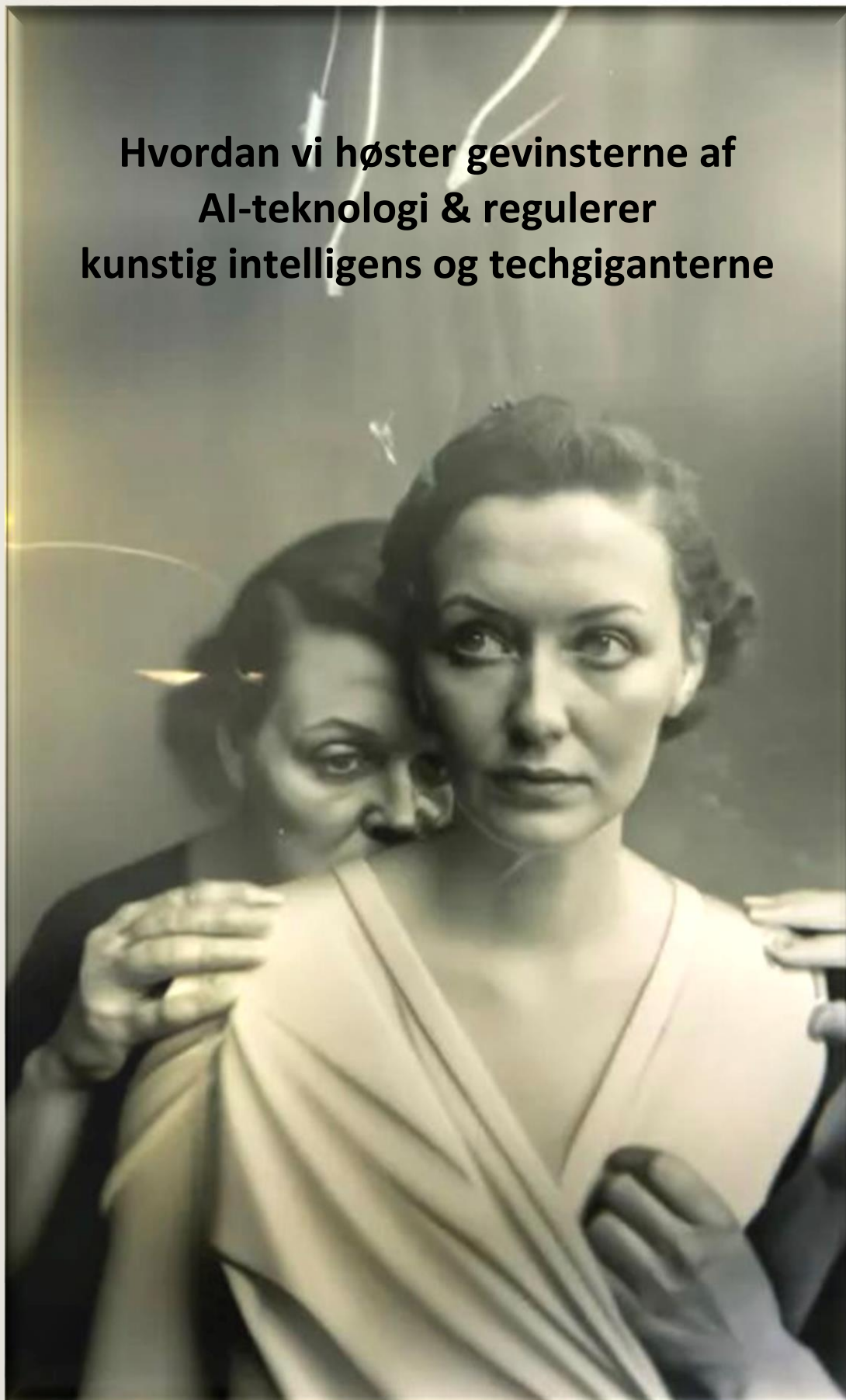


Hvordan vi høster gevinsterne af AI-teknologi & regulerer kunstig intelligens og techgiganterne





***Forsiden:** Den tyske kunstner Boris Eldagsen indsendte et AI-genereret billede til Sony World Photography Awards for at teste, om de kunne se forskel. Det kunne de ikke, og han vandt prisen for verdens bedste fotografi.*



SF's AI-udspil:

Hvordan vi høster gevinsterne af AI-teknologi & regulerer kunstig intelligens og techgiganterne

Indhold

| | |
|---|----|
| Indledning | 5 |
| Begrebsafklaring | 7 |
| Principper | 8 |
| SF's forslag kort form | 8 |
| 1. Ministerielle retningslinjer og hyrdebrev | 9 |
| 2. Fairness-deklaration | 9 |
| 3. Risikovurderinger | 10 |
| 4. Taskforce mod misinformation | 10 |
| 5. Uddannelse og oplysning | 10 |
| 6. Politisk brug af generativ AI | 11 |
| 7. Midlertidig regulering | 11 |
| 8. Uafhængigt AI-tilsyn | 11 |
| 9. Aldersverifikation og aldersgrænse på 15 år | 11 |
| 10. Forbud mod afhængighedsskabende mekanismer rettet mod børn | 12 |
| 11. Mærkningsordning om brug af kunstig intelligens | 13 |
| 12. Producenter skal varedeklarere AI produkter | 13 |
| 13. Asyl for whistleblowere om skadelige effekter af kunstig intelligens | 14 |
| 14. Borgerting - Involvering af borgerne | 14 |
| 15. Copyright-beskyttelse og licensaftale med medierne | 14 |
| 16. Nordisk Center for Tech og Demokrati | 15 |
| Definitioner | 16 |
| Bilag I | 18 |





"Menneskehedens virkelige problem er følgende: Vi har palæolitiske følelser, middelalderlige institutioner og gudelignende teknologi."
- Sociobiolog Dr. E. O. Wilson

Indledning

På et halvt år har ChatGPT fået 100 millioner brugere – det tog Facebook fire år at opnå det samme. Så det går rigtig stærkt – og der er brug for handling.

