Angrep og forsvar med KI

Hva slags innvirkning har KI på cybersikkerhet?

Pål H. Aaserudseter | Cybersecurity Evangelist

Office of the CTO

Oktober 2024

YOU DESERVE THE BEST SECURITY

Agenda

- Kort om Generativ AI
- Hacking lønner seg!
- Nye vektorer for angrep og forsvar
- GenAI: «all-in» eller bør vi vente litt
- Veien videre og oppsummering

GenAI

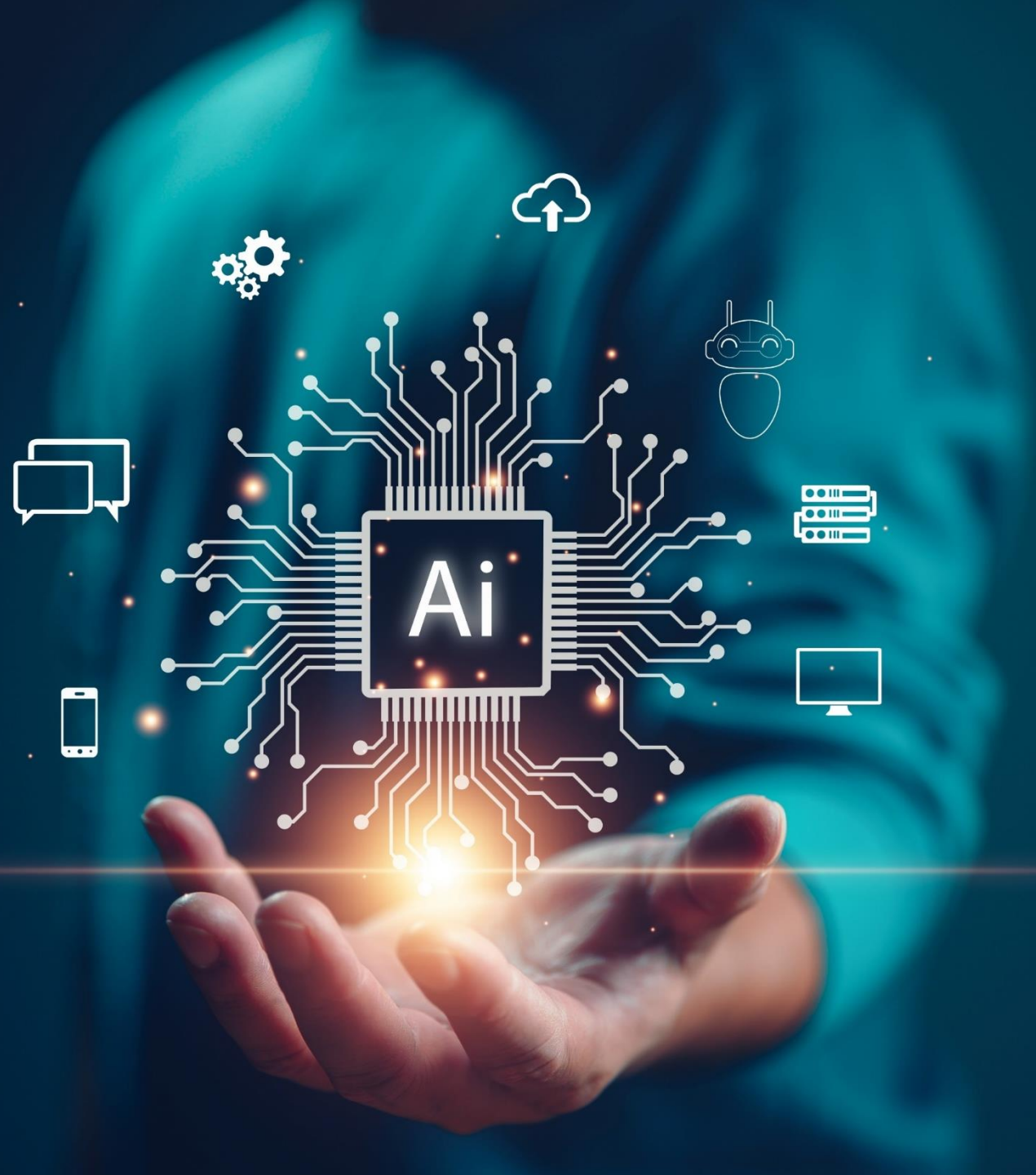
Sikkerhet!

Usikkerhet?

