

Sprogmodeller og "kunstig intelligens"



ONLINE VERSION

INFORMATION

INTRODUKTION

Et infodump er Cybernauterne og venners format, hvor vi dumper en masse viden om et digitalt eller teknologisk fænomen på meget lidt plads. Infodumps har derfor også masser af anbefalinger til yderligere læsning, som du kan fortsætte med, hvis du vil vide mere. Dette infodump er lavet af [Cybernauterne](#) og algoritmisk kunstner [Ada Ada Ada](#).

I dette infodump vil vi fortælle jer lidt om “kunstig intelligens” – eller “AI”. Det har vi valgt at gøre ved at zoome ind på den nok mest alment udbredte af disse teknologier, nemlig *store sprogmodeller*, også kendt som LLM'er (Large Language Models). Det gør vi, fordi de udgør kernen af tjenester som ChatGPT, Microsofts Copilot eller Googles Gemini. I resten af infodumpet refererer vi blot til dem som sprogmodeller, men vi snakker her specifikt om den store af slagsen.

Mange er meget begejstrede for det teknologiske fremskridt, de betragter disse tjenester som. Vi oplever dog også en ukritisk og urealistisk idé om, hvad sprogmodeller kan og bør bruges til.

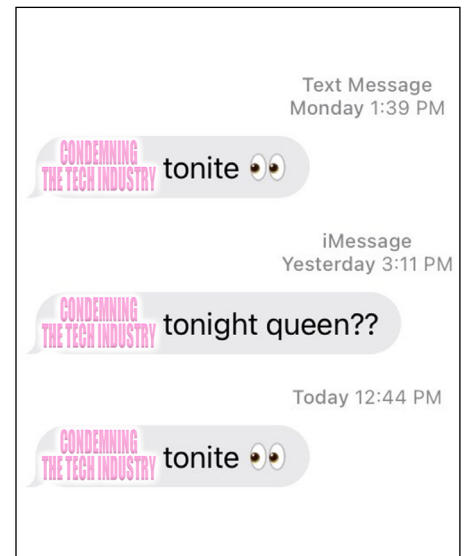
Derfor prøver vi i dette infodump at svare på: Hvad er sprogmodeller, hvordan er de lavet, hvad kan de, hvad kan de *ikke* og... er de egentlig intelligente?

HVAD ER EN SPROGMODEL?

Sprogmodeller er software, der er bygget til at generere eller behandle menneskeligt sprog. De er trænet på gigantiske datasæt, der består af ord og sætninger fra leksika, digitaliseret fag- og-skønlitteratur samt tekst skrabet fra sociale medier som Twitter eller Reddit. Mange datasæt består derudover af ophavsretsligt beskyttet materiale, som ikke nødvendigvis er blevet godkendt til at træne sprogmodeller.¹

Det er vigtigt at pointere, at sprogmodeller er trænet på *sprog* og designet til at simulere en menneskelig samtale. De er altså ikke trænet på *fakta* eller designet til at søge information, efterprøve logik eller afgøre om noget er *faktuelt korrekt*. En sprogmodel er, som udgangspunkt, kun i stand til at gætte hvilken kombination af ord, der er størst sandsynlighed for, at brugeren ønsker at se som svar på sit *prompt*. Sprogmodeller er altså en slags papegøje, der kan kæde ord og sætninger sammen, så det lyder som sprog – men de “forstår” ikke deres betydning. Forskere indenfor lingvistik og kunstig intelligens har derfor kaldt det en *Stochastic Parrot*.²

Dette står i stærk kontrast til den måde “AI” er blevet markedsført af techvirksomheder såsom OpenAI, der står bag ChatGPT. Det faktum, at teknologien brandes som “kunstig intelligens”, antyder, at der er tale om en form for *intelligens* i sprogmodellerne. Det giver en idé om, at programmet *forstår* dit spørgsmål, og at der foregår en art maskinel tankeproces, når den giver dig et svar.



Are you an LLM?

Do you:

- Easily forget what was said earlier in the conversation making you look dumb
- Not really understand what is talked about and instead say something barely appropriate but it doesn't quite fit?
- Feel like you need to use lots of energy and resources to be able to socialize?
- Sometimes remember things that aren't that way or things that didn't happen?

