NATIONELLT CYBERSÄKERHETSCENTER
En del av FRA

Rekommendationer

Juni 2026

Skydda dig mot AI-förstärkta angrepp

AI förändrar inte spelplanen – den höjer tempot.
Rekommendationer för beslutsfattare och organisationer.

Utvecklingen av AI

Utvecklingen av AI sker i rask takt. Nya modeller släpps med bara några månaders mellanrum och för varje ny generation förbättras dess förmågor. Dagens avancerade AI modeller har visat sig vara mycket kapabla inom cyberdomänen vilket kan leda till fler angrepp. AI förändrar sällan vilka typer av angrepp som förekommer, dock förändras takt och skala. Nationellt cybersäkerhetscenter (NCSC) bedömer att AI kommer öka både volymen och effekten av cyberattacker framöver, främst genom att förstärka befintliga metoder snarare än att skapa helt nya.

Ökar förmågan för både försvarare och angripare

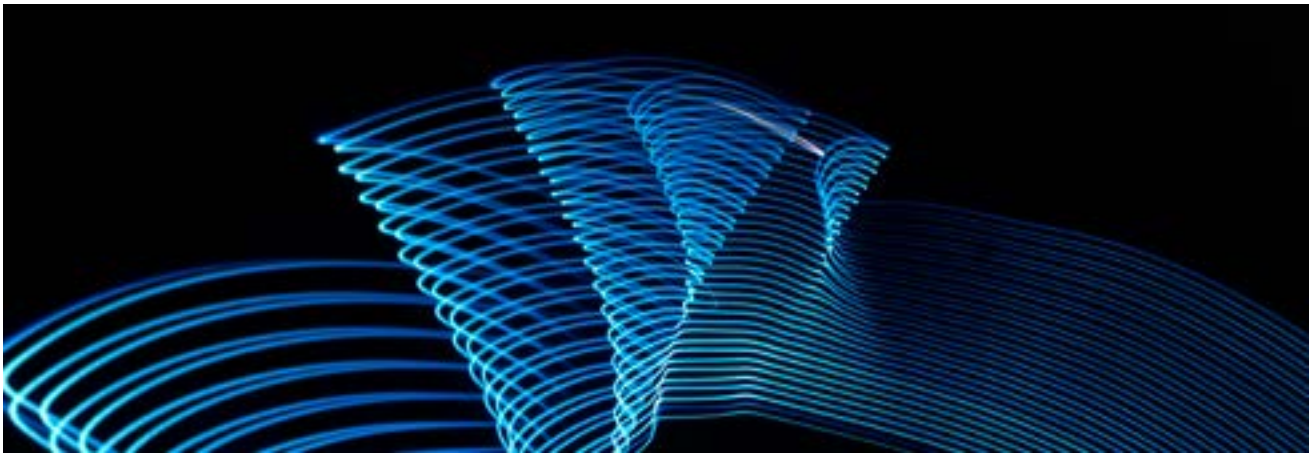
Genomförande av cyberangrepp, till följd av utvecklingen av AI, blir mindre en fråga om kompetens och mer en fråga om verktyg. Tröskeln sänks för mindre kunniga angripare och effektiviteten ökar för de med mer kompetens. AI kan exempelvis användas för att snabbare hitta och utnyttja sårbarheter i system, förbättra social engineering-tekniker såsom phishing och deep fakes samt att snabba upp dataanalys, både under och efter ett angrepp. AI kan även användas för att stärka skyddet mot cyberattacker. Det går att nyttja tekniken både i förebyggande syfte och för att snabbare hantera en attack när den väl inträffar.

Ytterligare en attackvektor

När organisationer själva använder AI uppstår en ny cyberattackyta eftersom systemen är kontextberoende och tolkar snarare än strikt exekverar input. Det gör dem sårbara för angrepp som prompt injection, datamanipulation och informationsläckage, särskilt när de kopplas till interna system och får agera autonomt.