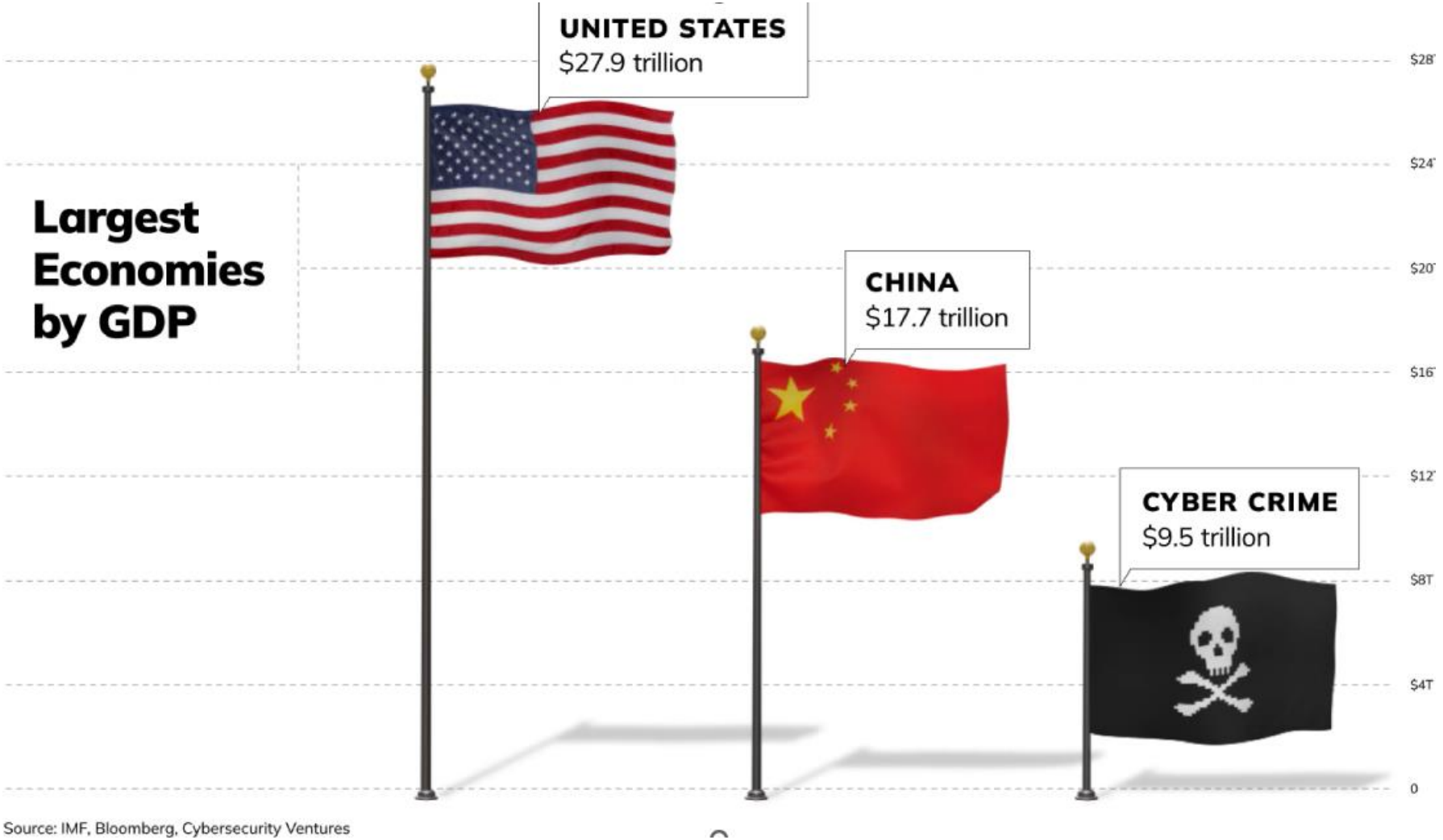
Hype?!

Frelse!

Trussel?



Cyberkriminalitet lønner seg!



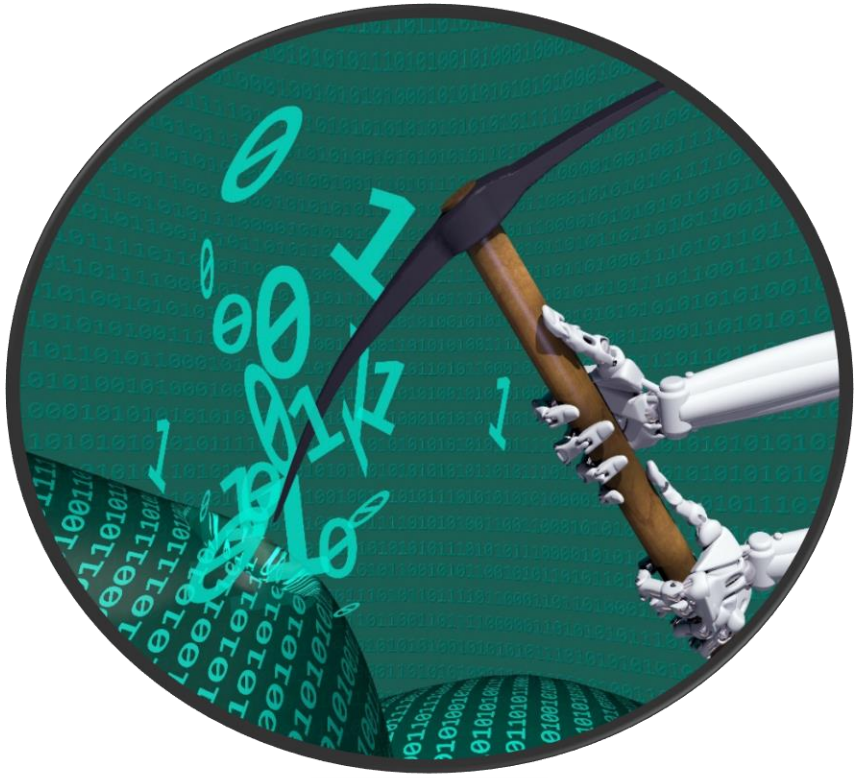

RANSOMWARE
Your Files is encrypted !
PAY $ 500 USD
TO RECOVER YOUR FILES
Time 09:59:58
PAY NOW

Hva med lover, regler og forskrifter?


I DON’T CARE!
- EU AI Act
- GDPR
- NIS2
- DORA
- SOX
- ISO 27xxx
- CISA
- NIST
- DPA
- SOCI
- EU CRA
- .....

HVORDAN BRUKES AI TIL ANGREP?

Hvordan brukes AI til **TEKNISKE** angrep?



Datafangst og Analyser

AI kan prosessere enorme datasett for å trekke ut verdifull informasjon som kan brukes til identitetstyveri, økonomisk svindel eller andre ondsinnede formål.

Maskinlæring kan analysere stjålne data for å finne de mest lukrative målene for ytterligere angrep.

