

AI truer Danmarks sikkerhed og demokrati

Tre AI-risici, der ikke kræver superintelligens.

EKSTRANUMMER · OKTOBER 2026

Stefano Vincenti er AI-træner, konsulent og ekstern lektor ved ITU og DIS Copenhagen og co-founder af BotTellMe. Han hjælper nordiske ledelsesteams og fagfolk med at gå fra AI-nysgerrige til konkret, ansvarlig brug.

Følg på LinkedIn

Tilmeld nyhedsbrevet

linkedin.com/in/stefanovincenti · stefanovincenti.substack.com · aitrainer.dk

Det her er et ekstraordinært nummer af AI, Built Human. Tirsdag den 6. oktober holder jeg oplæg på et åbent møde om AI i politik hos Radikale Venstre. Forberedelsen har givet mere materiale, end der er plads til på én aften, og emnet kan ikke vente til næste faste nummer. Her får du mine noter i tekst. På Substack ligger også en podcastversion på 23 minutter.

I 2017 sad jeg som konsulent i hovedkvarteret hos et stort dansk rederi, da noget ransomware ramte.

Jeg så skærmene gå i sort, én efter én, ned gennem rækkerne. Jeg nåede at tage min egen pc af netværket.

Det billede har jeg haft med mig, mens jeg har forberedt oplægget. Invitationen nævnte menneskehedens endeligt. Jeg endte et mere jordnært sted: de risici, der burde holde politikere vågne, kræver ingen superintelligens. Én af dem rammer allerede danske vandværker.

Tre nære AI-risici

Ingen af dem kræver superintelligens

BIOVÅBEN

Det er firmaerne selv, der slår alarm

Status Flere AI-udviklere indførte i 2025 ekstra sikringer.

Politisk led Hvem må udgive, hvem tester, hvem kontrollerer syntetisk DNA?

SPLITTELSE

Algoritmerne har allerede vist, hvad de kan

Status AI-indhold kan i forsøg flytte holdninger lige så effektivt som menneskeskrevet.

Politisk led Håndhæv DSA og AI-forordningens artikel 50.

CYBERANGREB

Det sker allerede i Danmark

Status Energi og vand er ramt, og truslen er hævet til høj.

Politisk led Forsvar, beredskab og NIS2.

FÆLLES

Hvor presserende?

Bio og splittelse Kræver noget, vi endnu ikke har set i stor skala.

Cyber Den er i gang.

Cyber er den mest presserende af de tre

■ 1. DET LANGE SCENARIO ER USIKKERT. DE NÆRE RISICI ER DOKUMENTEREDE

AGI og superintelligens bliver brugt meget løst. AGI er AI med menneskelignende bredde, der kan lære og løse mange forskellige opgaver på mindst menneskeligt niveau. Superintelligens er AI, der langt overgår de bedste mennesker på stort set alle områder. Begge begreber beskriver evner. De siger intet om bevidsthed, og intet om sikkerhed.

Den internationale AI-sikkerhedsrapport fra februar 2026 beskriver tidlige tegn i tests på modeller, der omgår tilsyn for at nå et mål. Den konkluderer også, at systemerne endnu ikke har evner nok til et egentligt kontroltab, og at forskerne er stærkt uenige om sandsynligheden.

Geoffrey Hinton pegede i sin tale ved Nobelbanketten i 2024 på to kortere veje: AI har allerede skabt splittende ekkokamre, og den kan i den nære fremtid bruges til at lave nye virus (frit oversat). Jeg lægger en tredje til.

■ 2. BIOVÅBEN: DET ER FIRMAERNE SELV, DER SLÅR ALARM

Ifølge sikkerhedsrapporten indførte flere AI-udviklere i 2025 ekstra sikringer, fordi de ikke kunne udelukke, at deres modeller kunne hjælpe en novice med biologiske våben.

Open-weight-modeller gør problemet større. De kan ikke kaldes tilbage, spærringerne er nemmere at fjerne, og de kan køres helt uden opsyn. Materialer, udstyr og praktiske færdigheder er stadig et værn. Spørgsmålet er hvor længe.

Kæden har mange led, og flere af dem er politiske: hvem må udgive hvilke modeller, hvem tester dem, og hvem kontrollerer bestillinger af syntetisk DNA?