¹ "OpenAI Punches Upward In Bigtime Legal Copyright Lawsuit That Will Surely Determine The Future Longevity Of Generative AI", Lance Eliot, forbes.com, 2024

² Wikipedia: [Stochastic parrot](#), lokaliseret oktober 2024

Der er altså ikke tale om intelligens, men sproglig statistik. Derfor ville en mere korrekt betegnelse for teknologien være “automatiseret statistik”. At kalde sprogmodeller for “kunstig intelligens” er altså mere et spørgsmål om branding og ikke, hvad teknologien reelt er og kan.

Vi, i Cybernauterne, mener, at det er misvisende at omtale sprogmodeller med terminologi, der kan være med til at antyde reel intelligens, men for forståelsens skyld bruger vi betegnelserne i dette infodump – dog med anførselstegn; “AI” og “kunstig intelligens”.

HALLUCINATIONER OG BULLSHIT

Efterhånden som sprogmodeller som ChatGPT eller andre “AI”-drevne chatprogrammer er blevet taget i brug af flere og flere mennesker, er det så småt begyndt at gå op for folk, at man ikke altid kan stole på sandhedsværdien af outputtet.

Det måtte en advokat fra New York sande, da han brugte ChatGPT til at forberede et søgsmål og bad om lignende sager, der kunne danne præcedens.³ ChatGPT svarede ved at opfinde en række sager, der ikke eksisterede i virkeligheden. Dette kom først frem, da modparten påpegede, at der ikke fandtes nogen dokumentation for sagerne. Ifølge advokaten selv, havde han aldrig brugt programmet før og anede ikke, at det kunne finde på at “opfinde sager”.

At ChatGPT genererer misinformation er så udbredt et fænomen, at en forskningsartikel i 2023 kaldte det “digitale hallucinationer”.⁴ Der findes endda et specifikt ord for de akademiske, juridiske og faglige referencer, som sprogmodellerne trækker ud af den blå luft: *hallucinations* – en sammentrækning af hallucinationer og citater.

Dog er det ikke den bedste idé at benytte sig af ordet “hallucination” til at beskrive dette fænomen. At bruge begrebet “hallucinere” understøtter nemlig ideen om, at sprogmodellerne er intelligente og *kan* hallucinere, hvilket fundamentalt fejlrepræsenterer, hvordan teknologien fungerer.

I deres paper “ChatGPT is bullshit” argumenterer forskerne for, at ChatGPT og andre sprogmodeller ikke hallucinerer men i stedet bullshitter i den forstand, at de ikke er designet med sandhed eller fakta for øje.⁵

ChatGPT is not designed to produce true utterances; rather, it is designed to produce text which is indistinguishable from the text produced by humans. Hicks et al., 2024

Dette betyder ikke, at sprogmodeller aldrig gengiver korrekte fakta. Hvis man spørger en sprogmodel om udbredte og alment kendte fænomener, vil der være større sandsynlighed for, at den har det korrekte svar mange gange i sit datasæt. Eftersom statistik er det eneste, som den

Guy Who Sucks At Being A Person Sees Huge Potential In AI

Published June 1, 2023



SAN MATEO, CA—After spending the past three decades of his life being totally unable and unwilling to engage in any meaningful way with the world around him, James Parker, a local guy who sucks at being a person, told reporters Thursday that he saw huge potential in AI. “While it’s still in its early phase,



³ “Lawyer Uses ChatGPT In Federal Court And It Goes Horribly Wrong”, Matt Novak, forbes.com, 2023

⁴ Alkaissi, H., McFarlane, S. [Artificial Hallucinations in ChatGPT: Implications in Scientific Writing](#). Cureus. (2023)

⁵ Hicks, M.T., Humphries, J. & Slater, J. [ChatGPT is bullshit](#). *Ethics Inf Technol* 26, 38 (2024)

arbejder ud fra, vil sprogmodellen derfor ofte vende tilbage med det rigtige svar på spørgsmål som "hvornår startede 2. Verdenskrig?" eller "hvad hedder kongen af Danmark?". Problemet opstår, når man spørger ind til mindre kendte ting, begivenheder eller personer. I dette tilfælde er der ikke nødvendigvis nok dokumentation i datasættet, så sprogmodellen vil ofte finde på ting, der lyder sandsynlige, men som er langt fra sandheden.