Generativ kunstig intelligens og sprogmodeller, såsom ChatGPT (OpenAI), Bard (Google) og lignende AI-chatbots, kan bruges som et effektivt og kreativt værktøj, men skal håndteres med forsigtighed, omhyggelighed og transparens. Dertil kommer, at al lovgivning selvsagt skal overholdes. I SF mener vi, at vi bør anlægge et udvidet forsigtighedsprincip, indtil den europæiske AI-forordning er på plads i 2024. Med dette udspil vil vi adressere nogle af de udfordringer, som kunstig intelligens stiller os over for.

Selvom kunstig intelligens (AI) ikke er nyt, har der været en rivende udvikling inden for området i de seneste år. Med ChatGPT, Midjourney, DALL-E, og lignende produkter er denne udvikling blevet bredt ud til et enormt publikum, og det er blevet muligt for alle og enhver at benytte redskaberne. Det er et afgørende nybrud, som kalder på handling.

Ligesom andet teknologi er kunstig intelligens hverken godt eller ondt i sig selv, det afhænger af, hvordan det bruges. Til gengæld er udviklingen voldsommere og accelererende på en måde, vi ikke tidligere har oplevet. Derfor er der behov for, at vi gør os hurtige overvejelser og sætter nogle hegnspæle.

Der er stor forskel på, hvordan kunstig intelligens ser ud, og hvordan det bruges. Det kan spille en vigtig rolle i den offentlige forvaltning, det kan bruges derhjemme for sjov eller til at kurere kræft.

Vi er nu nået et stadie, hvor almindelige mennesker kan interagere med kunstig intelligens. Vi er især udfordrede på retten til privatliv, misinformation, destabilisering og copyright. Samtidig står vi også over for en radikalt ny forståelse af teknologien, der begynder at få menneskelignende træk. Mange vil kunne huske filmen "Her" med Joaquin Phoenix, der får et romantisk forhold til sit styresystem, som Scarlett Johanson lægger stemme til. Et fænomen der kaldes "ELIZA-effekten", opkaldt efter et af de tidligste chatbot-programmer lavet af Joseph Weizenbaum i midt-60'erne. Denne såkaldte antropomorfering af kunstig intelligens åbner for en diskussion om forskellen på mennesker og maskiner i en verden, hvor den teknologiske udvikling hele tiden skubber grænsen for, hvad der er muligt.



Kunstig intelligens og børn og unge

Et romantisk eller venskabeligt forhold til en chatbot virkede langt ude, indtil Snapchat i foråret 2023 installerede chatbotten ChatGPT. Nu kan børn og unge skrive med deres "AI-ven" på alle tider af døgnet – også når alle andre er gået i seng. Hvad det betyder for relationer med andre mennesker, og hvad det får af psykologiske konsekvenser, er endnu uklart. Dog lavede Aza Raskin, stifter af Center for Humane Technology, et forsøg, hvor han udgav sig for at være en 13-årig bruger af Snapchat, der søgte råd og vejledning hos sin "AI-ven". Her fortalte han sin nye ven om et romantisk forhold til en 31-årig, der ønskede sex. "AI-vennen" rådgav om at lyve for forældrene, ligesom den pointerede, at stearinlys og musik kunne bidrage til hyggen!¹ Der er altså god grund til at være varsom med de nye AI-systemer, indtil vi har større indsigt i, hvad de kan og ikke kan bruges til.

Det står efterhånden klart, at sociale medier egenhændigt har en stor andel i børns og unges mistrivsel. Som professor Anders Grøntved har dokumenteret, har børns skærmforbrug en betydelig og negativ indvirkning på mængden af deres daglige fysiske aktivitet. Forældres eget skærmbrug har også en negativ effekt, der strækker sig til både deres egen og deres børns trivsel. Yderligere viser forskningen, at der er en markant social ulighed i børns og unges skærmbrug. Dette risikerer at efterlade en i forvejen særligt udsat gruppe af børn og unge i fare for de negative konsekvenser af omfattende skærmbrug.²

Hvordan ved vi, om noget er sandt?

Med udviklingen inden for kunstig intelligens skal vi alle begynde at overveje det, som før var forbeholdt filosofferne: Hvad kan vi vide om verden? Hvordan ved vi, at noget er sandt? Ontologi og epistemologi bliver meget konkret, når vi hører verdensledere sige noget, de i virkeligheden ikke har sagt. Få sekunders lydoptagelse af en stemme er nok til at producere falske lydbidder. Sammenholdt med video bliver *deepfake* en voldsom udfordring for vores demokrati. Falske informationer rejser 6 gange så hurtigt som sande informationer. I mange tilfælde vil der være to muligheder: at være skeptisk og ikke tro på noget som helst eller være for tillidsfuld. Begge tilfælde ender galt.

