

København, 26. juni 2023 (version 2)

Udvikling af kunstig intelligens (AI)

RettighedsAlliancens vigtigste budskaber

RettighedsAlliancen byder udviklingen af AI velkommen, og vi kan se mange måder, hvorpå AI kan bruges som endnu et digitalt værktøj for kunstnere og skabere af kreativt indhold. Som med alle nye teknologier med stort potentiale for at ændre vores samfund, er der behov for en fornuftig og velovervejet regulering, som tager højde for de forskelligartede samfundshensyn, herunder respekten for kreatives rettigheder.

Drivkraften bag den særlige variant af kunstig intelligens kaldet generativ AI, såsom AI-tjenesten ChatGPT, er algoritmer baseret på maskinlæring, der kan "forstå", skabe og forudsige nyt indhold. Generative AI-tjenester skaber menneskelignende indhold baseret på analyser af massive mængder tekst, billeder og musik, som er fundet på internettet.

Udviklingen af generativ AI er på det seneste gået rigtig hurtigt, og tjenesterne som ChatGPT (tekst), Midjourney (billeder) og MusicGen samt stemmemodeller, der kan efterligne virkelige kunstnere (musik), kan i dag bruges af alle med adgang til internet, og de vurderes at få enorm indflydelse blandt andet på den måde, vi bruger og skaber indhold på.

Derfor sker der i øjeblikket også en hurtig udvikling af reguleringen, der skal ramme-sætte udviklingen og brugen af generativ AI. Reguleringen sker internationalt, i EU, og nationalt.

På EU plan indeholder DSM-direktivet fra 2021 i direktivets Art. 3 og 4 nye undtagel-ser til ophavsretten, hvormed der gives adgang til at lave udtræk eller lave eksem-plarer af ophavsretligt beskyttet indhold, med henblik på at foretage analyser af indholdet for derved at generere oplysninger, herunder mønstre, tendenser og kor-relationer. Analysen betegnes som "tekst- og dataminering" i direktivet. Tekst- og dataminering (TDM) antages at dække over den proces, hvor generative AI-mo-deller trænes til at kunne forstå, skabe og forudsige nyt indhold.

Den danske regering har implementeret Art. 3 og 4 ved Lov nr. 680 af 6. juni 2023 om ændring af ophavsretsloven. De nye TDM-undtagelser findes i ophavsretsloven § 11b (Art. 4 om generel adgang til TDM) og 11c (Art. 3 adgang for forskning og kul-turarv TDM).

Den 14. juni 2023 vedtog EU-Parlamentet deres ændringsforslag til en kommende forordning om kunstig intelligens, EU AI Act, som bredt regulerer kunstig intelligens i EU. På nuværende tidspunkt indeholder EU AI Act betingelser for udbydere af "foundation models" (AI grundmodeller). Hertil kommer særlige betingelser for ud-byderen, hvis grundmodellen skal bruges i en AI-tjeneste, som kan generere ind-hold såsom tekst, billeder, audio og video (generativ AI), og hvis en udbyder har fintunet en eksisterende grundmodel til at kunne bruges som generativ AI-tjeneste. Forordningen er i juni 2023 gået ind i Trilogforhandlinger og den forventes færdigbe-handlet ved udgangen af 2023.

At sikre vores medlemmers ret til at kontrollere brugen af deres indhold er Ret-tighedsAlliancens kerneopgave. Derfor arbejder RettighedsAlliancen fokuseret med at forstå, kortlægge og udvikle tiltag, der kan sikre at vores medlemmers op-

havsretligt beskyttede indhold ikke indgår i udviklingen af grundmodeller og generativ AI uden deres samtykke. Dette papir beskriver RettighedsAlliancens position og budskaber i relation til de politiske processer, der er i gang i øjeblikket.