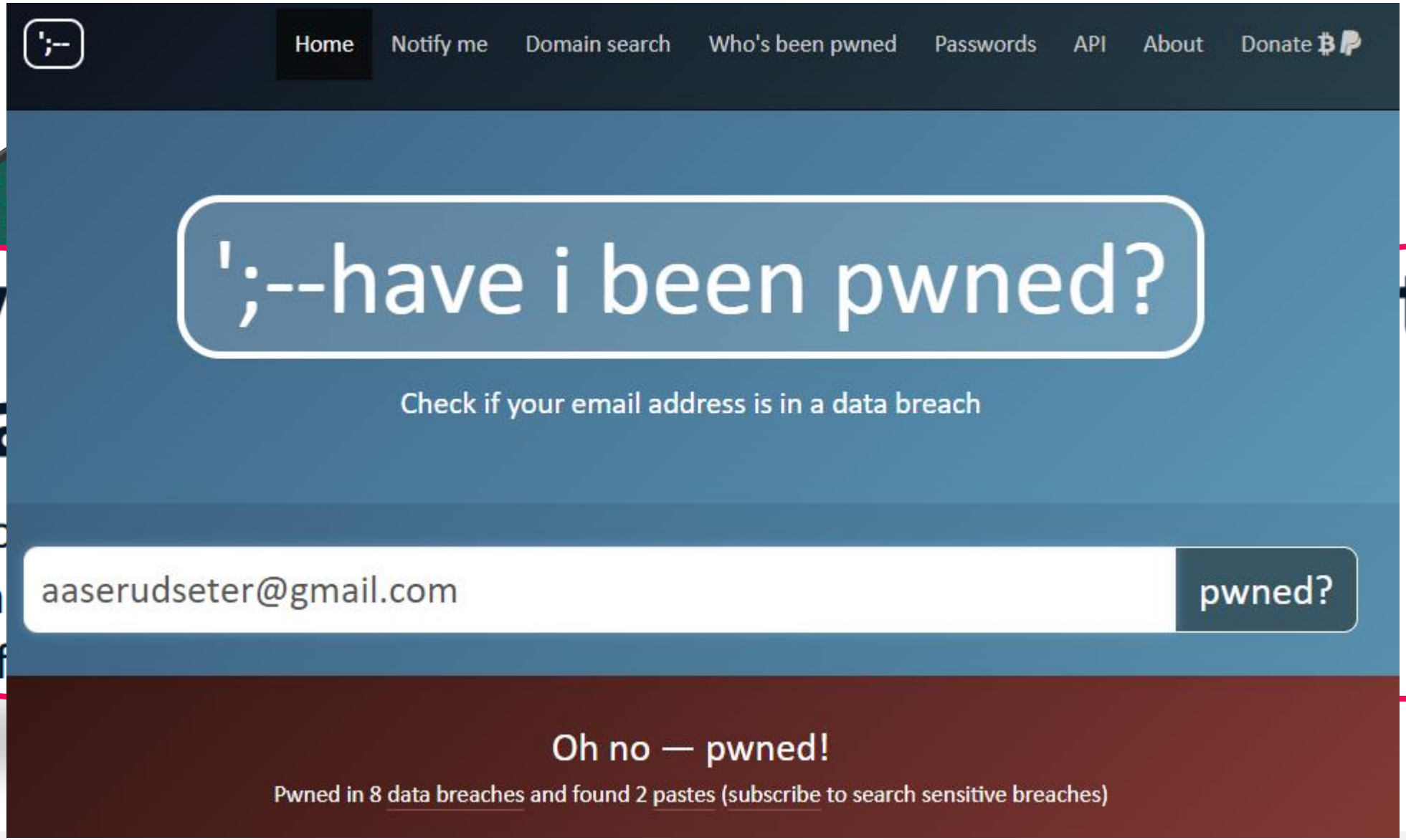
Hvordan brukes AI til **TEKNISKE** angrep?

Datafangst og Analyser

Advarer mot skreddersydd svindel etter datalekkasje

Personopplysningene som er på avveie etter det store datainnbruddet mot Norkart, kan åpne for mer utspekulerte former for svindel, tror ekspert.

Hvordan brukes AI til **TEKNISKE** angrep?



Hvordan brukes AI til **TEKNISKE** angrep?

Oppdage nulldagshendelser

GenAI kan brukes av sofistikerte trusselaktører for å automatisere deler av prosessen for å oppdage sårbarheter.

AI-systemer kan analysere store mengder kode eller nettverkstrafikk for å identifisere svake punkter raskere enn mennesker kan.

Hvordan brukes AI til **TEKNISKE** angrep?

Oppdage nulldagshendelser

Nulldagssårbarhet i Ivanti Endpoint Manager (MobileIron Core)

Publisert: 24.07.2023 Oppdatert: 25.07.2023

På pressemøtet mandag 24. juli 2023 informerte Nasjonal sikkerhetsmyndighet og Departementenes sikkerhets- og serviceorganisasjon (DSS) om at det var en nulldagssårbarhet som ble benyttet til å utføre dataangrepet mot Departementenes sikkerhets- og serviceorganisasjon (DSS).

Hvordan brukes AI til **TEKNISKE** angrep?

AI Forgiftning

Angripere manipulerer bevisst data som brukes til å trene eller oppdatere en AI-modell.

Målet med AI-forgiftning er å manipulere modellens atferd på en måte som gagner angriperen, noe som fører til feilaktige forutsigelser, klassifiseringer eller beslutninger.

Hvordan brukes AI til **TEKNISKE** angrep?

AI Forgiftning

Hvordan brukes AI til **SOSIALE** angrep?

Deepfake-teknologi

Cyberkriminelle bruker AI til å lage deepfake-videoer og stemmeklipp, som kan brukes til å bedra ofre eller spre desinformasjon.

Dette kan for eksempel være falske videoer av bedriftsledere som gir instruksjoner om økonomiske transaksjoner.

Hvordan brukes AI til **SOSIALE** angrep?

Deepfake-teknologi

Vår Staude svindelmisbrukt i falsk Nettavisen-artikkel

transaksjoner.

Hvordan brukes AI til **SOSIALE** angrep?

Deepfake-teknologi

Vår Staude svi
Nettavisen-art

transaksjoner.

Norske profiler misbrukes i videoer: – Veldig ubehagelig

Svindlere bruker nå kunstig intelligens for å lage falske videoer av norske profiler. – Det er et helt nytt nivå.

FOTO: SKJERMBILDE FRA SVINDELVIDEO.

Egil Aslak Hagerup
Journalist

Publisert 30. mai kl. 09:30
Oppdatert 30. mai kl. 10:34

Hvordan brukes AI til **SOSIALE** angrep?

Deepfake-teknologi

Ble svindlet for 275 millioner:

Trodde han var i videokonferanse med kolleger – ingen var ekte

Norske profiler misbrukes i videoer: – Veldig ubehagelig

Svindlere bruker nå kunstig intelligens for å lage falske

lt nytt nivå.

FOTO: SKJERMBILDE FRA SVINDELVIDEO.

Egil Aslak Hagerup
Journalist

Publisert 30. mai kl. 09:30
Oppdatert 30. mai kl. 10:34

Hvordan brukes AI til **SOSIALE** angrep?

Deepfake-teknologi

Ble svindlet for 275 millioner:

Trodde han var
videokonferans
kolleger – inge

er: – Veldig

Egil Aslak Hagerup
Journalist

Publisert 30. mai kl. 09:30
Oppdatert 30. mai kl. 10:34

Hvordan brukes AI til **SOSIALE** angrep?