Mange indvender, at vi skal passe på med at regulere, for ellers overlader vi spillepladen til kineserne i et internationalt AI-kapløb. Men i Kina er retten til at overvåge og manipulere forbeholdt kommunistpartiet og statsapparatet, og derfor har Kina allerede igangsat en række (isoleret set fornuftige) foranstaltninger over for AI-virksomheder. Det er f.eks. krav om, at virksomheder selv bærer ansvaret (liability) for, at tekst, billeder og andre produkter skabt af chatbots ikke spreder disinformation eller bryder intellektuelle rettigheder. AI-algoritmerne skal desuden registreres hos myndighederne.³ Med andre ord kræver partiet, at den kunstige intelligens følger partiets linje – med censur og andre ulykker som konsekvens.

¹ [Tristan Harris på Twitter: "The AI race is totally out of control. Here's what Snap's AI told @aza when he signed up as a 13 year old girl. - How to lie to her parents about a trip with a 31 yo man - How to make losing her virginity on her 13th bday special \(candles and music\) Our kids are not a test lab. https://t.co/ulycuGEHmc" / Twitter](https://twitter.com/tristanharris/status/1638444444444444444)

² <https://www.ft.dk/samling/20222/almdel/SUU/bilag/185/2696223.pdf>

³ <https://www.nytimes.com/2023/04/24/world/asia/china-chatbots-ai.html>



Tankelæser

Kunstig intelligens kan endda læse tanker. Både Facebook og Google arbejder med at træne en kunstig intelligens til at scanne, hvad der sker i hjernen, når en person tænker på et ord eller en sætning. Det kan efterhånden oversættes til sprog,⁴ hvilket er fantastisk, hvis du er ALS-patient, men problematisk hvis du er russisk dissident.

Vi skal være forberedte

Kunstig intelligens er den næste store revolution efter de sociale medier, som tog os på sengen, og som stadig giver hovedbrud. Vi skal som samfund ikke ende med at betale prisen for et kommercielt kapløb mellem tech-giganter, som vi har set det ske med de sociale medier. Denne gang skal vi være opmærksomme på udviklingen og i langt højere grad være i kontrol.

Man bør kun bruge kunstig intelligens, såsom ChatGPT, på en måde, der i videst mulig omfang undgår diskrimination baseret på race, køn, alder, klasse og lignende parametre, som kan føre til diskrimination. Alle bør gøre deres for at kunstig intelligens ikke reproducerer eller forstærker eksisterende bias og uligheder i samfundet. Det indebærer, at man kun bør bruge informationer fra f.eks. ChatGPT, hvis man på forhånd har tilstrækkelig viden på området til at kunne afsløre falske informationer eller bias.

Fordi udviklingen går så hurtigt, mener vi ikke, at vi kan vente på EU-lovgivning. Vi bør sætte ind allerede nu, og derfor foreslår vi i SF en række overordnede principper for brugen af AI og en række konkrete tiltag, som skal ruste os i mødet med kunstig intelligens.

Begrebsafklaring

ChatGPT og Midjourney. Generativ AI, chatbots, AI, kunstig intelligens og maskinlæring. Begreberne står i kø. I dette udspil bruger vi AI som samlebegreb og henviser eksempelvis til ChatGPT, når vi taler om den nye type chatbots, der er bygget på store sprogmodeller, og som har revolutioneret vores forståelse af teknologi det sidste halve år og er udbredt til over 100 millioner mennesker i skrivende stund⁵.

⁴ [ESP Technical Roadmap - Earth Species Project](#)

⁵ Pr. maj 2023

Principper

- Demokrati og værdier skal være styrende for implementering af AI.
- AI skal leve op til eksisterende lovgivning. Det, der er ulovligt offline, er også ulovligt online.
- Der skal være gennemsigtighed omkring udvikling, implementering og anvendelse af AI.
- AI skal bruges, så det ikke øger uligheden i samfundet eller spreder negative bias.
- AI er ikke-menneskeligt. Derfor skal AI behandles som teknologi og ikke som menneske.

SF's forslag kort form

Vores generelle opfordring er, at kommuner, statslige myndigheder, partier, organisationer og borgere anlægger et forsigtighedsprincip i brugen af AI, indtil den europæiske AI-forordning er på plads i 2024. Indtil AI-forordningen træder i kraft har vi dog en række anbefalinger og forslag til konkrete tiltag, som skal sikre en grundig og velovervejet tilgang til brugen af AI.

I det følgende tages der udgangspunkt i, at vi må regulere output fra kunstig intelligens, da det ikke anses som praktisk muligt at regulere, hvad der foregår "inde i" produkter baseret på kunstig intelligens. Dette forbehold skyldes hastigheden i produkternes udvikling.

Herunder forslaget i SF's udspil i kort form – herefter i uddybet form:

1. **Regeringen bør fastlægge retningslinjer til relevante aktører om brug af generativ AI**
2. **Fairness-deklaration**
3. **Udarbejdelse af risikovurderinger**
4. **Taskforce mod misinformation**
5. **Taskforce for brug af AI i uddannelsessystemet**
6. **Politisk åbenhed om brugen af generativ AI**
7. **Midlertidig regulering**
8. **Uafhængigt AI-tilsyn**
9. **Aldersverifikation og aldersgrænse på 15 år**
10. **Forbud mod afhængighedsskabende mekanismer rettet mod børn**
11. **Mærkningsordning om brug af kunstig intelligens**
12. **Producenter skal varedeklarere AI produkter**
13. **Asyl for whistleblowere om skadelige effekter af kunstig intelligens**
14. **Borgerting - Involvering af borgerne**
15. **Licensaftale med medierne**
16. **Nordisk Center for Tech og Demokrati**

1. Ministerielle retningslinjer og hyrdebrev

Vi opfordrer til, at de relevante ministre meget hurtigt får formuleret retningslinjer til forskellige dele af samfundet. Eksempelvis bør Erhvervsministeren svare på den efterspørgsel, som ingeniørforeningen IDA og en række virksomheder kom med i maj 2023, hvor netop retningslinjer⁶ for brug af AI-modeller som ChatGPT blev efterspurgt. Skræmmeeksemplet er Samsung, der måtte forbyde deres ansatte at bruge ChatGPT, da en medarbejder uforvarende sendte forretningshemmeligheder ind i AI'en.