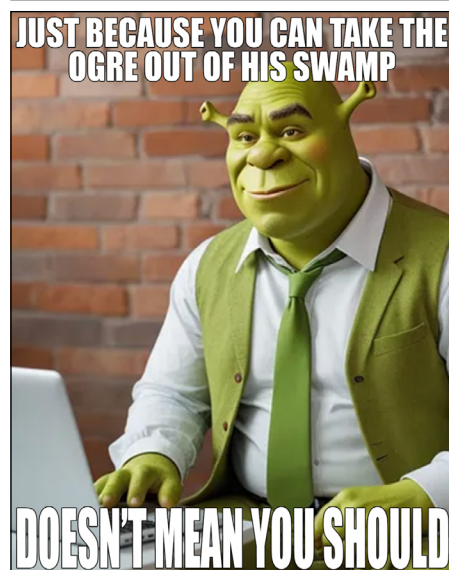
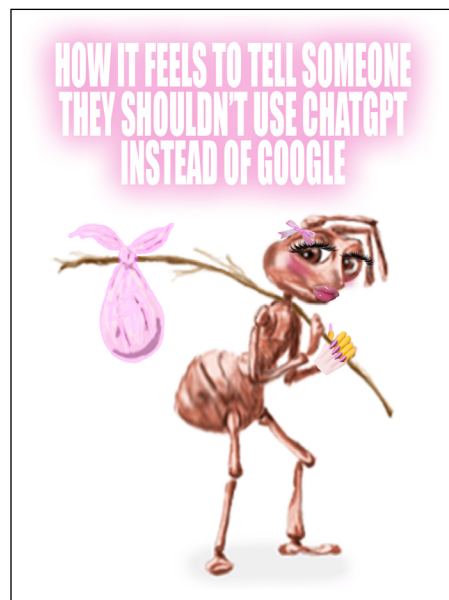
Et stærkt underholdende eksempel er Googles "AI"-søgefunktion Gemini, der blandt andet er trænet på datasæt fra Reddit. Derfor har den blandt andet anbefalet brugere at putte lim på deres pizza for at gøre osten mere klistret og fortalt, at det er sundt at spise en lille sten om dagen.⁶ Det var nemlig svar, der eksisterede i deres datasæt, men som folk har skrevet for sjov. Og så er der historien om en "AI"-chatbot, der rickrollede en kunde, der bad om en instruktionsvideo.⁷ Dette skete, fordi der er mange eksempler i træningsdatasættet med folk, der har sendt et link til Rick Astley's musikvideo "Never Gonna Give You Up" til folk som en joke. Sprogmodeller kan altså kun generere sprogligt overbevisende outputs men ikke afgøre kontekst eller sandhedsværdien af det.

DO'S AND DON'TS: HVAD KAN DU BRUGE DEM TIL?

Når folk, som den amerikanske advokat, alligevel bruger sprogmodeller til at søge viden om alt fra arbejde til sundhed, skyldes det, at sprogmodellerne er designet til at "svare" med største selvfølgelighed og overbevisning. Dertil kommer formen, der minder om en chat-samtale og giver indtrykket af et talende, tænkende væsen på den anden side, hvilket styrker troværdigheden.

Nogle af de problematiske aspekter ved brug af sprogmodeller til at op søge viden er derfor risikoen for misinformation. Et par eksempler er:

- At søge viden om sundhed og helbred kan resultere i direkte farlige eller dødelige råd. Som f.eks. svampeguides skrevet af sprogmodeller, der påstår at giftige svampe er spiselige.⁸
- At søge råd og vejledning om juridiske processer og lovgivning kan resultere i ulovlige råd, som da New York bys "AI"-chatbot gav udlejere vejledning, der går imod loven.⁹
- At skrive kode: Der er ofte fejl i, og det kan tage længere tid at rette, end hvis man havde skrevet det fra bunden.
- At udføre matematiske udregninger: Sprogmodeller kan ikke lave regnestykker. De kan kun generere sprog og laver derfor som regel fejl i spørgsmål, der handler om mere kompleks logik.
- Søgninger om information, når du egentlig har brug for et leksikon eller en fagbog: Sprogmodeller laver mange fejl og opfinder f.eks. bibliografier eller blander mindre kendte mennesker sammen.¹⁰



⁶ ["Google Is Paying Reddit \\$60 Million for Fucksmith to Tell Its Users to Eat Glue"](#), Jason Koebler, 404media.co, 2024

⁷ [@lazerwalker på x.com](#)



⁸ ["The AI grift that can literally poison you"](#), Constance Grady, vox.com, 2024

⁹ ["NYC's AI chatbot was caught telling businesses to break the law. The city isn't taking it down"](#), Jake Offenhardt, apcnews.com, 2024

¹⁰ [@cridhe_ på tiktok.com, the dangers of AI generated outdoors websites](#), 2024



Generelt er det risikabelt at bruge sprogmodeller til information, man ikke kan faktatjekke selv. Man skal gå til dem, som man går til en fremmed person på internettet. Der er en vis sandsynlighed for, at det passer, men du har ingen garantier, og du kan ikke stole på, at det, du får at vide, er sandt.

