SUMMARY CYBERSECURITY

Hacking med AI

Indledning

Keld giver et indblik i, hvordan kunstig intelligens fundamentalt har ændret angrebslandskabet, og er tager udgangspunkt i konkrete demonstrationer og hands on-eksempler, der viser, hvor lidt teknisk kunnen der i dag kræves for at udvikle avancerede angrebsværktøjer.

Praktiske erfaringer fra både incident response og red team arbejde viser, hvordan AI ikke blot effektiviserer eksisterende angrebsteknikker, men sænker adgangsbarrieren markant for nye angribere.

AI som praktisk angrebsværktøj

Kelds indlæg viste, hvordan kunstig intelligens i praksis har ændret, hvem der kan udføre avancerede cyberangreb. AI kan bruges til at udvikle, tilpasse og automatisere offensive værktøjer - uden at brugeren har dyb teknisk forståelse.

Funktioner som generativ kode, selvrettende scripts, ubegrænsede LLM modeller og autonome agenter betyder, at komplekse angreb kan bygges ved hjælp af instruktioner i naturligt sprog. Angreb, som tidligere tog uger eller måneder at udvikle, kan nu sammensættes på timer eller dage.

Trusselsbilledet: teknikker, værktøjer og aktører

På tværs af eksempler samler Keld et trusselsbillede, hvor især følgende elementer går igen:

- AI-baseret udvikling af hackingværktøjer ("vibe coding") uden manuel kodning
- Ubegrænsede LLM'er og jailbreaking, som omgår sikkerhedsfiltre
- Autonome AI-agenter, der scanner, angriber og udvider deres funktionalitet via plugins
- Plugins og open source komponenter, hvor en stor andel indeholder malware eller bagdøre
- Kriminelle økosystemer med adgangsbrokere, botnets og "hacking as a service"

AI fungerer her som en multiplikator, der gør angreb mere skalerbare, mere vedholdende og sværere at opdage.

Konsekvenser for virksomheder: når angreb bliver trivielle

Konsekvenserne af AI-drevet hacking primært opstår, fordi angreb bliver:

- **Hurtigere** – tiden fra adgang til fuldt kompromis er markant reduceret
- **Nemmere** – avancerede teknikker kræver ikke længere ekspertviden
- **Mere skjulte** – angreb udføres via legitime værktøjer, brugere og plugins

Virksomheder risikerer dermed kompromittering gennem helt almindelige handlinger, fx installation af plugins, brug af udviklingsværktøjer eller interaktion med overbevisende AI genereret indhold.

Key take aways

- Kunstig intelligens har gjort avancerede cyberangreb langt mere tilgængelige og skalérbare end tidligere.
- AI-baserede værktøjer og agenter kan udvikle, tilpasse og udføre angreb med minimal menneskelig indgriben.
- Plugins, vibe coding og supply chain komponenter udgør en voksende og ofte overset risiko.
- Virksomheder må gentænke styring, adfærdsfokus og sikkerhedsdesign for at forblive modstandsdygtige.

Vil du vide mere?

Kontakt vores specialist

Keld er vores specialist
Kontakt ham til en snak om dine virksomhedsbehov.

Keld Norman
Cybersecurity Consultant, Conscia
kno@conscia.com

Conscia
Conscia A/S
Østbanegade 135,
2100 København Ø
Denmark

Kontakt
+45 70 20 77 80
conscia.com

Conscia
Secure progress