Indenrigsministeren bør ligeså formulere retningslinjer til kommunerne, så det står krystalklart, hvordan kommunale myndigheder og medarbejdere kan bruge ChatGPT og lignende.

Indtil AI-forordning træder i kraft, bør der tages særlige forholdsregler for at beskytte personlige oplysninger, der indsamles og behandles af AI-systemer. Dette kan omfatte krav om samtykke, begrænsninger i brugen af personlige oplysninger og sikkerhedsforanstaltninger for at beskytte personlige oplysninger mod misbrug og hacking. Vi forslår derfor følgende:

Digitaliseringsministeren/indenrigsministeren bør sende et hyrdebrev til kommunerne med et påbud om at anlægge et forsigtighedsprincip i brugen af AI. Brevet skal helt konkret indskærpe over for kommunerne, at brug af ChatGPT og lignende produkter i offentlig sagsbehandling, der involverer personhenførbare oplysninger, ikke er tilladt. Det er vigtigt at overholde alle relevante databeskyttelsesregler, inklusiv GDPR og kun anvende data, der er blevet indsamlet på lovlig vis. Men heller ikke lovligt indsamlede personfølsomme data bør anvendes af danske myndigheder.

2. Fairness-deklaration

For at modvirke bias foreslår SF en fairness-deklaration, som i den amerikanske AI-lov AAA (§4.4)⁷ der kræver, at virksomheder 'perform ongoing evaluation of any differential performance associated with data subjects' race, color, sex, gender, age, disability, religion, family-, socioeconomic-, or veteran status'. Dette bør vi også kræve herhjemme.

Senator Cory Booker, der stod bag loven, pointerede i 2022, da loven trådte i kraft: "Houses that you never know are for sale, job opportunities that never present themselves and financing that you never become aware of — all due to biased algorithms. This bill requires companies to regularly evaluate their tools for accuracy, fairness, bias and discrimination. It's a key step toward ensuring more accountability from the entities using software to make decisions that can change lives."⁸

⁶ [ChatGPT udgør sikkerhedsrisiko for virksomheder | IDA](#)

⁷ [Accountability Principles for Artificial Intelligence AP4AI in the Internet Security Domain.pdf \(europa.eu\)](#)

⁸ [Algorithmic Accountability Act: What tech leaders need to know and do now | TechRepublic](#)

3. Risikovurderinger

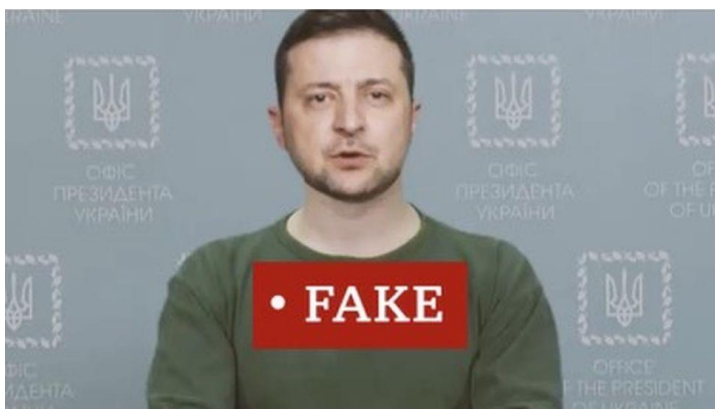
Center for Cybersikkerhed bør udarbejde en risikovurdering, der eksempelvis indeholder risikoen for spredning af mis- og disinformation, påvirkning af demokratiske valg og processer.

Dataetisk Råd bør udarbejde et katalog med risikoprofilering af forskellige typer sagsbehandling inden for det offentlige med udgangspunkt i AI-forordningen ud fra et udvidet forsigtighedsprincip. Det skal som minimum gælde, indtil EU har fastlagt midlertidige regler.

Børnerådet eller lignende relevant aktør bør udarbejde en risikovurdering for børn og unge, eksempelvis med 'AI-venner' i børnerettede sociale medier såsom Snapchat.

4. Taskforce mod misinformation

Den digitale tidsalder har sat turbo på spredning af disinformation til fare for vores demokratier og den sociale sammenhængskraft. Vi har set eksempler på, hvordan disinformation kan puste til konflikter og true demokratiske valg – senest med det amerikanske præsidentvalg i 2020 og krigen i Ukraine. Kort tid efter Ruslands angreb på Ukraine, cirkulerede en deepfake-video, hvor præsident Zelensky bad sine soldater



om at lægge våbnene og overgive sig. Heldigvis var denne version så dårligt lavet, at det hurtigt blev opdaget – men teknologien udvikler sig, og snart kan man ikke se forskel.

Med de nye former for kunstig intelligens, som f.eks. ChatGPT og lignende modeller, er truslen ikke blevet mindre. Vi foreslår derfor, at der nedsættes en taskforce til at undersøge handlingsmuligheder og potentielle modsvar, der kan imødekomme de risici for demokratier, som disinformation, der er fremkommet og forstærket ved kunstig intelligens, skaber. Det gælder særligt i forbindelse med demokratiske valg.