Det som sprogmodeller kan bidrage til er beskedent – omend brugbart. De kan fx bruges til:

- Opsummering af tekster
- Omformulere en sætning
- Finde stavfejl
- Lave simple oversættelser
- Skrive meget formulariske tekster

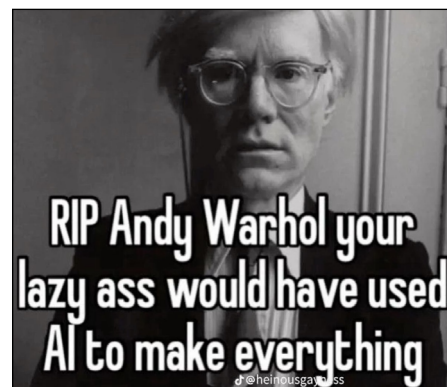
“AI”-HYPECYKLUSSEN

Der har været meget hype omkring “AI”, især at det skulle revolutionere arbejdsmarkedet og øge produktiviteten. Det er dog endnu ikke bevist, hvorvidt “AI” skulle give en stor produktivitetstigning. I en undersøgelse fra The Upwork Research Institute siger 96% af direktørerne i de adspurgte virksomheder, at de regner med at “AI” øger produktiviteten.¹¹ Samtidig siger 77% af de adspurgte ansatte, at “AI”-værktøjerne faktisk har gjort dem *mindre*, ikke *mere*, produktive. Ansatte og chefer har altså meget forskellige idéer om, hvad “AI” er i stand til at gøre.¹²

Hypen omkring “kunstig intelligens” er i høj grad et produkt af techindustriens evne til at hype en bestemt ny innovation og dermed sikre massive investeringer. Tidligere eksempler på disse hypecyklusser er *blockchain*, *kryptovaluta* og *metaverset*. Det, vi i dag kender som “kunstig intelligens”, var tidligere betegnet som neurale netværk, maskinlæring og deep learning. At brande det som “AI” har fået interessen – og investorkapitalen – til at stige til astronomiske højder, men har ikke så meget at gøre med, hvad teknologien reelt kan eller er på vej til at blive.

Et af problemerne med samtalen omkring “kunstig intelligens” er, hvor ofte teknologien bliver tillagt magt og evner, der ikke er evidens for, at den har. “Kunstig intelligens” er ikke klogere end mennesker, og gør for det meste et langt dårligere stykke arbejde.¹³ Der er ikke grund til at tro, at den om føje tid vil tage alle vores jobs eller overtage magten over menneskeheden og gøre os til slaver a la The Matrix. Der er heller ikke grund til at tro, at det vil revolutionere verden.

Investeringerne i “AI”-relaterede virksomheder og startups har nået astronomiske niveauer.^{14 15} Men i andet halvår af 2024 er flere investorer begyndt at få kolde fødder.¹⁶ Det skyldes blandt andet, at teknologien ikke indfrier forventningerne til indtjening, hvilket har fået flere analytikere til at tale om “AI”-optimismen som en boble, der er i risiko for at briste.¹⁷



premierls.dk
Sponsored · Expand story

IS LAVET MED AI

Udviklet med KUNSTIG INTELLIGENS og 90 års egne erfaring

Premier Ice
AI SCREAM
SCREAMING GOOD!

NYHED: Prøv den nye AI SCREAM, med smag af cremet grøn bananis med forfriskende blåbærsaube, chokoladeovertrukne majs-knager og sprøde chokoladestykker.

