AI – venn eller fiende?

Innovasjonsdagen 6. mars 2024

The world is how we shape it

sopra steria

Racin Gudmestad

Senior Cybersecurity Engineer

Hvem er vi

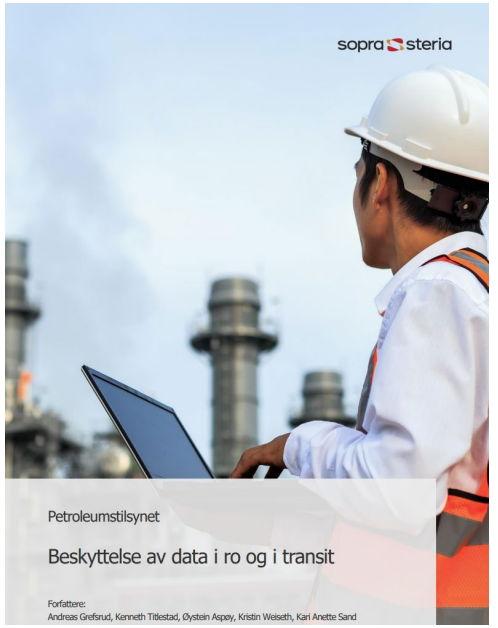
- Sikkerhetsanalytiker i Sopra Steria i 3 år
- Jobber for kunder innen Olje og Gass
- Veldig glad i å dykke dypt ned i tekniske aspektene av cyber sikkerhet

Andreas Grefsrud

Consulting director

Hvem er vi

- Prosjektleder for rapporten «Beskyttelse av data i ro og i transit»
- CISO i Klaveness i to år før retur til Sopra Steria i februar i år
- Veldig glad i risiko…
- Ingen AI-ekspert, men forstår teknologien og hvilke muligheter og farer den kan utgjøre



sopra steria

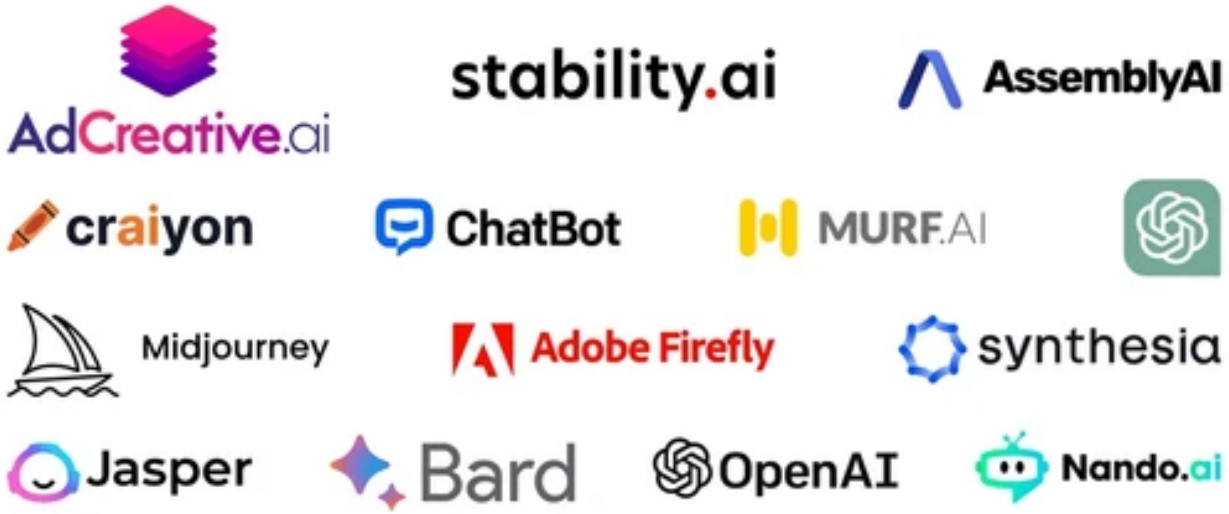
Hva vi skal snakke om i dag

1. 2023 - Transformasjonsåret
2. Hvordan trusselaktører benytter AI og hvordan vi bruker AI for å forsvare oss
3. Styring og kontroll - Noe nytt under solen?
4. Refleksjoner om risiko og hva vi må forberede oss på

Utviklingen reflekteres i markedene

Mye har skjedd på kort tid

Hvor mange av disse logoene kjente vi til før 2023?