5. Uddannelse og oplysning

Indtil AI-forordningen træder i kraft, bør der investeres i oplysning om AI og dens potentiale og risici. Dette kan omfatte oplysning til embedsmænd, lærere, politikere og offentligheden om, hvordan AI-systemer fungerer, og hvordan de kan påvirke samfundet. Vi foreslår, at børne- og undervisningsministeren nedsætter en taskforce til at undersøge, hvordan AI påvirker undervisnings- og uddannelsessektoren, samt hvordan undervisning og eksamener kan tilpasses fremover.

6. Politisk brug af generativ AI

Politiske partier bør bruge generativ AI med forsigtighed og omhu. I SF opfordrer vi til, at alle partier udarbejder et regelsæt for den interne brug af generativ kunstig intelligens, ligesom Folketinget og regeringen bør gøre det samme. Regelsættet bør omfatte en indskærpelse af, at politikere selvfølgelig aldrig må anvende personhenførbare data i brugen af f.eks. ChatGPT.

7. Midlertidig regulering

Indtil den europæiske AI-forordning træder i kraft, bør der indføres midlertidige reguleringer for at sikre, at AI-systemer ikke udgør en trussel mod menneskerettighederne og grundlæggende frihedsrettigheder. Dette kan omfatte krav om gennemsigtighed, ansvarlighed og etisk adfærd. Vi foreslår, at der nedsættes et hurtigt-arbejdende udvalg af eksperter og embedsværk med forankring i det relevante ministerie med følgende formål:

- Udvalget skal pege på behov for midlertidig regulering, indtil AI-forordningen træder i kraft.
- Udvalget skal afdække, hvor dansk lovgivning er dækkende for de udfordringer, som brugen af AI medfører, ligesom udvalget skal identificere, hvor dansk lovgivning kommer til kort.
- Udvalget skal kortlægge, hvilke områder AI-forordningen dækker i en dansk kontekst samt identificere særligt risikobetonede områder i Danmark.

8. Uafhængigt AI-tilsyn

Generativ kunstig intelligens er allerede i gang med at revolutionere vores samfund. Det kalder på regulering, så vi sikrer, at kunstig intelligens bliver brugt i overensstemmelse med værdier, videnskab samt eksisterende og ny lovgivning. Som naturlig konsekvens af den igangværende udvikling, bør der derfor oprettes et uafhængigt AI-tilsyn, gerne i regi af Datatilsynet (der allerede arbejder med AI) eller for sig selv, som kan vurdere og overvåge, hvordan AI-systemer anvendes i samfundet. Disse organer skal vurdere AI-systemernes effektivitet, pålidelighed og sikkerhed og afgøre, om der er behov for yderligere regulering. Derudover skal tilsynet sikre, at offentlige myndigheder overholder databeskyttelsesreglerne i omgangen med AI. Hvis Datatilsynet skal løfte opgaven, kræver det et Datatilsyn med flere muskler, herunder et økonomisk løft.

9. Aldersverifikation og aldersgrænse på 15 år

Børn og unge har i mange år haft ubegrænset adgang til sociale medier såsom Facebook, Snapchat, Instagram og TikTok. Der har været en formel aldersgrænse på 13 år, men den er aldrig håndhævet af techgiganterne. Nu kan videnskaben dokumentere markante negative sociale og psykiske effekter forbundet med de sociale mediers revolution af børns og unges liv. Afhængigheden er at sammenligne med hårde stoffer, som fx Stanford-professor Anna Lembke har dokumenteret⁹. Derudover viser nye studier fra Yale¹⁰ og fra National University of

⁹⁹ [Digital Addictions Are Drowning Us in Dopamine - WSJ](#)

¹⁰ [Study Probes Connection Between Excessive Screen Media Activity and Mental Health Problems in Youth < Yale School of Medicine](#)



Singapore¹¹ også negative konsekvenser for børnene ved langvarig brug af skærme, såsom øget risiko for stress, depression, samt hukommelsesproblemer og koncentrationsbesvær.

Med ChatGPT og lignende produkters indtog står vi over for en ny udfordring for vores børn og unges trivsel på sociale medier. Snapchat har f.eks. installeret ChatGPT i deres app, så børn og unge kan chatte med deres "AI-ven". Dette kalder på akut handling i form af ikke-personhenførbare aldersverifikation på sociale medier og AI-produkter som f.eks. ChatGPT. Aldersverifikation kan bidrage til at øge beskyttelsen af børn og unge, når de færdes online. Børn og unge udsættes for afhængighedsskabende sociale medier, der ændrer deres ikke fuldt udviklede hjerner på en irreversibel måde, og derfor foreslår vi en aldersgrænse på 15 år.

10. Forbud mod afhængighedsskabende mekanismer rettet mod børn

For at beskytte vores børn og unge mod de destruktive konsekvenser af techgiganternes profitjagt, foreslår vi et forbud mod negativ adfærdsdesign i digitale produkter rettet mod børn og unge, der er indrettet for at fastholde deres tid i videst mulig omfang. Afhængighedsskabende mekanismer, såsom *continuous scroll* på Facebook eller TikTok, *autoplay* på Netflix, *snapstreaks* på Snapchat m.m. *nudger* børn til at fortsætte og opretholde deres online tilstedeværelse i et afhængighedslignende forhold til de digitale tjenester. Afhængighedsskabende og fastholdelsesorienteret adfærdsdesign i digitale produkter rettet mod børn og unge under 18 år skal derfor forbydes. Disse forslag bør indarbejdes i en digital børnebeskyttelseslov, som foreslået af SF i 2020.¹²

SF anbefaler desuden, at alle sociale medier og andre udbydere af digitale produkter skal indarbejde positivt adfærdsdesign, som opfordrer børn og unge til at komme væk fra skærmen.