De primære budskaber er:

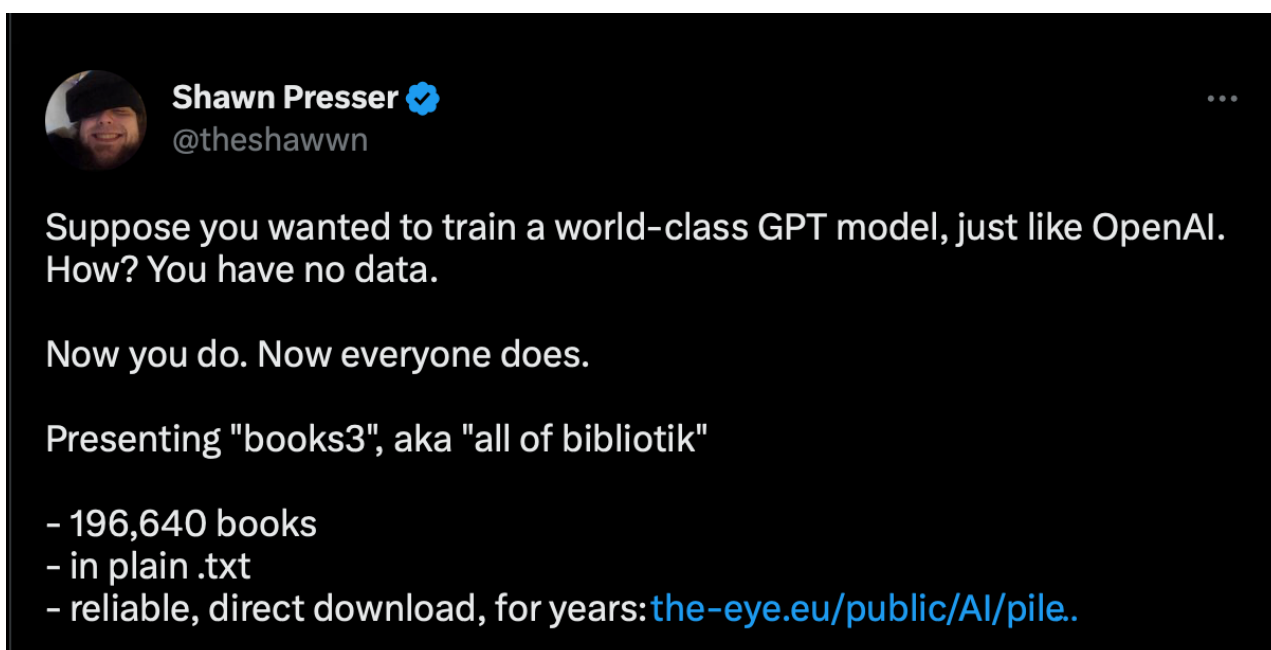
- Ulovlige eksemplarer af indhold, der tilhører RettighedsAlliancens medlemmer, må ikke indgå i datasæt, der bruges til at udvikle generative AI-modeller
- Indførelsen af TDM-undtagelsen i ophavsretslovens § 11b, stk. 2 har skabt et behov hos rettighedshavere for at vide, (1) **hvilket** beskyttet indhold, der har været kopieret til brug for TDM, (2) **hvor** på internettet indholdet er tilgået og kopieret fra, (3) og **hvordan** indholdet er tilgået.
- Uden indsigt i ovennævnte 3 punkter, har rettighedshavere ikke mulighed for at kontrollere,
 - o Om der er anvendt *ulovligt indhold* til at træne AI-modeller,
 - o Om der har været en *lovlig adgang* til de værker/frembringelser, der har været udsat for eksemplar fremstilling til brug for TDM, og
 - o Hvorvidt evt. *forbehold* er overholdt.
- EU-Parlamentets ændringsforslag til AI Act art. 28b (4)(c) om transparens er ikke tilstrækkelig til at opfylde rettighedshavernes behov for at kontrollere ovenstående.

Eftersom området udvikler sig hurtigt, vil RettighedsAlliancen også løbende tilpasse budskaber og position på området. Vores positionspapir beskæftiger sig udelukkende med den generelle adgang til TDM i ophavsretslovens § 11b. Papiret beskæftiger sig således ikke med de særlige regler, der gælder for forskningsorganisationer og kulturarvsinstitutioner.