2023

Det er ikke vanskelig å se at dette er AI-generert

OpenAI’s SORA i februar 2024

Microsoft, OpenAI: Nation-States Are Weaponizing AI in Cyberattacks

iOS, Android Malware Steals Faces to Defeat Biometrics With AI Swaps

AI will empower ransomware over the next two years

FraudGPT: The Villain Avatar of ChatGPT

Hvordan angripere benytter AI?

The dark artificial intelligence (AI) business

How to weaponize LLMs to auto-hijack websites

Cybercriminals train AI chatbots for phishing, malware attacks

Statlige aktører

Baser på rapporter fra Microsoft og OpenAI

- **Charcoal Typhoon (Kina):**
  - Forsket på ulike selskaper og cybersikkerhetsverktøy.
  - Debugget kode og genererte skripter.
  - Generere innhold til phishing-kampanjer.
- **Salmon Typhoon (Kina):**
  - Oversette tekniske artikler.
  - Hentet offentlig tilgjengelig informasjon om etterretningsetater og trusselaktører.
  - Forsket på metoder for å unngå oppdagelse av skadelig programvare.
- **Crimson Sandstorm (Iran):**
  - Genererte innhold for målrettede phishing-kampanjer.
  - Forsket på metoder for å unngå oppdagelse av skadelig programvare.
- **Emerald Sleet (Nord Korea):**
  - Identifiserte eksperter og organisasjoner fokusert på forsvarsspørsmål i Asia-Stillehavsområdet.
  - Forsket på offentlig tilgjengelige sårbarheter i software.
  - Generere innhold til phishing-kampanjer.

Effekt av AI for trusselaktører

	Statlige aktører	**organiserte kriminelle aktører**	**Hacktivister, "script kiddies", Hackers for hire**
Kapabilitet til å bruke AI	Høyt kvalifisert innen AI og cyber, godt utrustet - vil kunne lage egen AI-løsninger	Dyktig innen cyber, noen begrensninger i AI ressurser – Mest open-source AI, noe custom	Nybegynner innen cyber, begrensede til open-source AI
Sosial manipulasjon/Phishing:	Moderat forbedret og effektivisert (Allerede veldig god evne)	Sterkt forbedret og effektivisert – fra 1-til-1 --> 1-til-mange	Sterkt forbedret og effektivisert (opp fra å vær **meget** begrenset)
Implikasjoner	Best egnet til å utnytte AI sitt potensiale	Mest kapasitetsheving innen rekognosering og, sosial manipulasjon/phishing.	**Mye laver inngangs barriere**

Oppsummert:

- **Rekognosering**: Bruk av AI til å samle inn og prosesser nyttig informasjon om mennesker, selskaper, og teknologier

- **Sosial manipulasjon**: AI assistanse med oversettelser og kommunikasjon, med hensikt å etablere forbindelser eller manipulere mål.

- **Utvikling**: Utvikle og forbedre verktøy, inkludert malware. Gjennom:
  - **Sårbarhetsforskning**:Bruke AI til å finne svakheter i programvare og systemer for videre utnytelse
  - **Unntaksdeteksjonsevasjon**: utvikle metoder og verktøy som hjelper ondsinnede aktivitet med å blande seg inn med normal atferd for å unngå deteksjon

Hvordan vi som forsvarere benytter AI

Som verktøy!