Prioritera grundläggande cyberhygien

Grundläggande cybersäkerhetsprinciper gäller fortfarande, men de måste fungera bättre än någonsin. Dessutom måste det finnas flexibilitet och utrymme för att utveckla nya förmågor. Grundprinciperna för cybersäkerhet gäller fortfarande och bör prioriteras för de risker som inte kan hanteras på annat sätt. Segmentering av nätverk, patchning av kända sårbarheter, identitets- och åtkomsthantering samt försvar på djupet och bredden ökar svårigheten för angriparen. Att utöka dessa insatser medan tid finns är en klok investering.



AI-drivna cyberhot är ett område med snabb utveckling. Dessa rekommendationer kan förändras över tid.

Administrativa rekommendationer

- **Risk ägs av ledningen.** Gamla och ej uppdaterade system, kod som inte underhålls och oklart ägarskap över system medför affärsrisker med direkta konsekvenser för verksamhetens kontinuitet och anseende.
- **Uppdatera hot- och riskanalysen.** Säkerställ heltäckande analys av både AI-förstärkta angrepp utifrån och hot från interna AI-system, om sådana används. Fastställ en tydlig riskaptit och tidsätt de beslut som krävs för att uppnå den.
- **Säkerställ resurser och reservkapacitet.** Analysera behovet av utökad bemanning, budget och extern kompetens så att kritiska resurser inte uttöms. Det är viktigt för att leveranser ska kunna upprätthållas och personal inte överbelastas. Verifiera att leverantörer har proportionerliga resurser.
- **Samverka med sektorsgrupper, sektorsforum och Sveriges nationella enhet för hantering av it-incidenter (nationell CSIRT), CERT-SE.** Försvarare bör använda samverkans forum minst lika aktivt som angripare använder sina nätverk. Engagemang i kunskapshöjande nätverk och forum kan stärka tillgången till kompetens, forskning, erfarenhetsutbyte och relevant hotinformation.
- **Kravställ säkerhet vid upphandling.** Ställ säkerhetskrav i avtalen, säkra kontinuitets- och utträdesplaner och följ upp efterlevnaden. Det som inte kravställs är svårt att kräva i efterhand.
- **Besluta vilka AI-verktyg som är tillåtna och vad de får se.** Avsaknad av beslut och policy ökar risken för skugg-IT. Policy bör inkludera vilken som får delas med AI systemen och förbjud mot känsliga data i öppna AI-tjänster.
- **Kräv mänskligt godkännande (human in the loop).** Innan AI tillåts påverka verksamhetens processer bör en riskanalys genomföras. Analysen ska avgöra om åtgärden behöver granskas och godkännas av en människa.
- **Öva på AI-förstärkta attackscenarier.** Genom övning kan organisationen öka förmågan att hantera AI-assisterade angrepp. Övningar kan anpassas för att testa delar av organisationen eller mer heltäckande.

Tekniska rekommendationer

- **Grundläggande cyberhygien.** Avancerade angrepp bygger oftast vidare på redan kända metoder, AI gör dem snabbare och billigare, men sällan är de helt nya. En organisation med ordning på grunderna gällande cybersäkerhet står betydligt starkare, oavsett hur sofistikerad angriparen är. Grundläggande cyberhygien är därför inte bara en miniminivå att bocka av, utan själva förutsättningen för att de mer avancerade och AI-specifika åtgärderna ska ge effekt. Säkerställ en god cyberhygien genom att följa NCSC:s 10[1] rekommenderade säkerhetsåtgärder, och se till att de implementeras korrekt och upprätthålls över tid.

- **Uppdatera mjukvara skyndsamt** och så automatiserat som möjligt. Tiden från att en sårbarhet blir känd till att den utnyttjas krymper, och AI förstärker den utvecklingen genom att göra det snabbare och billigare att hitta och utnyttja kända brister. Prioritera internetexponerade och affärskritiska system, där sårbarheter får störst konsekvens.

  Eftersom det inte längre är realistiskt att vänta på servicefönster, använd tekniker som exempelvis hot patching[2], blue/green deployment[3] och canary release[4] för att uppdatera utan driftavbrott. Säkerställ att processer och riskaptit i högre grad tillåter frekvent och storskalig uppdatering och fasa ut produkter som inte längre får säkerhetsuppdateringar.