Eksempel på ulovlige eksemplarer der indgår i udvikling af grundmodeller, som kan generere indhold såsom tekst, billeder, audio og video (generativ AI)

RettighedsAlliancen har konstateret, at ulovlige eksemplarer af ophavsretligt beskyttet indhold tilhørende danske rettighedshavere, indgår i offentligt tilgængelige træningsæt til udviklingen af grundmodeller og generative AI-modeller/tjenester. Et eksempel herpå er datasættet "Books3", der består af ca. 200.000 ulovlige kopier af e-bøger, som alle stammer fra en ulovlig tysk fildelingstjeneste "bibliotik".

Nedenstående skærmpoint viser opslag fra en person ved navn Shawn Presser, hvor han offentliggør datasættet og beskriver dets indhold og formål til træning af grundmodeller som Open AI's GPT modeller.



Vi ved, at Books3 datasættet er indgået i træningen og udviklingen af følgende generative AI-modeller:

- Facebook's Large Language Model (LLaMA)¹

¹ LLaMA: Open and Efficient Foundation Language Models, <https://arxiv.org/abs/2302.13971>

- BloombergGPT²
- StableLM³
- GPT-Neo⁴
- GPT-J⁵

Det ved vi, fordi udviklerne har besluttet at offentliggøre den data, der har været brugt til at træne og udvikle deres AI-modeller. Men det er langt fra alle udbydere, der gør det. Fx er det indhold, som OpenAI har anvendt til at træne deres grundmodeller GPT-3 og senest GPT-4 (der fungerer som grundmodel for den generative AI-tjeneste ChatGPT), ikke blevet offentliggjort i tilstrækkelig detaljeret grad til at identificere, hvor alt data stammer fra.

RettighedsAlliancen har i juni 2023 påbegyndt en sag med det formål at begrænse offentlighedens adgang til Books3 datasættet mest muligt.

² BloombergGPT: A Large Language Model for Finance, <https://arxiv.org/pdf/2303.17564.pdf>

³ StableLM, <https://github.com/stability-AI/stableLM/>

⁴ <https://github.com/EleutherAI/gpt-neo>

⁵ <https://huggingface.co/EleutherAI/gpt-j-6b>

RettighedsAlliancens position og beskrivelse af de væsentligste udfordringer

1. Behovet for en effektiv transparenspligt

Det er vigtigt at se TDM-undtagelsen i § 11b, stk. 2 og transparenspligten i EU AI Act i sammenhæng, da dette i første omgang bliver reguleringen af AI på ophavsretsområdet. I dette lys er det væsentligt, at transparensforpligtelsen i AI Act lever op til de behov, som TDM-undtagelsen har skabt for rettighedshavere, herunder behovet for at kontrollere, om der har været en *lovlig adgang* til værker/frembringelser, der har været udsat for udtræk og eksemplar fremstilling til brug for TDM, og hvorvidt evt. *forbehold* er blevet overholdt.

Rettighedshavere har brug for at vide, (1) **hvilket** beskyttet indhold, der har været kopieret til brug for TDM, (2) **hvor** på internettet indholdet er tilgængeligt og kopieret fra, (3) og **hvordan** indholdet er tilgængeligt (lå det frit tilgængeligt, eller krævede det en aftale/abonnement for at få adgang, og hvad dækkede aftalen/abonnementet).

Tekst- og dataminering

Der stilles i dag krav om, at der skal være en *lovlig adgang* til det ophavsretligt beskyttede indhold, der udtrækkes eller eksemplar fremstilles til brug for tekst- og dataminering (TDM). Det fremgår af reglerne i den vedtagne lov nr. 680 af 6. juni 2023 om ændring af ophavsretsloven, § 11b, stk. 2, og af de almindelige bemærkninger til lovforslaget punkt 2.1.2, hvorefter kravet om *lovligt forlæg* eksplicit fremhæves som en forudsætning for at indhold må udtrækkes eller eksemplar fremstilles til brug for TDM.