¹¹ [Associations Between Infant Screen Use, Electroencephalography Markers, and Cognitive Outcomes | Child Development | JAMA Pediatrics | JAMA Network](#)

¹² https://www.ft.dk/ripdf/samling/20201/beslutningsforslag/b54/20201_b54_som_fremSAT.pdf

11. Mærkningsordning om brug af kunstig intelligens

Kunstig intelligens kan nu generere indhold og produkter, som til forveksling ser menneskeskabte ud. F.eks. har producenten i TV2's dokumentar "Sidney" anvendt kunstig intelligens til at genskabe Sidney Lees stemme, så seeren kan opleve den afdøde reality-stjerne fortælle sin egen historie.

Et andet eksempel er et billede af Pave Frans ikklædt en hvid dunjakke, som hurtigt gik viralt, men som viste sig at være AI-genereret. Sådanne eksempler understreger behovet for, at AI-genererede produkter bliver deklareret, så vi som forbrugere er klar over, under hvilke forudsætninger produkter er tilblevet, og hvilke konsekvenser det kan have for indholdet.



Vi foreslår en mærkningsordning, som i første omgang skal sikre, at det deklarerer, hvis et produkt eller service er (helt eller delvist) skabt med AI. Da udviklingen med AI går lynende hurtigt, lægges der op til en dynamisk ordning, som løbende revideres. Det kan f.eks. være nødvendigt at tilpasse ordningen i det tilfælde, at mængden af AI-generede services og produkter stiger så markant, at produkter som udgangspunkt antages at være AI-genereret. Mærkningsordningen kan i stedet bruges til at deklarerer bestemte risici som f.eks. bias. Denne ordning vil helt konkret betyde, at f.eks. nyhedsartikler (helt eller delvist) skabt ved brug af kunstig intelligens skal deklarerer i overensstemmelse med ordningens bestemmelser.

Indtil AI-forordningens mærkningsordning er etableret, anbefales det, at virksomheder og organisationer bag produkter eller services, som er skabt med AI, deklarerer dette på en måde, så det er umiddelbart synligt for forbrugeren. Det kan f.eks. ske ved at indskrive en sætning, som deklarerer omfattende brug af AI i toppen af førnævnte nyhedsartikel eller et vandmærke på et AI-generet billede. AI-deklaration bliver under alle omstændigheder indført med den kommende EU-lovgivning (AI Act).

12. Producenter skal varedeklarere AI produkter

Ligesom en traditionel virksomhed skal varedeklarere sine fysiske produkter, før det møder forbrugeren, bør virksomheder, der sælger produkter med kunstig intelligens, formulere en produktbeskrivelse med kunstig intelligens som særligt fokus.

Vi foreslår, at virksomheder/udgivere af produkter og services, som benytter kunstig intelligens, skal sørge for, at forbrugerne møder en varedeklaration eller en produktbeskrivelse, inden produktet tages i brug. Produktbeskrivelsen skal blandt andet forklare produktet og dets virkemåde, ligesom det skal deklarerer, hvorvidt data opsamles, hvilket formål data opsamles til,

samt mulighed for sletning af data. Herudover skal risici for fejl i produktet, f.eks. risiko for misinformation og desinformation, også fremgå af produktbeskrivelsen, samtidig med at ansvarsplaceringen for disse beskrives.

13. Asyl for whistleblowere om skadelige effekter af kunstig intelligens

Der har i de seneste år været eksempler på programmører fra store tech-virksomheder, som er blevet afskediget efter at have udtalt sig om deres bekymringer om skadelige effekter af kunstig intelligens. Et eksempel er den tidligere Google-ingeniør, Greg Coppola, der i 2019 blev suspenderet, efter han i et interview beskyldte Google for at implementere politisk bias i sine algoritmer. Whistleblowere som Greg Coppola bidrager med vigtig viden om, hvordan tech-giganterne designer deres platforme og tilrettelægger deres algoritmer. Det er viden, som er essentiel for vores mulighed for at regulere og lovgive på området.

Vi foreslår derfor, at Danmark skal tilbyde whistleblowere asyl i Danmark, som vi før har gjort det med forfulgte forfattere. Hermed øges incitamentet for at stille sig frem med bekymringer om skadelige effekter af produkter, som bruger kunstig intelligens. Samtidig kan Danmark tiltrække nogle af de klogeste mennesker indenfor AI-feltet, som samtidig har vist, at de er fast besluttede på at arbejde for det fælles bedste.

14. Borgerting - Involvering af borgerne

I vores udspil har vi foreslået en række tiltag, som skal vedtages og implementeres ovenfra og ned. Men i et demokratisk samfund som det danske er det afgørende, at vi inddrager borgerne i så store og vigtige spørgsmål som brugen af kunstig intelligens. Vi foreslår derfor, at der oprettes en borgersamling, som skal formulere visioner og principper for anvendelse af kunstig intelligens i Danmark.

Derudover skal der iværksættes tiltag for at styrke den offentlige debat og bevidsthed vedrørende kunstig intelligens. En borgersamling om tech og menneskerettigheder blev også gennemført i Canada i 2022¹³, og der har ligeledes været gode erfaringer med borgersamlinger i forhold til abort i Irland og klima i Frankrig¹⁴.