LEARN MORE

Katherine
@mycorporatebestie

"chat GPT is going to take over your job!"
me: take it

¹¹ ["From Burnout to Balance: AI-Enhanced Work Models"](#), Kelly Monahan og Gabby Burlacu, upwork.com, 2024

¹² ["77% Of Employees Report AI Has Increased Workloads And Hampered Productivity, Study Finds"](#), Bryan Robinson, forbes.com, 2024

¹³ ["10 reasons why AI may be overrated"](#), Greg Rosalsky, npr.org, 2024

¹⁴ ["AI Is the Hot Topic on Earnings Calls This Quarter"](#), Jennifer Ryan, bloomberg.com, 2024

¹⁵ ["Four-week-old AI start-up raises record €105mn in European push"](#), Tim Bradshaw og Leila Abboud, ft.com, 2023

¹⁶ ["Stock market has worst day since 2022 as Tesla, Google parent Alphabet sink"](#), Rob Wile, nbcnews.com, 2024

¹⁷ ["Is The AI Bubble About To Burst?"](#), Bernard Marr, forbes.com, 2024

EN KLIMASYNDER

En lidt bredere betegnelse for "AI"-teknologier, der genererer indhold er "generativ AI". Dette dækker over sprogmodeller, men også lyd-, billed- og videogenererende teknologier som ElevenLabs, Stable Diffusion og Runway. "AI"-software i denne kategori har én meget håndgribelig påvirkning på verden til fælles: De har et massivt klima- og miljøaftryk.

Der er bred konsensus, om at "generativ AI" bruger utroligt meget strøm. Ifølge Goldman Sachs er en informationssøgning med sprogmodeller 10 gange så energikrævende som en almindelig Google søgning.¹⁸ I dansk kontekst har forskere fra Datalogisk Institut anslået, at ét prompt med ChatGPT bruger ligeså meget strøm som 40 mobilopladninger.¹⁹ Derfor kan det at bruge en sprogmodel til at rette sine mails for stavfejl, hurtigt ende med at bruge strøm nok til at lade en telefon op i et år.

"Kunstig intelligens" bruger også utroligt meget vand blandt andet for at nedkøle de datacentre, der driver processerne. Ifølge et studie fra OECD vil "kunstig intelligens"-teknologier i 2027 bruge 4-6 gange så meget vand som Danmark bruger årligt.²⁰ Et nyt studie fra University of California estimerer ChatGPT bruger to liter vand på 10-50 prompts.²¹

De massive energiomkostninger ved "AI" har endda fået Googles CO²-udledninger til at stige med 48% de seneste fem år, hvilket sår tvivl om, hvorvidt tech-giganten kan nå sit erklærede mål om at reducere sine udledninger til nul i år 2030.²²

"Generativ AI" er ikke kun forslugen, når det gælder strøm og vand. Der er også et stigende behov for ny data til at forbedre "AI"-modeller, og samtidig ser vi en forringelse af kårerne for de mennesker, hvis arbejde det er at rense, bearbejde og gøre al den data brugbar.^{23 24}

VI DRUKNER I HABSBERG AI

Det indhold vi ser på internettet er i stigende grad produceret af "generativ AI". ChatGPT og alle de andre sprogmodeller har gjort det nemt at producere spam, og vi ser derfor mere og mere "AI"-genereret bullshit i vores internetsøgninger, sociale medier og indbakker. Eftersom alle sprogmodeller er trænet på data skrabet fra internettet, vil videreudviklingen af dem kræve, at man skraber mere data. Problemet er bare, at man ikke kan træne sprogmodeller på data, der er produceret af sprogmodeller. Dette resulterer nemlig i at sprogmodellerne degraderer og mister evnen til at skabe overbevisende tekst. Dette fænomen har podcasteren Jathan Sadowski kaldt for *Habsburg AI*, opkaldt efter det adelige dynasti, der blev berygtet for at kollapse på grund af indavl.²⁵ Dette sker samtidig med, at flere og flere techvirksomheder burer deres data inde bag betalingsmure, så konkurrenterne ikke "stjæler det" til deres sprogmodeller.²⁶



¹⁸ ["AI brings soaring emissions for Google and Microsoft, a major contributor to climate change"](#), Dara Kerr, npr.org, 2024

¹⁹ ["ChatGPT er en forslugen robot"](#), Ole Hoff-Lund, ida.dk, 2023

²⁰ ["How much water does AI consume? The public deserves to know"](#), Shaolei Ren, oecd.ai, 2023

²¹ ["Thirsty ChatGPT uses four times more water than previously thought"](#), Mark Sellman og Adam Vaughan, 2024

²² ["Google emissions jump nearly 50% over five years as AI use surges"](#), Camilla Hodgson og Stephen Morris, ft.com, 2024

²³ ["For Data-Guzzling AI Companies, the Internet Is Too Small"](#), Deepa Seetharaman, wsj.com, 2024

²⁴ ["Exclusive: OpenAI Used Kenyan Workers on Less Than \\$2 Per Hour to Make ChatGPT Less Toxic"](#), Billy Perrigo, time.com, 2023