En ulovlig adgang kan forekomme, hvis det der udtrækkes eller laves eksemplarer af stammer fra **ulovlige kilder** på internettet. Eksempelvis på ulovlige tjenester, hvor indholdet stilles til rådighed uden tilladelse fra rettighedshaverne. Det kunne være

ulovlige fildelingstjenester som thepiratebay.org eller b-ok.org. Her er kravet om lovligt forlæg ikke opfyldt, og TDM-undtagelsen kan ikke gøre de ulovlige eksemplarer lovlige.

Der kan også være tale om en ulovlig adgang, hvis det der udtrækkes eller laves eksemplarer af stammer fra **lovlige kilder**, men hvor rettighedshaverne har taget et *forbehold* i medfør af § 11b, stk. 2, hvorefter TDM-undtagelsen i § 11b, stk. 2 kun "*finder anvendelse på betingelse af, at rettighedshaveren ikke udtrykkeligt har forbeholdt sig anvendelsen af værket på passende vis*". Derfor er der også tale om uretrettiget eksemplarfremstilling, hvis rettighedshaveren har taget et passende forbehold for brugen af deres indhold til TDM, og der desuagtet foretages udtræk eller eksemplarfremstilling.

En ulovlig adgang kan også forekomme ved udtræk eller eksemplarfremstilling på lovlige kilder såsom Saxo.com, men hvor betingelserne for adgangen for vedkommende, som har lavet udtræk eller eksemplarfremstilling mhp. TDM, ikke dækker over den faktiske brug. Fx vil der efter RettighedsAlliancens vurdering ikke være lovlig adgang i det tilfælde, hvor det abonnement, som vedkommende har anvendt til at få adgang til indholdet, er forbeholdt en personlig adgang, hvorimod vedkommende har lavet udtræk eksemplarfremstilling mhp. TDM til brug for en virksomhed.

Rettighedshaverne har imidlertid, uanset disse regler, **ikke mulighed for at kontrollere**, om der har været en lovlig adgang til deres indhold. Det skyldes, at der ikke stilles krav om transparens vedrørende brug af indhold til TDM.

EU AI Act

I EU-Parlamentets ændringsforslag til kommende AI Act er der i **Art. 28b** lagt op til en række forpligtelser for udbydere af grundmodeller på det indre marked.

Art. 3 (1) og (1)(c) indeholder en definition af kunstig intelligens system (*artificial intelligence system*) og af grundmodeller (foundation model):

'artificial intelligence system' (AI system) means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments;

'foundation model' means an AI system model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks;

Den centrale bestemmelse for transparens vedrørende ophavsretligt beskyttet indhold brugt til at træne AI-modeller er **Art. 28b (4)(c)**. Bestemmelsen pålægger (1) udbydere af *grundmodeller*, som anvendes i AI-tjenester med det specifikke formål at generere tekst, billeder, audio og video (*generativ AI*), og (2) udbydere som fin-tuner en grundmodel til at blive en *generativ AI-tjeneste*, at dokumentere og offentliggøre en **tilstrækkelig detaljeret opgørelse** over indhold beskyttet af ophavsret, som er indgået i træningsdata for modellen.

Art. 28b (1) pålægger udbyderen af en grundmodel at sikre sig overholdelsen af betingelserne i 28b før udgivelse af grundmodellen i det indre marked.

Art. 28b (2)(g) indeholder en registreringspligt for udbydere af grundmodeller på det indre marked. Grundmodellen skal registreres i en EU database omtalt i Art. 60 i overensstemmelse med Annex VII pkt. C. I dette annex's nr. 5 fremgår det, at der skal indgives oplysninger om data kilder ("data sources"), som har været brugt til at udvikle grundmodellen. Oplysningerne skal holdes opdaterede.

RettighedsAlliancens vurdering af reglerne om transparens i AI Act

RettighedsAlliancen vurderer, at indholdet af 28b(4)(c) og (2)(g) ikke i tilstrækkelig grad opfylder de behov, som TDM-undtagelsen har skabt for rettighedshavere.