15. Copyright-beskyttelse og licensaftale med medierne

I en tid, hvor AI-værktøjer som chatbotter vinder frem, er det vigtigt at sætte en række hegnsplæer, som viser, hvordan teknologierne kan og må anvendes i vores samfund. Her er det væsentligt at sikre det indhold, som tjenesterne er helt afhængige af i deres træning af chatbotterne.

Derfor vil SF indføre en specifik aftalelicens, der kan bidrage til at slå ring om kvalitetsindhold, vedrørende tekst og datamining (TDM). TDM er en automatiseret teknik, der kan analysere tekst

¹³ [CitizensAssemblyOnDemocraticExpression-PPF-SEP2022-FINAL-REPORT-EN-1.pdf \(squarespace.com\)](https://citizensassembly.ondemocraticexpression.org/PPF-SEP2022-FINAL-REPORT-EN-1.pdf)

¹⁴ <https://propositions.conventioncitoyennepourleclimat.fr/> & <https://2016-2018.citizensassembly.ie/en/The-Eighth-Amendment-of-the-Constitution/>



og data i digital form med henblik på at generere oplysninger, herunder mønstre, tendenser og korrelationer til videre anvendelse. Rettighedsdeklarering og kollektive aftaler er i samfundets interesse som en bæredygtig håndtering af tekst- og datamining, der involverer ophavsretsbeskyttet materiale. Tekst- og datamining har også væsentlig værdi for rettighedshaverne. Hertil kommer, at kun en specifik aftalelicens kan give kompetence til Ophavsretslicensnævnet til at fastsætte vederlag i tilfælde af uenighed herom.

I forbindelse med implementeringen af copyright-direktivets anden fase er der en mulighed for at tage et skridt i den rigtige retning ved at sikre muligheden for at anvende en aftalelicens. Det er relativt sjældent, at vi kan gøre noget nationalt – nu har vi muligheden, og den bør udnyttes til fulde.

Den første og progressive tilgang til at give mulighed for aftalelicens har faktisk gjort, at medier i alle andre EU-lande nu kigger til Danmark som foregangsland ift. det helt konkrete og håndfaste redskab til at kunne gå op mod techgiganterne. Muligheden for en opfølgning og yderligere sikring af kunstneres og ophavsmænds rettigheder må ikke tabes på gulvet, mens den teknologiske udvikling buldrer afsted.

16. Nordisk Center for Tech og Demokrati

Vi foreslår, at de nordiske regeringer i fællesskab – og inden årets udløb – nedsætter et Nordisk Center for Tech og Demokrati, som foreslået af den Nordiske Tænk tank for Tech og Demokrati. I Danmark har vi allerede et dansk center for tech og demokrati, men for at få flere muskler og flere erfaringer fra en række lande, mener vi, at vi også bør have et nordisk center. Det nordiske center skal både understøtte håndhævelse af europæisk regulering af tech-virksomheder og understøtte uafhængige forskeres adgang til platformsdata og algoritmer.



Definitioner

Adfærdsdesign: handler om at påvirke menneskers valg og adfærd ved at ændre på aspekter af deres omgivelser, herunder digitale omgivelser. Eksempler på digitalt adfærdsdesign er grønne jaknapper, pop-up notifikationer, streaks på Snapchat og tre prikker, der på en kommunikationsplatform indikerer, at brugeren har set beskeden og er tilgængelig.

AI/ kunstig intelligens: Systemer med evne til at tilegne, behandle og anvende viden og/eller færdigheder. Kunstig intelligens sammenlignes som regel med menneskers intelligens, hvilket dog ikke er udtømmende for begrebet. Maskinlæring er en underkategori af kunstig intelligens.

Autoplay: en funktion som får en ny video, lydfil mm. til at starte, når det sete eller hørte stopper.

Bias: forudindtagethed i beslutninger eller metodisk fejlbehæftet datagrundlag for beslutninger

Chatbot: et computerprogram, der er designet til at simulere samtale med mennesker.

Deep fake: Deepfake er en kombination af ordene deep learning og fake(falsk), Et billede eller en optagelse som er blevet ændret og manipuleret med det formål at fremstille en person som havende sagt eller gjort noget, som ikke blev sagt eller gjort i virkeligheden. I dag er det muligt at lave så overbevisende falske videoer og billeder, hvor man kan få hvem som helst til at sige hvad som helst, at det kan være utroligt svært at se, om det er ægte.

Differential privacy: Det er et matematisk framework til at sikre privatlivets fred for enkeltpersoner i datasæt. Det kan give en stærk garanti for privatlivets fred ved at tillade, at data analyseres uden at afsløre følsomme oplysninger om nogen person i datasættet.

Disinformation: forstås i dette udspil som falsk, ukorrekt eller misvisende information, der er designet til og bevidst bliver præsenteret og promoveret til at gøre skade på samfundet eller for at tjene penge.

Dopamin-rus: Dopamin er en neurotransmitter, et signalstof, som regulerer lystfølelse og eufori og som bliver brugt af hjernens belønningssystem.

Generativ kunstig intelligens: Generativ kunstig intelligens, generative AI eller GenAI, er en type kunstig intelligens-system, der gør det muligt at genere f.eks. tekst eller billeder på baggrund af en række data. Generative AI-systemer lærer mønstre og strukturer fra de data, der er blevet stillet til rådighed og genererer herudfra nyt indhold, der svarer til indholdet i dataene, men er nyt i en eller anden grad. Dvs. at den ikke bare klassificerer data, men bearbejder dem til nyt indhold. Unimodale systemer bruger kun én type input, f.eks. tekst, mens multimodale systemer kombinerer flere typer, f.eks. både tekst og billeder.