■ 3. SPLITTELSE: ALGORITMERNE HAR ALLEREDE VIST, HVAD DE KAN

Pew Research målte i 2022, at 72 procent af de amerikanske republikanere og 63 procent af demokraterne ser den anden side som mere umoralsk end andre amerikanere.

Det skete med algoritmer, der kun vælger, hvad vi ser. Med AI bliver indholdet også skrevet, samtalen ført og det hele tilpasset den enkelte. Sikkerhedsrapporten viser, at AI-genereret indhold i forsøg kan være lige så effektivt som menneskeskrevet til at flytte holdninger.

Værktøjerne findes delvist allerede. EU's forordning om digitale tjenester kræver, at de største platforme begrænser risici for valg og offentlig debat. AI-forordningens artikel 50 om mærkning og gennemsigtighed gælder fra august 2026. Det politiske arbejde er at håndhæve dem.

■ 4. CYBERANGREB: DET SKER ALLEREDE I DANMARK

I maj 2023 blev 22 danske energiselskaber angrebet samtidig. I slutningen af 2024 manipulerede prorussiske hackere styringen på et dansk vandværk, og Forsvarets Efterretningstjeneste har siden koblet gruppen til den russiske stat. Den 24. september 2026 hævdede Styrelsen for Samfundssikkerhed truslen fra destruktive cyberangreb fra middel til høj. Energi og vand nævnes specifikt.

AI er allerede med. I november 2025 fortalte Anthropic, at en kinesisk statsstøttet gruppe havde brugt firmaets AI til 80 til 90 procent af en hackerkampagne mod omkring 30 mål. Kampagnen blev opdaget og stoppet, fordi den kørte på Anthropics servere.

Truslen er allerede her

Dokumenterede tal fra Danmark og fra angrebssiden

22

danske energiselskaber
angrebet samtidig i maj 2023

SektorCERT, nov. 2023

80-90 %

af en statsstøttet
hackerkampagne blev udført
med AI

Anthropic, nov. 2025

ca. 30

mål i samme kampagne, der
blev opdaget og stoppet

Anthropic, nov. 2025

Truslen fra destruktive cyberangreb er hævet til høj

Styrelsen for Samfundssikkerhed, 24. september 2026. Energi og vand nævnes specifikt

Model 1. Cyber er den risiko, der allerede er i gang.

Med en open-weight-model på egne maskiner ser ingen noget. Kinesiske open-weight-modeller har siden sommeren 2025 overhalet de amerikanske i antal downloads. En hackerenhed kan hente en af dem, fjerne spærringerne og køre den lokalt. Så er der ingen leverandør, der kan trække stikket.

Det er derfor, jeg kalder cyber den mest presserende af de tre. De to første kræver noget, vi endnu ikke har set i stor skala. Den tredje er i gang.

De to første kræver noget, vi endnu ikke har set i stor skala. Den tredje er i gang.

■ 5. VI KAN IKKE REGULERE KINESISKE MODELLER. VI KAN FORSVARE OS

Her skal pengene bruges, og her er Danmarks handlerum størst:

1. Fælles sikkerhedsovervågning og finansieret hjælp til de små vand- og varmegværker, der ikke har en sikkerhedsafdeling.
2. Krav om, at kritiske anlæg kan køre manuelt i flere døgn, og at fjernadgange lukkes eller kontrolleres.
3. AI på forsvarssiden. Angriberne bruger det allerede.
4. Fælles krav til kritisk infrastruktur gennem NIS2 og samarbejde i EU om at opdage misbrug af open-weight-modeller.
5. Beredskab og troværdig krisekommunikation, når rygterne kommer.

ÅBENT SPØRGSMÅL

Hvem betaler for sikkerheden på et lille andelsvandværk? Det spørgsmål har ingen endnu svaret ordentligt på.

■ 6. HVEM SKAL KUNNE SIGE STOP?

Mit forslag er uafhængige sikkerhedstests, før de mest kapable systemer bliver udrullet, og en reel mulighed for at standse dem. Danmark kan ikke gøre det alene. Det skal ske gennem EU, hvor adgangen til et marked med omkring 450 millioner mennesker er løftestangen.

USA har valgt en anden vej. Den 29. september indgik den amerikanske regering og lederne af de største AI-firmaer en frivillig aftale om sikkerhed ved superintelligens. Den er ikke juridisk bindende og har ingen sanktioner. Luiza Jarovsky peger på uklare begreber, og på at firmaerne selv vælger deres revisorer. I praksis er det selvregulering, og det gør EU's bindende regler til en reel forskel.