Det skyldes, at rettighedshavere, ud over at vide **hvilket** beskyttet indhold, der har været brugt til træning, også har behov for at vide, **hvor** på internettet indholdet er tilgængeligt og lavet udtræk eller eksemplarfremstilling fra, og **hvordan** indholdet er tilgængeligt dvs. om indholdet lå frit tilgængeligt på internettet, eller om rettighedshaverne havde begrænset adgangen til indholdet (fx en betalingsmur), og i givet fald om de adgangsplysninger, som vedkommende anvendte til at få adgang, dækkende den faktiske brug. Eksempelvis, hvor et privat abonnement anvendes af en person, der i arbejdsregi tilgår og laver eksemplarer af beskyttet indhold til brug for TDM. Dette mener RettighedsAlliancen ikke vil være en lovlig adgang, da abonnementet er privat og eksemplarer lavet til brug for en virksomhed vil mangle *lovligt forlæg*).

EU AI Act transparenspligten mangler altså **hvor** og **hvordan** beskyttet indhold er tilgængeligt.

Derudover viser RettighedsAlliancens egne erfaringer med at undersøge datasæt til brug for træning af AI-modeller, at det kan være **meget vanskeligt at beskrive indholdet tilstrækkeligt** til at identificere rettighedshavere. Der findes et datasæt ("Books3" omtalt ovenfor på side 3-5) med ca. 200.000 ulovlige e-bøger, der er blevet brugt til at træne alt fra Meta's AI-model LLaMA til Bloomberg's BloombergGPT.

De ulovlige e-bøger stammer fra en ulovlig tysk fildelingstjeneste, og det er ikke alle e-bøger, der stadig har en kolofon med angivelse af rettighedshaver i. I dette konkrete eksempel har hverken Meta eller Bloomberg mulighed for på tilstrækkelig vis at beskrive, hvilket ophavsretligt beskyttet indhold, der er blevet brugt til at træne deres AI-model.

Den mest optimale løsning, der sikrer rettighedshaverne kontrol med deres indhold, vil være, at udviklerne af AI grundmodeller pålægges at lagre og opbevare det

indhold, der er brugt til at træne AI grundmodeller på. Men da denne løsning er forbundet med store sikkerhedsmæssige risici, fx at indhold lækkes, eller der hackes adgang til det, er RettighedsAlliancen forsigtige med at anbefale dette.⁶

Det skal for nuværende understreges, at den foreslåede artikel 28b(4)(c) skal fortolkes således, at udbydere af AI grundmodeller skal dokumentere og offentliggøre tilstrækkelig detaljerede oversigter over det **ophavsretligt beskyttede indhold og de relevante rettighedshavere**, der er anvendt til træning af en AI grundmodel. Det fremgår ligeledes af artikel 28b(1), at disse informationer skal offentliggøres inden AI grundmodellen sættes på det indre marked. Derfor bør det også understreges, at hvis ikke udbyderne inden AI modellen sættes på markedet dokumenterer og offentliggør tilstrækkelig med information om indhold og rettighedshavere til træningsdata, så følger det af 28 b (1), at udgiveren **ikke lovligt kan udgive AI tjenesten**.

2. Rettighedshavere skal kunne forhindre at brugere uploader ulovlige eksemplarer til AI-tjenester

Brugere af generative AI-tjenester, som ChatGPT eller Midjourney, kan fodre/uplade beskyttet indhold, såsom tekst fra bøger eller beskyttet billedkunst/fotografier, som AI-tjenesten benytter i dets genering af nyt output.

Selvom det uploadede beskyttede indhold ikke har været en del af træningen af grundmodellen eller AI-tjenesten, er der en formodning for, at AI-tjenesten anvender indhold uploadet af brugere til at videreudvikle grundmodellen/AI-tjenesten.