Sosial manipulering

AI kan analysere sosiale medier og andre data for å forstå og forutsi menneskelig atferd.

Denne innsikten brukes til å utføre målrettede sosiale manipulasjonsangrep, hvor angripere utnytter menneskelige svakheter for å få tilgang til informasjon eller systemer.

Hvordan brukes AI til **SOSIALE** angrep?

Sosial manipulering

Falske kontoer spredte drapsanklager: – Profitterer på tragedie

Det siste døgnet har en navngitt person blitt hengt ut som gutten bak knivdrapene i Southport. Men det var noe som skurret.

Hvordan brukes AI til **SOSIALE** angrep?

Sosial manipulering

Falske kontoer spredte

drap … **å**

trag

Det siste …

knivdrap…

Kortversjonen

- En 17 år gammel gutt er pågrepet for å ha knivstukket flere barn på et dansekurs i Southport i England. Tre er så langt drept.
- På sosiale medier spres falsk informasjon om gjerningspersonens identitet ved hjelp av boter.
- Professorer advarer om spredning av desinformasjon og mener det er kommersielle motiver bak disse botene.

Vis mindre

Hvordan brukes AI til **SOSIALE** angrep?

Sosial manipulering

Falske kontoer spredte

drap

trag

Det siste d

knivdrap

Kortversjonen

- En 17 år gammel gutt er pågrepet for å ha knivstukket flere barn på et dansekurs i Southport i England. Tre er så langt drept.
- På sosiale medier spres falsk informasjon om gjerningspersonens identitet ved hjelp av boter.
- Professorer advarer om spredning av desinformasjon og mener det er kommersielle motiver bak disse botene.

Vis mindre ^

Trump Pleads Ignorance After Sharing AI-Generated Taylor Swift Images

The former president says he's not worried about getting sued because he didn't generate the images himself

BY NAOMI LACHANCE

AUGUST 22, 2024

Taylor Swift and Donald Trump. NEILSON BARNARD/GETTY IMAGES/THE RECORDING ACADEMY; PETER ZAY/AFP/GETTY IMAGES

Kriminelles bruk av **AI** gir store utfordringer innen **sikkerhet**

Vi må bruke like **avanserte** tiltak for å motvirke **truslene**

Hvordan brukes AI til forsvar?

Hvordan brukes AI til forsvar?

AI gir Bedre Sikkerhet

AI-Drevet Trussel Beskyttelse

AI-Drevet Assistent for Admins & Analytikere

Beskytte AI Servere

Aktivere Trygg Bruk av GenAI

ThreatCloud AI

50+ **trusselforebyggende** motorer for 99,8 % fangstrate for skadelig programvar.
Trusselintelligens i sanntid

AI Copilot

Sparer opptil 90 % av tiden som trengs for å utføre vanlige administrasjonsoppgaver
Akselererer SecOps trusseljakt, analyse og automatisert respons

AI Cloud Protect

Nvidia-partnerskap for å beskytte AI-skyinfrastruktur som brukes av bedrifter for deres egne AI-apper

GenAI Sikkerhet

Muliggjør sikker adopsjon av GenAI i bedriften; leverer oppdagelse, risikoinnsikt og databeskyttelse i sanntid

Eksempler på bruk av AI hos Check Point

- **Totalt 90+ Trusselmotorer**
  - 50 AI-baserte
  - 40 Tradisjonelle

- **Av 50 AI-baserte motorer er**
- 10 Deep Learning teknologi
  - 40 Klassisk Machine Learning teknologi

Eksempler på bruk av AI hos Check Point

Big data threat intelligence:

2,800,000,000
Websites and files inspected

146,000,000
Full content emails

53,000,000
File emulations

20,000,000
Potential IoT devices

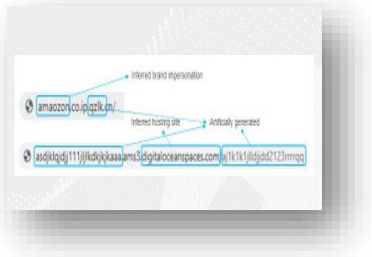
2,600,000
Malicious indicators

1,500,000
Newly installed mobile apps

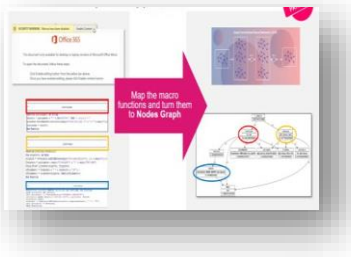
1,200,000
Online web forms

Prosessert
HVER DAG!

ThreatCloud AI **Stopper** Angrep!