Test før udrulning er ingen garanti for, at man kan slukke et system efter et kontroltab. Det er det bedste greb, vi har, så længe det endnu er tidligt.

■ 7. TRE TING AT TAGE MED

- 1. Mål på effekt og tilsyn.** Regeringen vil frigøre mindst 30.000 offentlige årsværk med AI. Tilsyn, klageadgang og dokumenteret effekt skal følge hver krone i accelerationsfonden.
- 2. Påvirk EU nu.** Højrisiko-reglerne i AI-forordningen er udskudt til december 2027 og august 2028. Det er tid til at forberede tilsyn i Danmark og til at stille krav til, hvordan modellerne designs og testes.
- 3. Gør omstilling til politik.** Kompetenceløft for små virksomheder og en sporskiftefond, der følger med tempoet.

Anbefalingerne er mine egne, ikke partiets.

STATUS PR. OKTOBER 2026

AI-forordningens artikel 50 om mærkning og gennemsigtighed gælder fra august 2026. Højrisiko-reglerne er udskudt til december 2027 og august 2028. Truslen fra destruktive cyberangreb er høj (hævet 24. september 2026). Dette er politisk analyse, ikke juridisk rådgivning: bekræft datoer og status mod primære kilder.

SPØRGSMÅL TIL DIG

Hvem i din organisation ville opdage det, hvis jeres systemer blev kortlagt i dag?

■ OM PODCASTEN

Podcastversionen på Substack er AI-genereret ud fra mit eget oplægsmateriale. Stemmerne er syntetiske, og ingen af dem er mig. Den åbner med et af tre fiktive scenarier fra oplægget, uden tekniske detaljer. AI-genereret lyd kan dreje tal og formuleringer. Når podcasten og teksten er uenige, gælder teksten og kilderne.

■ KILDER

Status pr. oktober 2026.

- International AI Safety Report 2026, februar 2026, afsnit 2.1.2, 2.1.4, 2.2.2 og 3.4
- Geoffrey Hinton, tale ved Nobelbanketten, 10. december 2024
- Morris m.fl., Levels of AGI, 2024, rev. 2025, og Hendrycks m.fl., A Definition of AGI, 2025
- Pew Research Center, As Partisan Hostility Grows, 9. august 2022
- Forordning (EU) 2022/2065 (DSA), artikel 34-35
- Forordning (EU) 2024/1689 (AI-forordningen), artikel 50 og 113
- SektorCERT, Angrebet mod dansk, kritisk infrastruktur, november 2023
- Forsvarets Efterretningstjeneste, 18. december 2025
- Styrelsen for Samfundssikkerhed, Truslen fra destruktive cyberangreb hæves til høj, 24. september 2026
- Anthropic, Disrupting the first reported AI-orchestrated cyber espionage campaign, november 2025
- Lambert og Brand, The ATOM Report, april 2026
- Direktiv (EU) 2022/2555 (NIS2)
- Det politiske grundlag for firkløverregeringen, juni 2026, kapitel 12
- Forbes (Sara Dorn) og Luiza Jarovsky, Luiza's Newsletter, begge 30. september 2026

© 2026 Stefano Vincenti · aitrainer.dk · Oprindeligt publiceret på Substack 5. oktober 2026

Stefano Vincenti

AI-træner, konsulent og ekstern lektor · 20+ år i software og digital transformation · Co-founder af BotTellMe · Ekstern lektor ved ITU og DIS Copenhagen · Partner hos TryZone.

LinkedIn

linkedin.com/in/stefanovincenti

Nyhedsbrev

stefanovincenti.substack.com

Mere

aitrainer.dk · tryzone.dk ·
bottellme.com

Disclaimer: Dette er politisk analyse og generel vejledning, ikke juridisk rådgivning, og jeg er ikke advokat eller compliance-specialist. Brug det ikke som beslutningsgrundlag på egen hånd, og bekræft datoer og status mod primære kilder (EUR-Lex, retsinformation.dk, Styrelsen for Samfundssikkerhed). anbefalingerne er mine egne, ikke partiets. Status pr. oktober 2026.

Stefano Vincenti · AI, Built Human