⁶ IFPI London anbefaler en position, hvorefter de ønsker: *Push for stricter record-keeping and disclosure obligations for foundation models which would oblige the providers to disclose data used to train their systems.*

RettighedsAlliancen vil arbejde for, at der iværksættes filtrering eller lignende løsninger, der kan forhindre, at ulovlige eksemplarer af rettighedsbeskyttet indhold kan uploades til generative AI-tjenester. Erfaringerne fra samarbejderne med online-platforme om automatiske genkendelses-værktøjer og filtre kan indgå i udviklingen af dette område.

3. Fremtidig mulighed for sikkerhedsscreening af træningsdata

RettighedsAlliancen vil arbejde for en model, hvor udbydere af grundmodeller og generative AI-tjenester forud for offentliggørelsen af AI-modellen på det danske marked skal give rettighedshavere til indhold, der har været anvendt til træning af modellen, mulighed for at gennemgå træningsdata og forbeholde sig retten til, at deres beskyttet indhold må anvendes til træning af den endelige udgave af modellen. Det kunne ske via et offentligt (sikkerheds)organ, der nedsættes til det formål. Klarhed om lovligheden af anvendelsen af ophavsretligt beskyttet indhold i et træningssæt bør således være en del af en overordnet sikkerhedsscreening af dataset til træning og udvikling af AI-modeller, som skal sikre, at offentliggjorte AI-modeller og AI-tjenester ikke skader vores samfund.

EU AI Act

Indeholder bestemmelser om pligter for udbydere af grundmodeller i det indre marked, herunder at udbyderne sikrer sig, at pligterne overholdes forud for udgivelse på det indre marked, samt at oplysninger om modellen og den data, som har været brugt til at træne modellen, registreres, obevares og holdes opdateret i en EU database oprettet til formålet.

Det er alle elementer, som der kan bygges videre på på vejen mod et scenarie, hvor rettighedshavere har bedre muligheder for at forhindre ophavsretskrænkelser og vores samfund kan beskyttes mod de potentielle farer ved kunstig intelligens.

4. Opsummering

Det er vigtigt at se TDM-undtagelsen i § 11b, stk. 2 og transparenspligten i EU AI Act i sammenhæng, da dette i første omgang bliver reguleringen af AI på ophavsretsområdet.

I første omgang har TDM-undtagelsen i ophavsretslovens § 11b skabt et akut behov for, at rettighedshavere ved, **hvilket** beskyttet indhold, der har været brugt til TDM, **hvor** på internettet, det beskyttede indhold er blevet tilgået og udsat for udtræk eller eksemplar fremstilling til brug for TDM, og **hvordan** indholdet er tilgået.

EU AI Act transparenspligten mangler **hvor** og **hvordan** beskyttet indhold er tilgået.

Derudover viser RettighedsAlliancens egne erfaringer med at undersøge datasæt til brug for træning af AI-modeller, at det kan være **meget vanskeligt at beskrive indholdet tilstrækkeligt** til at identificere evt. rettighedshavere. Derfor skal det understreges, at artikel 28b(4)(c) skal forstås sådan, at udviklerne af AI grundmodeller skal dokumentere og offentliggøre data, der på tilstrækkelig vis identificerer værker og rettighedshavere, der er brugt til træning af AI-modeller. Det fremgår af artikel 28b(1) at disse oplysninger skal gives **inden en AI grundmodel sættes på markedet**, hvorfor en AI grundmodel, der sættes på markedet uden at opfylde disse betingelser, er ulovlig.

På sigt er der ligeledes et behov for at etablere løsninger, der skal forhindre, at brugere af AI-tjenester uploader beskyttet indhold til tjenesten. Her kan der også med fordel ses på DSM art. 17, og de forpligtelser der knytter sig til udbydere af onlineplatforme med brugergeneret indhold såsom Facebook og YouTube.

På den lange bane er der et behov for at etablere en **sikkerhedsscreening af datasæt** til træning og udvikling af AI-modeller, som skal sikre, at offentliggjorte AI-modeller og AI-tjenester ikke skader vores samfund, herunder bruger ulovlige kopier af ophavsretligt beskyttet indhold.