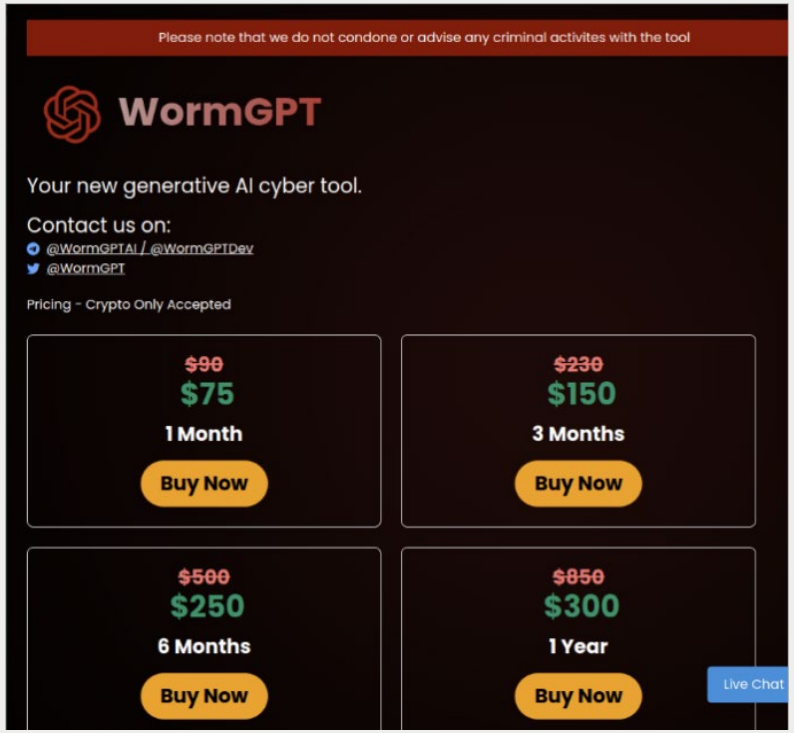
Akkurat som angripere, må vi forsvarere bruke AI
som **verktøy** til å **effektiviser** og **automatisere** det vi allerede gjør

- **Avviksdeteksjon:** avvik i system oppførsel
- **Atferdsanalyse:** avvik i bruker oppførsel
- **Trusseletterretning:**
- **Oppdage og forhindre Phishing / Sosial manipulasjon:**
- **Malware deteksjon og fil analyse:**
- **Automatisert respons:**

Hva blir det neste?

Hvordan vil AI påvirke IT sikkerhet i tiden fremover?

- Trusselen vokser like eksponentielt som AI
- Hvem som helst kan bli en "sofistikert" hacker med hjelp av AI
  - Men, hvem som helst kan også bli en "sikkerhetsanalytiker"
- Kriminelle vil ikke slutte å innovere - så det kan heller ikke vi!



sopra steria

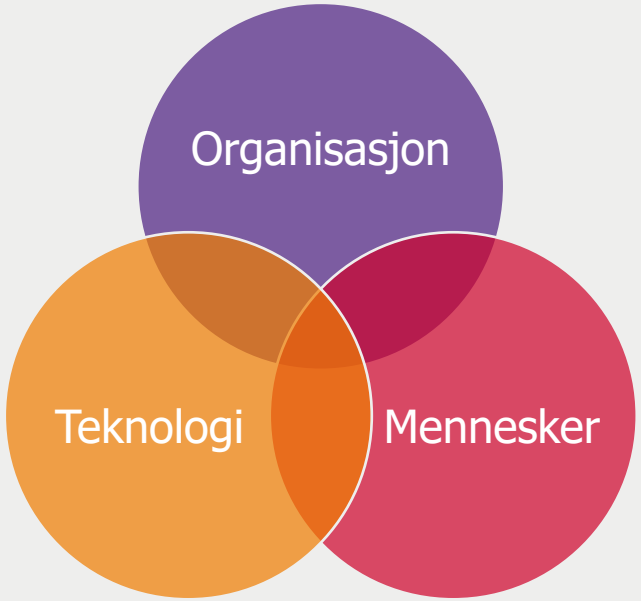
Motstandsdyktighet – hva vi alle må gjøre

Styring og kontroll blir bare viktigere

2014: «Vellykket digitalisering krever fokus på informasjonssikkerhet»

2024: «Vellykket utnyttelse av AI, krever fokus på informasjonssikkerhet»

