



Association of IT Professionals

Vester Farimagsgade 37A  
1606 København V

Søren Frichs Vej 38K  
8230 Åbyhøj

Overgade 54  
5000 Odense

prosa.dk  
+45 33 36 4141

Til:

Digitaliseringsministeren

Kopi til:

DIU, ekspertgruppen

København, 2024-03-01

## Kommentar til "Grænser for tech-giganternes udvikling og anvendelse af kunstig intelligens"

I PROSA har vi med interesse læst delrapporten, og vi synes der er mange gode anbefalinger.

Vi vil dog gerne knytte et par kommentarer til 3 af anbefalingerne.

### Rapportens 1.1: Indsats mod skadeligt og vildledende indhold genereret af kunstig intelligens

Ekspertgruppen foreslår at mærke AI-genereret indhold. Det vil næppe komme til at virke, for allerede i dag kan man køre softwaren på ens eget udstyr, og dermed undgå at få mærket ens indhold som AI-genereret.

PROSA tror, at vi vil have større succes med at mærke det indhold, som kan verificeres. Det kunne være med teknologier som <https://contentcredentials.org/> hvor man med få klik kan verificere, hvor indholdet kommer fra.

Måske kunne man også opfordre sociale medier til, at verificerbart indhold bliver ranket højere af algoritmerne, og at mulighederne for at verificere bygges ind i de sociale medier, så det bliver helt naturligt for brugerne at verificere indholdet.

### Rapportens 3.1: Tech-giganterne skal dokumentere, at de ikke bryder ophavsretten

Perfect10-dommen ([https://en.wikipedia.org/wiki/Perfect\\_10,\\_Inc.\\_v.\\_Amazon.com,\\_Inc.](https://en.wikipedia.org/wiki/Perfect_10,_Inc._v._Amazon.com,_Inc.)) er den dom, der gør, at Google Image Search kan vise miniature-billeder: *This was because they were "highly transformative."* Det er svært at se, at billeder til træning af AI ikke også er "highly transformative", hvis en miniature er det.

Hvis vi kræver, at der skal indgås aftaler med leverandører af træningsdata, som er offentligt tilgængelige, så skal vi være opmærksomme på, at vi de-facto udelukker alle andre end tech-giganterne fra at bruge disse træningsdata: Det er kun giganterne, der vil have ressourcerne til at indgå sådanne aftaler; små forskningsgrupper vil ikke. Dermed vil det blive langt sværere at konkurrere med tech-giganterne på det punkt.

For at kunne give plads til forskere kunne man kræve, at hvis man vil bruge offentligt tilgængelige træningsdata uden aftale, så skal modellen gøres frit tilgængelig under en fri licens. Det vil også gøre det nemmere at samarbejde om en dansk sprogmodel, idet andre små sprogområdet vil stå med lignende udfordringer.

## **Rapportens 4.2: Retningslinjer for brug og indkøb af kunstig intelligens i den offentlige sektor**

For at beskytte borgerens privatliv vil det være et rimeligt krav, at sprogmodellen kører på hardware, der er 100% under det offentliges kontrol og ikke blot kører som en service hos en tech-gigant. Derved bliver det nemmere at sikre at data ikke lækker og undgå sager som Chromebook-sagen.

Til selve træningen af modellen vil dette næppe være realistisk: Træning kræver voldsomt meget mere udstyr end brug af en model.

Med venlig hilsen

Niels Bertelsen  
Formand  
PROSA

Ole Tange  
IT-politisk rådgiver  
PROSA