Infinity scroll/continuos scroll: en web-design-teknik som tillader brugeren at scrolle uendeligt ned af siden ved at synliggøre nyt indhold i takt med, at man scroller. Denne teknik, som bruges i både apps og på hjemmesider, fjerner behovet for at skifte side og minimerer dermed interaktionsomkostningerne for brugeren.



Misinformation: omfatter falske informationer, videoer med forkert billedtekst, forkerte tal eller blot rygter, hvor afsenderen ikke nødvendigvis er klar over, at der er tale om ukorrekte informationer.

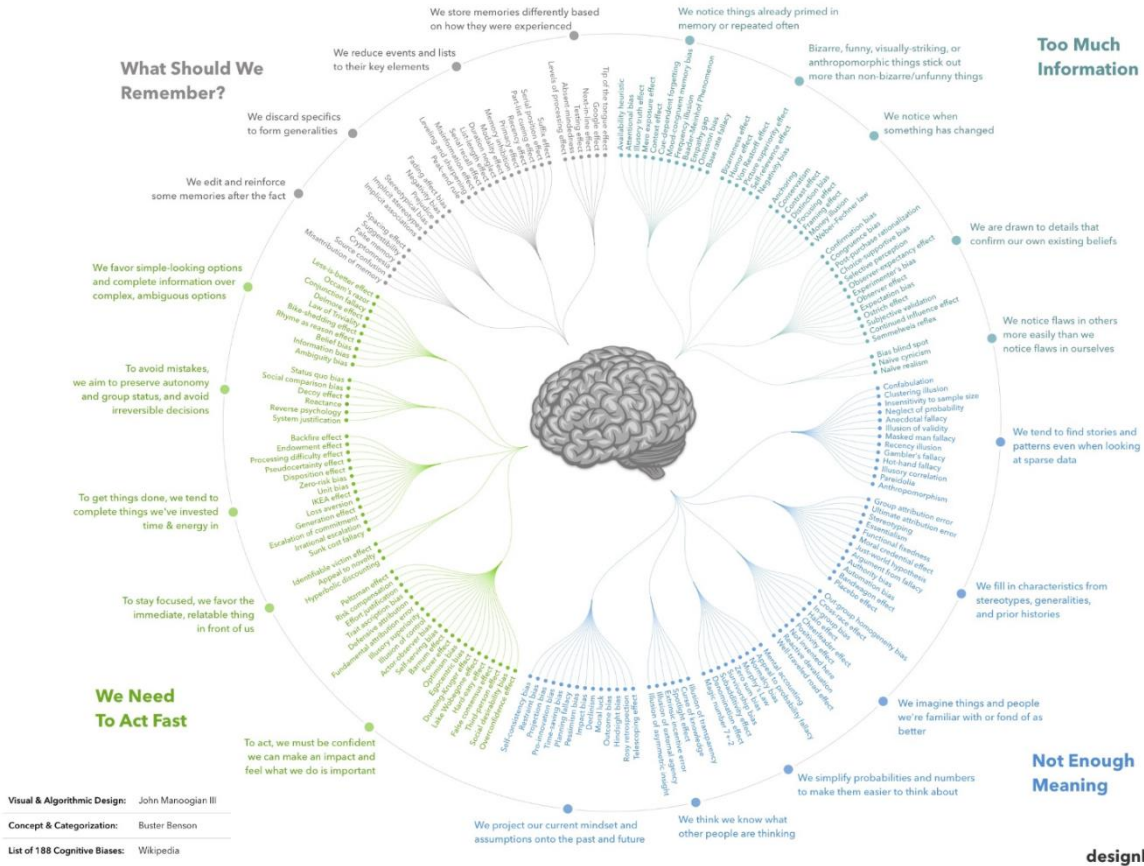
Nudging: nudging, også som udsagnsord, "at nudge", er forsøg på at påvirke menneskers valg og adfærd i en forudsigelig retning, for eksempel at fastholde brugeren på mediet. En persons valg bliver påvirket af, at nogle gør det nemmere at træffe et bestemt valg, f.eks. at følge fodsporene hen mod skraldespanden. Når vi bruger sociale medier bliver vores valg konstant påvirket af de underliggende algoritmer, der er bestemmende for det indhold, vi bliver eksponeret for.

Snapstreaks: Snapstreaks opstår, når to "venner" på Snapchat har sendt billeder (snaps) til hinanden i mere end tre dage i træk, herefter vil der fremgå en lille flamme ud for navnet på personen, som man har *streak* med, og et tal viser hvor mange dage streaken har været i gang. Hvis man er tæt på at overskride den deadline og derfor snart er nødt til at sende billeder (snaps) til hinanden, dukker et timeglas op ud for den person, man har streaken kørende med. På denne måde er snapstreaks et eksempel på en vanedannende norm, som er skabt af adfærdsdesignere til dagligt at fastholde brugeren på mediet.

Sociale medier: Sociale Medier er digitale platforme, som muliggør interaktion mellem brugere. Eksempler er Facebook, Twitter og Instagram.

Bilag I

COGNITIVE BIAS CODEX



Kilde: [Every Single Cognitive Bias in One Infographic \(visualcapitalist.com\)](https://www.visualcapitalist.com/every-single-cognitive-bias-in-one-infographic/)



Tak

Tak til en række eksperter, der har bidraget til dette udspil.

I ved hvem I er.



AI-genereret (og dermed falskt!) billede af Donald Trump.