Factory	Dell AI Factory skyddar AI-arbetsbelastningar genom en säker leverantörskedja, vilket säkerställer en betrodd infrastruktur från utveckling till distribution. Med funktioner som dataoföränderlighet, isolering och kryptering skyddar den känsliga modeller och datauppsättningar, skyddar mot cyberhot och möjliggör skalbar, effektiv och sömlös AI-drift i dynamiska, datadrivna miljöer.
Cyberelasticitet	PowerProtect säkrar AI-arbetsbelastningar med avancerade funktioner som oföränderlighet och isolering, vilket säkerställer dataintegritet och skydd mot cyberhot. Den erbjuder heltäckande kryptering och avvikelsedetektering, samtidigt som snabb återställning minimerar längden på driftavbrott.
Dell Trusted Workspace (slutpunktssäkerhet)	En kombination av inbyggda och valfria tilläggsfunktioner som utformats för att skydda kommersiella AI-datorer och AI-arbetsbelastningar som körs på dem. Inbyggda funktioner är byggda med metoder för säker leverantörskedja och inkluderar SafeBIOS och SafeID med TPM. Valfria tillägg inkluderar säker komponentverifiering, SafeID med ControlVault och partnermjukvaran CrowdStrike och Absolute för att maximera säkerheten över arbetsytan.
AI-baserade säkerhetsrådgivningstjänster	En svit av tjänster som kan hjälpa dig att utveckla och implementera en omfattande AI-säkerhetsstrategi. Erbjudandena omfattar rådgivningstjänster, AI vCISO och datasäkerhetsplanering.
Hanterad säkerhetsdrift för AI	Skapar ett synlighetsdjup över stacken för att göra det möjligt att snabbt upptäcka och reagera på hot. Funktionerna omfattar Managed Detection and Response, Managed AI Guard, intrångstester för AI samt incidentsvar och återställningstjänster.
Integrering av säkerhetsmjukvara	Utforma, installera och konfigurera säkerhetsverktyg som skyddar åtkomsthantering, program, nätverk, moln med mera.

Lär dig att hantera några av dagens främsta cybersäkerhetsutmaningar på
dell.com/cybersecuritymonth

DELL Technologies