²⁵ ["AI Appears to Be Slowly Killing Itself"](#), Maggie Harrison Dupré, futurism.com, 2024

²⁶ ["Twitter replaces its free API with a paid tier in quest to make more money"](#), Jess Weatherbed, theverge.com, 2023

En af de reelle trusler fra "kunstig intelligens" er altså, at det risikerer at undergrave kvaliteten og sandhedsværdien af det materiale, vi kan finde online, mere og mere – samtidig med at der lukkes ned for forskere og journalisters adgang til at indhente og analysere data fra digitale platforme.²⁷

EN DEHUMANISERENDE TEKNOLOGI

De hjælpsomme kvaliteter, som sprogmodeller tilbyder, bør ikke overskygge de fejl og mangler, som de også har. De burde ikke få os til at se bort fra de etiske, moralske, klimamæssige, miljømæssige og praktiske problemer som den store, ukritiske udbredelse af sprogmodeller bringer med sig.

Det er værd at stoppe op og spørge om vi ønsker at bruge "AI" til alle de her ting, bare fordi vi kan. Hvis du bruger "AI" til at skrive en mail og din modtager får "AI" til at læse den, kan vi så ikke droppe den mail i det hele taget? Hvis "AI" kan opsummere dit møde og lave referat, kunne mødet så droppes til fordel for en tekst? Man kan sige, at "AI" automatiserer arbejdsprocesser, men det er også nemt at forestille sig hvor meget ekstra pseudoarbejde, det kan generere. Der er lige nu meget lidt, der tyder på, at "AI" kan indfri de håb og drømme om produktivitet, som tech- og finanssektoren har hængt deres vækstdrømme op på.²⁸

Science Fiction forfatteren Ted Chiang har kaldt "AI" for en *fundamentalt dehumaniserende teknologi*, fordi det i hans øjne reducerer os til at blive mindre end det, vi egentlig er: Væsener der kan skabe og forstå mening.²⁹ Vi efterlader dig derfor med en sidste overvejelse:

Hvis generativ "AI" er fremtiden, så er det en fremtid, hvor vi fortsat skal vaske tøj, kede os i møder og støvsuge, mens maskinen bliver sat til at skabe kunst, kreere musik og skrive digte. Er det den fremtid, vi ønsker?

Vi håber, at de perspektiver der er delt i dette infodump kan bidrage til den samtale.[‡]

Cybernauterne & Ada Ada Ada

Cybernauterne er et netværk af eksperter i cybersikkerhed og digitale medier som tilbyder kurser, rådgivning og foredrag.

Vi fandt sammen i 2018 for at forsøge at fylde det hul, vi så i mediebildet omkring formidling af internetkultur og retligheder. Siden da har vi afholdt hundredevis af workshops og foredrag for skoler, NGO'er, fagforeninger, SSP og politigrupper, rettighedsorganisationer og virksomheder.

Vi er specialiseret i at formidle til folk, der ikke er it-nørder eller har specielle tekniske færdigheder. Vi bruger inddragelse, kreative metoder og åbne workshop-formater til at indføre folk i alt fra falske nyheder, privatliv på nettet, algoritmernes magt, omgangsform på de sociale medier og sikkerhedstips til at beskytte sig mod hackere og it-kriminelle. Fælles for os er en kærlighed til teknologi kombineret med en kritisk tilgang til den digitale udvikling.



²⁷ ["Google cut a deal with Reddit for AI training data"](#), Emma Roth, theverge.com, 2024

²⁸ Wikipedia: [Pseudoarbejde](#), lokaliseret oktober 2024

²⁹ ["Why AI Isn't Going to Make Art"](#), Ted Chiang, newyorker.com, 2024

[‡] Er du interesseret i at høre mere? Så lyt til vores episoder om "AI" i vores podcast *Cybernormer* hvorend, du lytter til podcast. Vi anbefaler:
→ "Cybernormer #26 FOMO for milliarder"
→ "Cybernormer #17 Almageddon"
→ "Algotratiet #1: Når din sagsbehandler er en kunstig intelligens"
↓ Find dem [her](#)



Du kan støtte vores arbejde via patreon.com/cybernauterne →

Du kan finde os her ↓
Web: cybernauterne.dk
Podcast: [Cybernormer](#)
Instagram: [@cybernauterne](#)
TikTok: [@cybernauterne](#)