- **Minimera den internetexponerade attackytan** genom att snabbt identifiera och minska externt exponerade ytor. Det blir ännu viktigare när en angripare snabbare kan kartlägga system och hitta svagheter i dem med hjälp av AI. Att ha full kontroll över vad som är exponerat, hålla exponeringen på ett minimum och stänga allt som inte behöver vara åtkomligt utifrån är inte en engångsinsats utan ett kontinuerligt arbete.

- **Använd AI-assisterad kodgranskning**, både för egenutvecklad kod och för kod som utvecklas av tredje part. AI kan granska stora mängder kod snabbt och fånga sårbarheter, osäkra mönster och kända brister tidigt, då de är som billigast och enklast att åtgärda.

  Bygg in granskningen som ett naturligt steg i utvecklingsflödet, så att kod kontrolleras löpande i takt med att den skrivs. Låt scanningen även omfatta de externa bibliotek och beroenden som dras in. För mjukvara som levereras av tredje part bör krav på AI-assisterad kodgranskning skrivas in i avtalet och följas upp.

- **Skydda hela AI-livscykeln.** Säkra AI genom att bygga in säkerhet i varje del av AI-systemet från början. Det innebär att skydda de system, data och arbetsflöden som stödjer designen, utvecklingen, driftsättningen, driften och avvecklingen av AI.

- **Begränsa agentiska system.** Agentiska system kräver särskild försiktighet, se till att de inte når eller kan utföra mer än den faktiskt behöver och kräv mänsklig bekräftelse innan känsliga eller oåterkalleliga åtgärder utförs. Etablera spårbarhet för agentens beslut och en möjlighet att snabbt stoppa systemet om det beter sig oväntat. Ett agentsystem med bred åtkomst och svag tillsyn är ett tydligt högriskscenario. Utgå från värsta tänkbara scenario och se till att det finns ett tydligt ägarskap för säkerheten genom hela livscykeln.

- **Inför AI i cyberskyddet,** bevara försvararens fördel. Angripare använder redan AI för rekognosering, exploateringsutveckling och förflyttning i nätverk. Lösningen är inte så enkel som ett enskilt verktyg som går att köpa in, det handlar om att stegvis bygga AI-stödda förmågor i försvaret. Den som avstår riskerar att hamna strukturellt på efterkälken. Samtidigt har försvararen en hemmaplansfördel som angriparen saknar: privilegierad insyn i de egna näten, systemen och loggarna. Den fördelen består dock bara om investeringarna görs tidigt. Den som väntar låter angriparen sätta tempot.

- **Använd AI där maskinhastighet gör störst skillnad**. Prioritering och berikning av larm, identifiering av sårbarheter, logg- och anomalianalys samt automatiserad inneslutning av pågående hot. Ingripande åtgärder ska ske med mänsklig översyn och med möjlighet till säker återställning. AI:n får agera snabbt, men människan behåller kontrollen. Grunderna först, AI för att accelerera, människan i kontroll.

Att blicka framåt

Den nuvarande framdriften av gränsöverskridande AI förväntas fortsätta avancera. Bara under det gånga året har mycket hänt inom utvecklingen av AI och dess förmåga att påverka både cybersäkerhetshot och -skydd. Fortsätter utvecklingen i samma eller ökad takt som tidigare kommer det troligtvis innebära fortsätta förändringar kopplade till cybersäkerhetsrisker. Det innebär att framtiden är oviss och att organisationer måste ha utrymme att hantera osäkerhet. För att ha så bra förutsättningar som möjligt att snabbt kunna reagera på kommande tekniskt utveckling hjälper det att ha en god självkännedom om den egna IT miljön. Målet för allt cybersäkerhetsarbete bör inte enbart vara att återgå till en balans mellan angripare och försvarare, utan också att bygga en uthållig förmåga och klarar av att upprätthålla den över tid.

1. 10 rekommenderade säkerhetsåtgärder; https://www.ncsc.se/sv/cybersakerhet/10-rekommenderade-sakerhetsatgarder/
2. https://en.wiktionary.org/wiki/hot_patch
3. https://en.wikipedia.org/wiki/Blue%E2%80%93green_deployment
4. https://martinfowler.com/bliki/CanaryRelease.html

Nationellt cybersäkerhetscenter juni 2026

För mer information och kontakt
www.ncsc.se och ncsc@ncsc.se