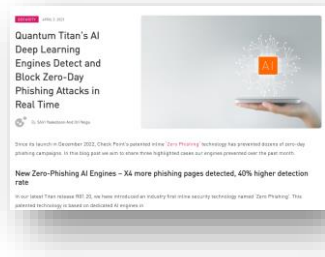


Skadelige URL'er

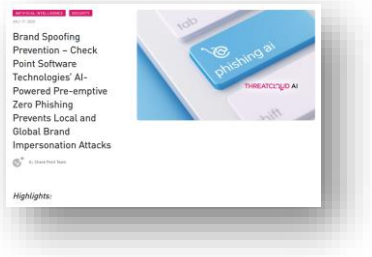


Avanserte Makroer

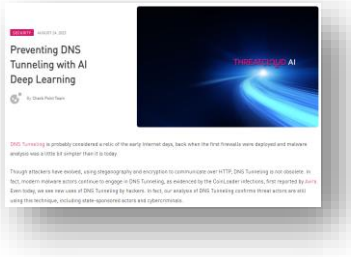
Skadelig kode



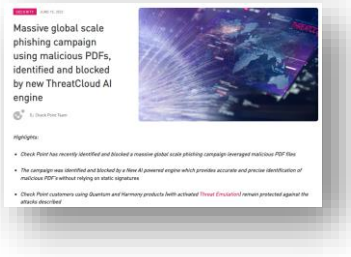
Phishing angrep



Forfalskede Nettsteder

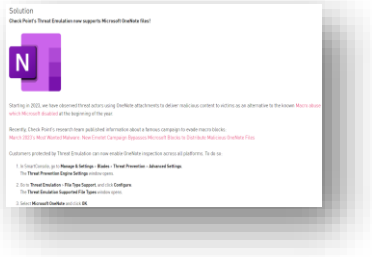


DNS Tunnelering

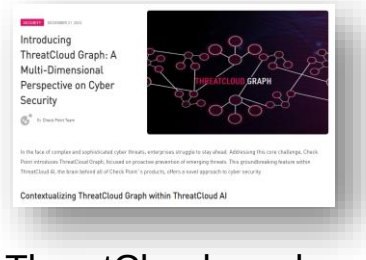


Avanserte Skadelige PDF filer

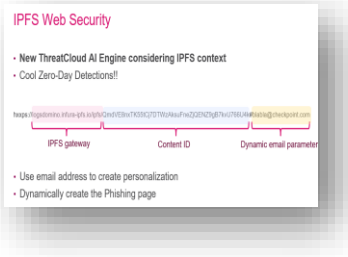
Skadelige LNK filer