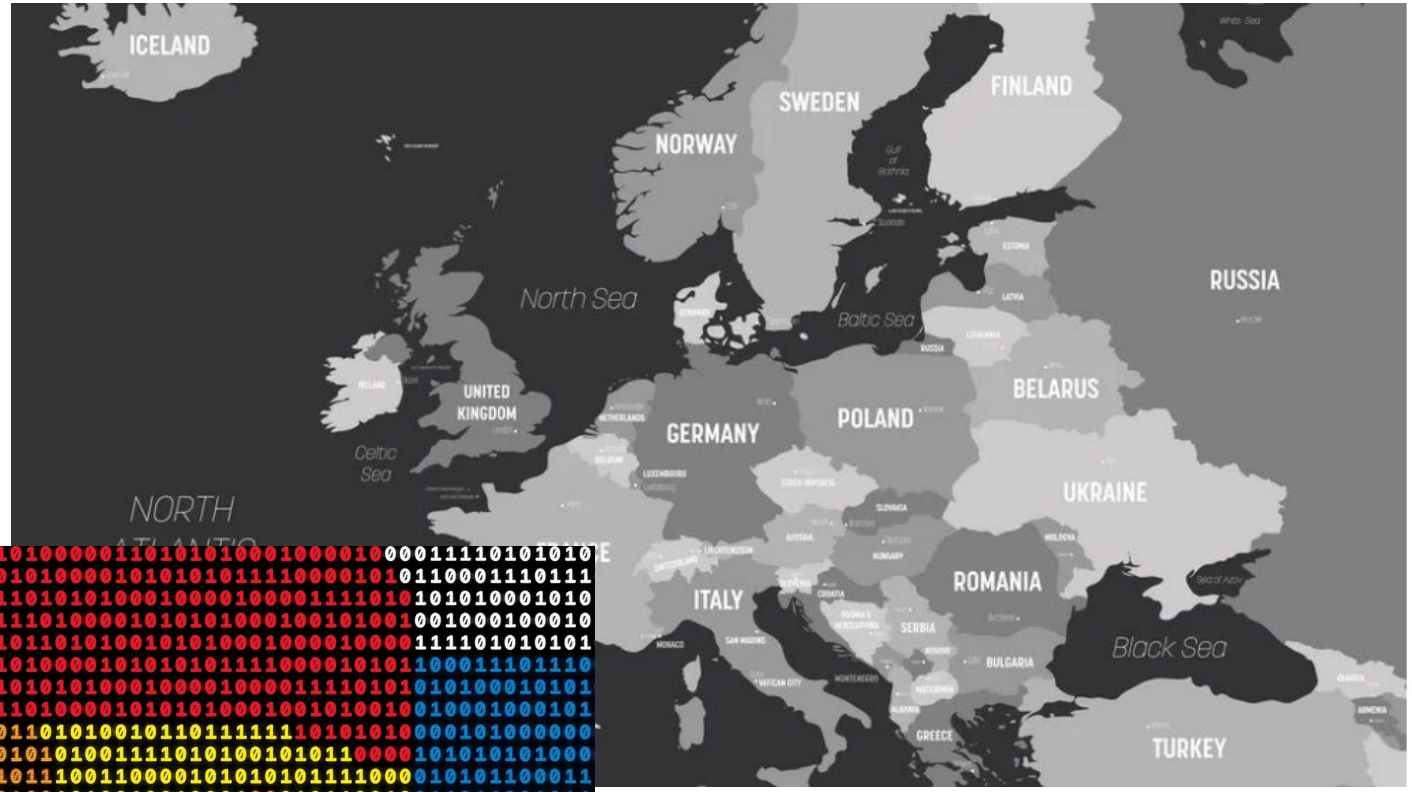
- Forretningssiden må ta mer eierskap til sikkerhetsrisiko og slutte å overlate det til IT.
- Teknologiske endringer betyr nye måter å jobbe på. Endring i seg selv krever ny kompetanse. Ny teknologi introduserer nye sårbarheter
- Opplæring!



Det store risikobildet og hva vi kan forvente mer av fremover – hva skjer hvis vi ikke lykkes?

Makrobildet

En mindre trygg verden på mange måter



"Cybersecurity is not only a series of technical threats, but a component of human security in today's digital age; and limited or poor cybersecurity can infringe upon democratic space and democracy itself"


sopra steria