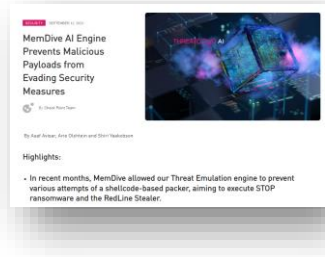
OneNote trusselbeskyttelse



ThreatCloud graph



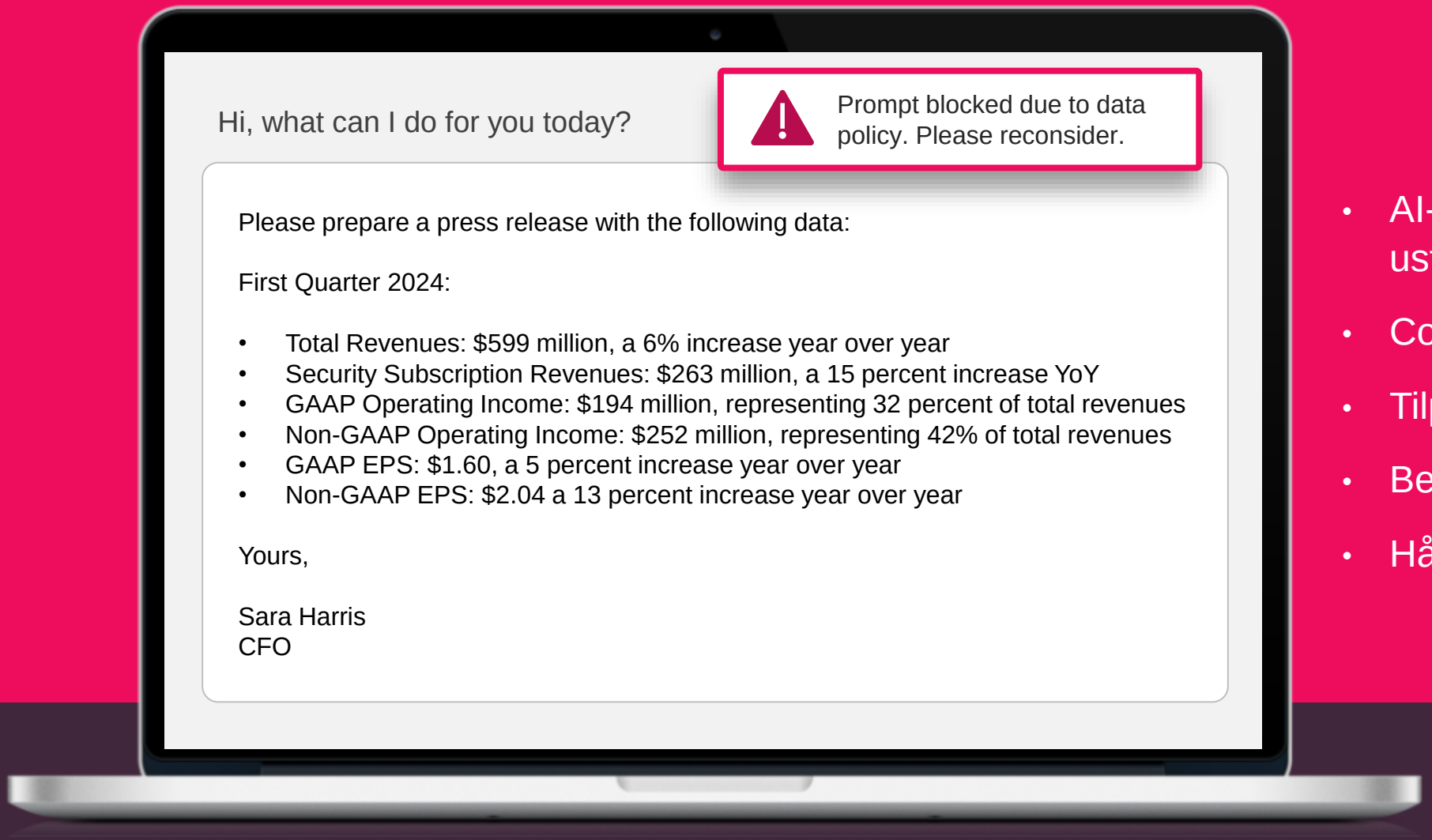
IPFS angrep



Skadelig programvare

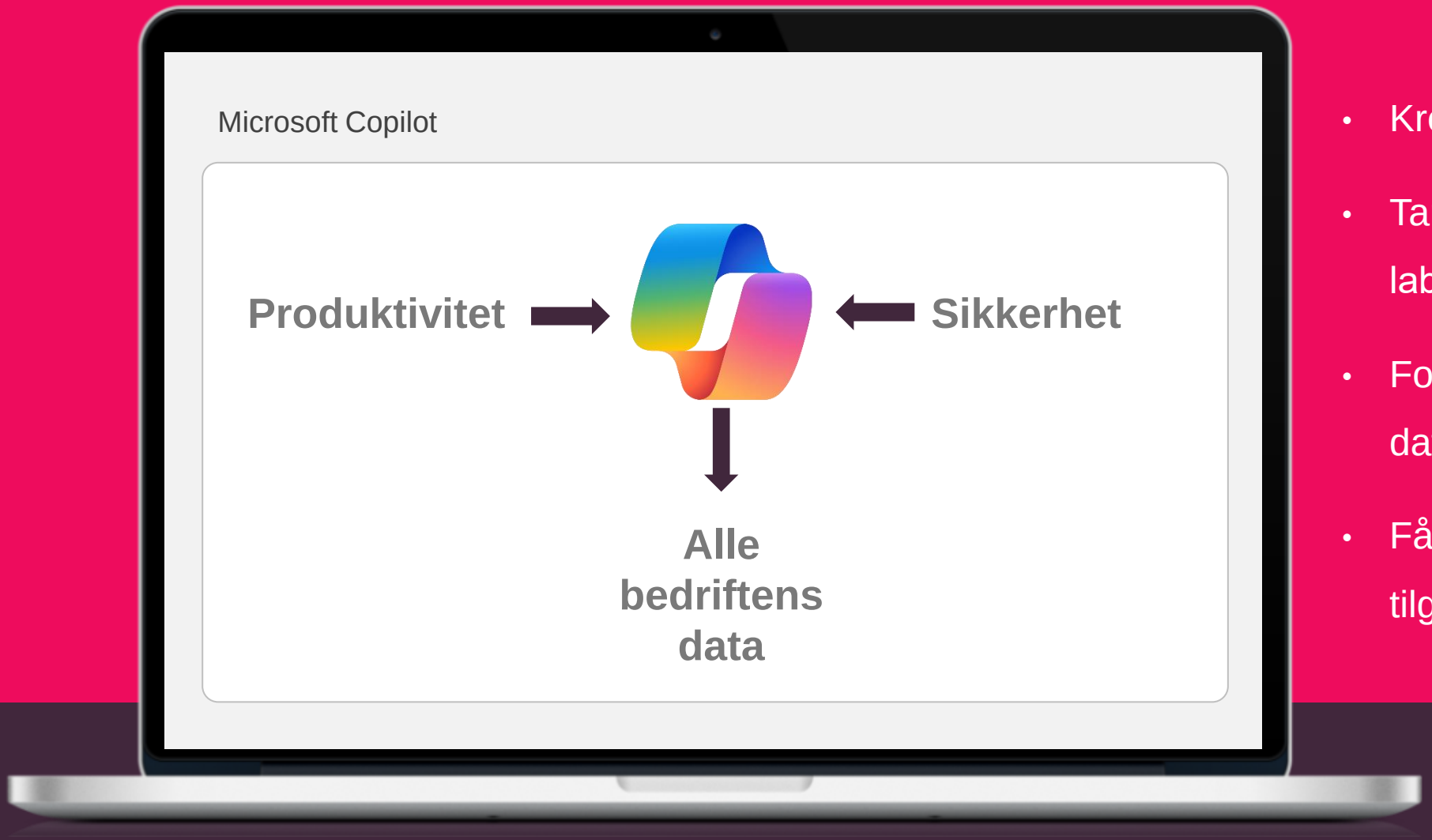
Generativ AI
Øker bedriftens forsprang

Stopp Datalekkasjer i Sanntid



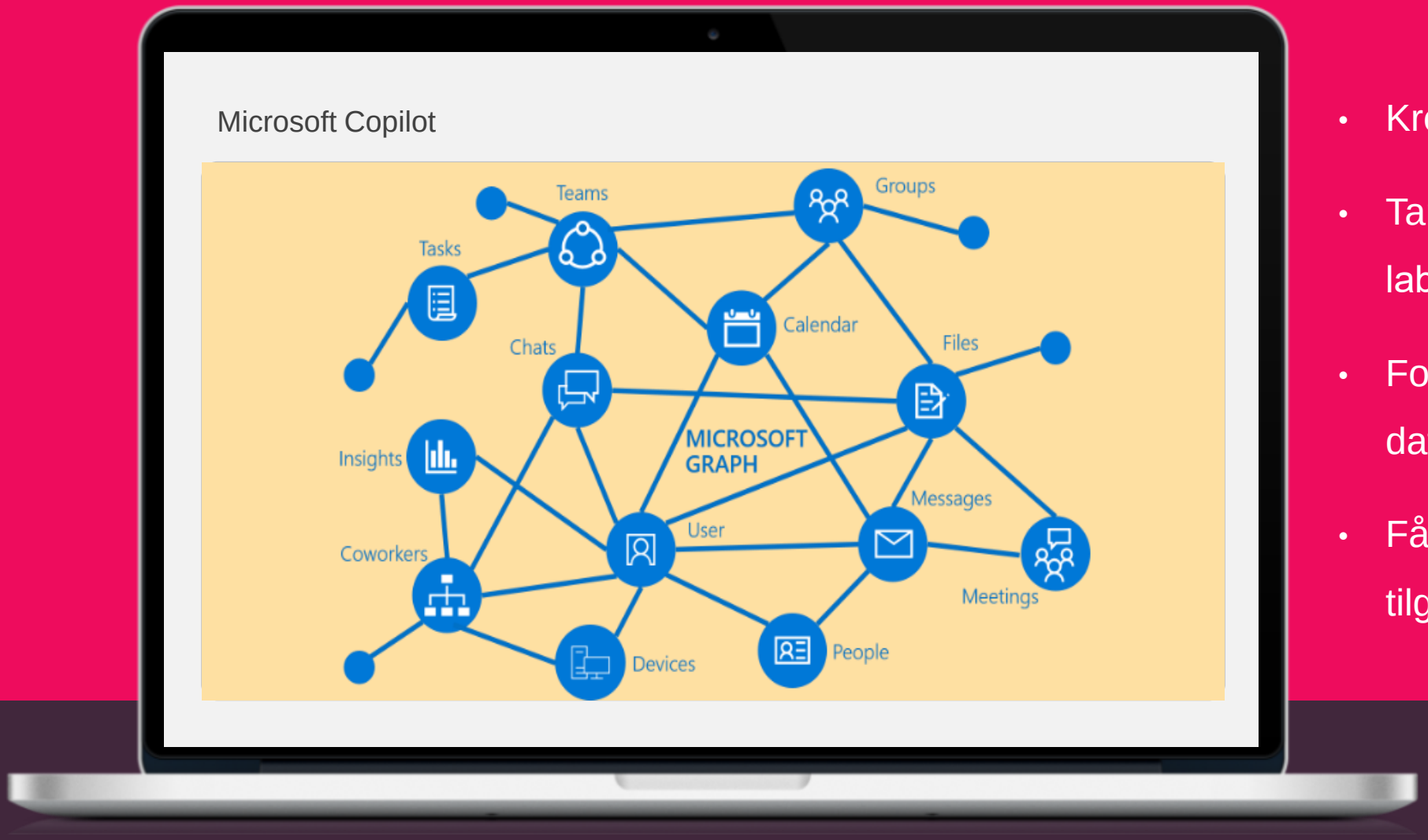
- AI-drevet klassifisering av ustrukturerte data
- Copy/paste restriksjoner
- Tilpasses bedriftens behov
- Beskytt dine data/åndsverk
- Håndtering av personvern

Stopp Datalekkasjer i Sanntid



- Krever **stålkontroll** på rettigheter
- Tar ikke hensyn til “Security labels” eller klassifisering
- Fort gjort å eksponere sensitive data (gjelder all GenAI)
- Får de tilgang til mailkonto har du tilgang til “alt” (MS Graph API)

Stopp Datalekkasjer i Sanntid



- Krever **stålkontroll** på rettigheter
- Tar ikke hensyn til “Security labels” eller klassifisering
- Fort gjort å eksponere sensitive data (gjelder all GenAI)
- Får de tilgang til mailkonto har du tilgang til “alt” (MS Graph API)

Hvorfor er vår AI trygg, og ikke andres?

Begrenset tilgang (ML/DL)

Telemetri

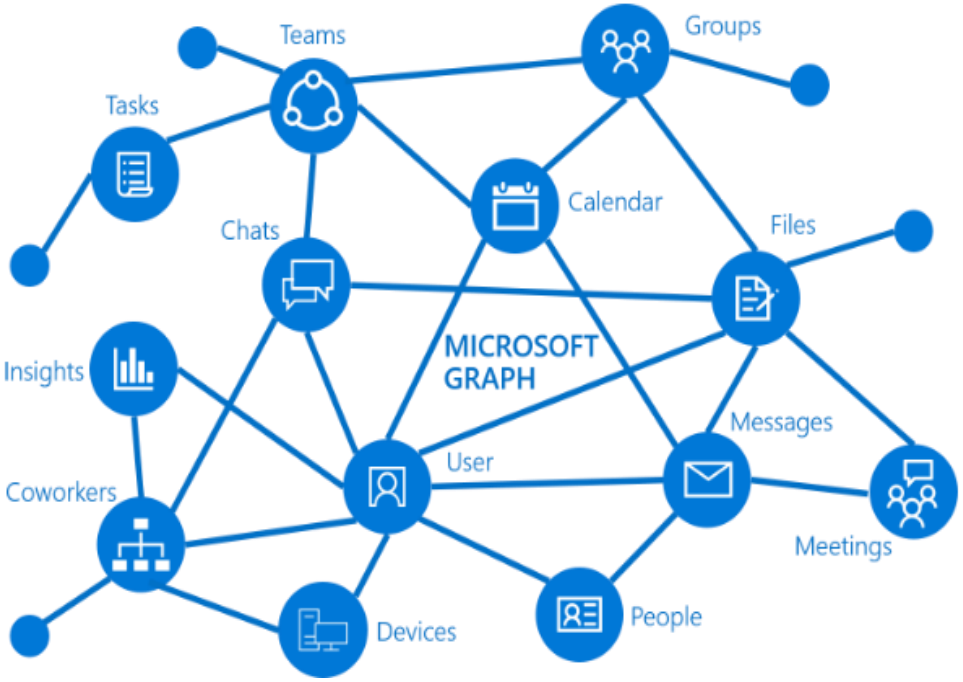
IoC'er

Trusler

THREATCLOUD AI

Resultat:

Full tilgang (LLM)



Resultat: Alt!

Hvorfor er vår AI trygg, og ikke andres?

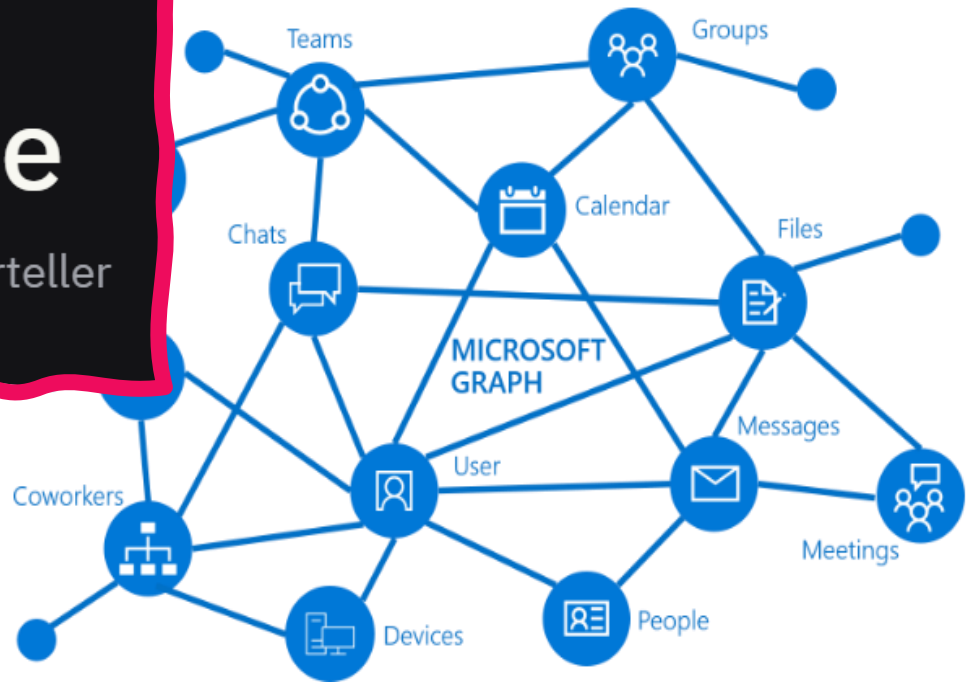
NTNU vil neppe innføre Copilot – frykter overvåkning av ansatte

Det er lite sannsynlig at NTNU vil ta i bruk Copilot for Microsoft 365, forteller IT-direktøren etter fersk rapport om en pilot med løsningen.

Trusler

Resultat:

Full tilgang (LLM)



Resultat: Alt!

Hvorfor er vår AI trygg, og ikke andres?

NTNU vil neppe innføre Copilot – frykter ove

Det er lite s

IT-direktøre

NTNUs Copilot-rapport møter kritikk

Trusler

– Hvis Copilot behandler persondata feil, så er det fordi at man har lagret persondata feil

Full tilgang (LLM)

Resultat:

Resultat: Alt!

HVA BRINGER FREMTIDEN?

- Avanserte angrep
- Etikk & regler
- Spesialister
- Hybridforsvar
- GenAI krever kontroll

Who you gonna call?

TLP:CLEAR

NØDPLAKAT FOR DIGITALE ANGREP

1. VARSLE

Varsle IKT driftsorganisasjon, og mobiliser beredskap i egen virksomhet. Husk krise-kommunikasjon mot eksterne og interne.

Politiets nasjonale cyberkrimsenter (NC3):
23 20 80 00
Kripos forebygger, avdekker og bekjemper trusler og kriminalitet i det digitale rom. Tidlig varsling bidrar til at politiet kommer raskt i gang med å sikre digitale spor, slik at trusselaktøren kan straffeforfølges gjennom nasjonalt og internasjonalt politisamarbeid. Døgnbemannet.

Nasjonalt cybersikkerhetssenter (NCSC):
23 31 07 50
NSMs nasjonale responsfunksjon for digitale angrep.Kan blant annet bidra med rådgivning. Drifter også det nasjonale varslingssystemet for digital infrastruktur. Døgnbemannet.

Økokrim (FIU):
23 29 14 05
Enheten for finansiell etterretning kan i mange tilfeller bistå ved internasjonal stans av penge-transaksjoner relatert til hvitvasking og terror-finansiering. Åpent 09 til 15 på hverdager.

2. HÅNDTERE

Be om akutt bistand fra leverandører du har avtale med, eller andre som tilbyr tjenester for håndtering av dataangrep. Se godkjente leverandører til høyre. Innhent råd om umiddelbare tiltak, og iverksett dette. Gi gjerne hendelses-håndteringsleverandøren fullmakt til å ha direkte kontakt med myndighetene (NC3/NCSC), og å dele informasjon med disse fortløpende.

Vurder om saken må meldes til Datatilsynet.
Det er i de fleste tilfeller lovpålagt å melde ulovlig tap, spredning eller endring av personopplysninger til Datatilsynet innen 72 timer. Se mer info på www.datatilsynet.no/brudd eller skann QR-koden til venstre.

Anmeld saken til ditt lokale politidistrikt.
Henvis lokalt politi til forutgående kontakt med NC3/NCSC. Politiets næringslivskontakter (se under) har ingen operativ funksjon ved hendelser, men kan bistå med råd og veiledning. Hvis du ikke vil anmelde, kan du velge å bare tipse. Se QR-kode for mer info.

3. GJENOPPRETTE

Iverksett beredskapsplaner for nøddrift. Etter et dataangrep kan det ta mange uker å gjenopp-rette normal drift, så det er viktig å iverksette alternative driftsmetoder tidlig.

Godkjente leverandører:
Følgende leverandører har døgnåpne vakttelefoner for håndtering av data-angrep, og er godkjent av NSM i henhold til deres kvalitetskrav.

Defendable
91 80 80 30
ir@defendable.no

mnemonic
23 20 47 47
Reserve: 23 20 28 25
mirt@mnemonic.no

ATEA
03060
irt@atea.no

Netsecurity
92 24 73 65
irt@netsecurity.no

sopra steria
24 14 04 56
no.irt@soprasteria.com

Orange Cyberdefense
23 96 63 20
csirt-no@orangecyberdefense.com

Vi gjør oppmerksom på at flere andre aktører tilbyr tilsvarende tjenester, selv om de ikke er en del av den frivillige kvalitetsordningen til NSM.

Politiets næringslivskontakter:

Økokrim:	Fredrik Roe Bee	408 69 977	fredrik.roe.boe@politiet.no
KRIPOS:	Else Gun Ommundsen	91 69 51 19	ommundsen@politiet.no
KRIPOS:	Espen Skogstad	95 06 45 01	skogstad@politiet.no
Oslo PD:	Christina T. Rooth	404 34 545	christina.rooth@politiet.no
Agder PD:	Erik Berg Thorsrud	904 11 051	eirik.berg.thorsrud@politiet.no
Sør Øst PD:	Hanna Holmelin	47 46 89 34	hanna.holmelin@politiet.no
Sør Vest PD:	Kyrre Lindanger	916 71 726	kyrre.lindanger@politiet.no
Vest PD:	Sonja Lund	924 97 728	sonja.lund@politiet.no
Trøndelag PD:	Terje Lunde	481 61 969	terje.lunde@politiet.no
Troms PD:	Merethe Samuelsen	915 17 939	merethe.samuelsen@politiet.no
Møre og Romsdal PD:	Kjell Arne Hestad	926 06 045	hestad@politiet.no
Nordland PD:	Håvard Fjærli	918 83 382	havard.fjarli@politiet.no
Øst PD:	Jan Kevin Brunvoll	975 77 266	jan.kevin.brunvoll@politiet.no
Innlandet PD:	Lene Espelund	958 99 584	Lene.Espelund@politiet.no
Finnmark PD:	Jan Arne Pettersen	900 96 960	jan.arne.pettersen@politiet.no

Næringslivets sikkerhets-råd

Næringslivets Sikkerhetsråd (NSR) er stiftet av Finans Norge, NHO, Virke, Spekter, Rederiforbundet og Den Norske Krigsforsikring for skib, og arbeider kontinuerlig med å kartlegge, forebygge og bekjempe tilsiktede sikker-hetstruseler mot norsk næringsliv. Nødplakaten er kvalitets-sikret av NSRs informasjonssikkerhetsutvalg.

Nødplakaten er utarbeidet i samarbeid med:
POLITIET
NASJONAL SIKKERHETSMYNDIGHET

KORT OPPSUMMERT

- (Generativ) AI…
  - Er et veldig bra hjelpemiddel
  - Bør vel ikke stoles helt på
  - Må også sikres og krever stålkontroll på rettigheter
  - Er til stor hjelp for Cyberkriminelle
  - Er helt nødvendig for å stoppe nye og ukjente angrep

CHECK POINT™

Tusen takk!

paala@checkpoint.com

YOU DESERVE THE BEST SECURITY