Specialeafhandling: Udvikling af generative AI i lyset af ophavsretten

Fagområde: Immaterialret

Problemformulering: Hvornår er kopieringen af træningsdata, der finder sted ifm. treningen af en ML-algoritme mhp. implementering i et generative AI-system, omfattet af eneretten til eksempelafremstilling, jf. OPHL § 2, stk. 1, når den anvendte træningsdata består af ophavsretligt beskyttede værker, i hvilket omfang er kopieringen dækket af de lovbestemte indskrænkninger i ophavsretten, og er de nugældende indskrænkninger hensigtsmæssige?

Navn                           KU-brugernavn
Christian Andreas Jul Gjølbo   JPV644

Vejleder: Jens Hemmingsen Schovsbo

Antal tegn: 140.369 tegn ekskl. mellemrum

Må benyttes på Det Juridiske Fakultets bibliotek med karakter (sæt kryds).

Må ikke benyttes på Det Juridiske Fakultets bibliotek (sæt kryds).

Studerende på Cand. soc i jura (sæt kryds)

Studerende på Cand. jur (sæt kryds)

Afleveringsdato: 30. maj 2023

Karakter:
Abstract

While the world has long relied on human creativity to differentiate creative output like art, music and text, generative AI has revolutionized the creation and speed of machine-generated creative content.

However, in order to generate new content, the generative AI system uses a machine learning ("ML") algorithm trained by analyzing data, which often consist of thousands, or millions, of copyrighted works; these data are subsequently subjected to copying.

This thesis examines when the copying of copyrighted works during ML training is subject to the exclusive right of reproduction under s. 2(1) of the Danish Copyright Act ("DCA"); when such copying is subject to the limitations to the right of reproduction; and whether the existing limitations are appropriate.

The thesis concludes that the copying of training data consisting of copyrighted works is subject to the right of reproduction under s. 2(1) of the DCA provided that the works are copied in their original and complete form, whereas the copying of individual elements of the works in their original form is only covered by s. 2(1) if such elements meet the originality requirement. Copying is also subject to s. 2(1) if the works are copied in modified form.

However, copying of training data is not subject to approval by the rights holders when such copying is subject to the limitations to the right of reproduction set out in s. 11a(1) of the DCA, Article 4(1) of the DSM Directive and s. 50(2) of the DCA. These limitations apply to some of the copies made during ML training.

The thesis also concludes that these limitations are not appropriate from a user perspective, as the limitations are narrow in scope and remain unclear and as copying in connection with the algorithm's analysis is subject always to permission. Conversely, these conditions are appropriate from the rights holders' perspective, as the rights holders' legitimate interests are rarely prejudiced as the copying during ML training requires their permission in most cases.

Intended by legislators to provide a high level of protection for rights holders, the limitations are also appropriate from a perspective of intent, but less so from a global competition perspective, as copying of works for ML training purposes in the EU is to a wider extent subject to permission compared with non-EU countries, including the US and Japan, thereby inhibiting the development of generative AI.
In conclusion, it is suggested that the scope of the existing limitations should be clarified until new limitations are introduced, which could take years. A clarification would indeed promote the development of *generative AI* and provide a more navigable legislative framework for EU-based *generative AI* developers in the future.
Indholdsfortegnelse

1 Introduktion ................................................................................................................................. 5
   1.1 Indledning ................................................................................................................................. 5
   1.2 Problemformulering .................................................................................................................. 6
   1.3 Specialets struktur .................................................................................................................... 6
   1.4 Afgrønsning .............................................................................................................................. 7
   1.5 Metode ..................................................................................................................................... 9

2 Generative AI, ML og kopiering af træningsdata ...................................................................... 10
   2.1 Hvad er generative AI? ............................................................................................................. 10
   2.2 Udvikling af Generative AI på baggrund af ML .................................................................... 13
      2.2.1 ML indebærer en kopiering af træningsdata .................................................................... 15
      2.2.2 Et eksempel på kopieringen af træningsdata ifm. træningen - Stable Diffusion .............. 16

3 Retsgrundlaget - den ophavsretlige regulering, eneretten og indskrænkningerne heri ............ 18
   3.1 Ophavsretten til litterære og kunstneriske værker .................................................................. 18
      3.1.1 Beskyttelsens genstand .................................................................................................... 19
      3.1.2 Beskyttelsens indhold ....................................................................................................... 19
   3.2 Indskrænkninger i ophavsretten ............................................................................................. 20
   3.3 Regler (eller mangel på samme?) om brug af værker til træningen af en algoritme ............ 21

4 Kopiering af træningsdata – en ophavsretligt relevant handling? ........................................... 22
   4.1 Eksemplarfremstilling på en hvilken som helst måde og i en hvilken som helst form .......... 22
      4.1.1 Kopiering af værker i oprindelig skikkelse ifm. træningen ............................................. 25
      4.1.2 Kopiering af værker i ændret skikkelse ifm. træningen ................................................... 28
   4.2 Delkonklusion ......................................................................................................................... 31

5 Kopieringen af træningsdata i lyset af de lovbestemte indskrænkninger i ophavsretten ......... 32
   5.1 Hvilke lovbestemte indskrænkninger er relevante for træningen af en ML-algoritme? ........ 32
   5.2 Undtagelsen for midlertidig eksemplarfremstilling, jf. OPHL § 11 a .................................. 34
      5.2.1 Flygtige eller tilfældige eksempler .................................................................................... 35
      5.2.2 Integreteret og væsentlig del af en teknisk proces .......................................................... 40
      5.2.3 Muliggørelse af mellemmands transmission eller en lovlig brug af et værk ............... 43
      5.2.4 Ingen selvstændig økonomisk værdi .............................................................................. 48
      5.2.5 Delkonklusion ................................................................................................................... 50
   5.3 Undtagelsen for reproduktion af værker mhp. tekst- og datamining, jf. DSM-direktivets art. 4 50
      5.3.1 Betingelser for anvendelsen af DSM-direktivets art. 4(1) ............................................. 52
      5.3.1.1 Betingelse 1: "Med henblik på" TDM ........................................................................... 52


5.3.1.2 Betingelse 2: "Lovligt tilgængelige værker" ................................................................. 56
5.3.2 Opbevaring så længe det er nødvendigt, jf. DSM-direktivets art. 4(2) .......................... 59
5.3.3 Delkonklusion.................................................................................................................. 60

5.4 Den generelle aftalelicens, jf. OPHL § 50, stk. 2 ................................................................. 60
5.4.1 En aftale om en brug af et værk inden for et nærmerede defineret område, jf. § 50, stk. 2..... 61
5.4.2 En repræsentativ rettighedsorganisation som er godkendt af Kulturministeriet, jf. § 50, stk. 4 64
5.4.3 Begrensnings: Rettighedshavers forbud efter OPHL § 50, stk. 2, 2. pkt.......................... 65
5.4.4 Delkonklusion.................................................................................................................. 65

6 Er de nugældende indskrænkninger hensigtsmæssige? .............................................................. 66
6.1 Indskrænkningernes hensigtsmæssighed fra hhv. et bruger- og et rettighedshaverperspektiv... 66
6.2 Indskrænkningernes hensigtsmæssighed fra et formålsperspektiv ........................................ 71
6.3 Indskrænkningernes hensigtsmæssighed fra et globalt konkurrenceperspektiv .................... 74
6.4 Nye regler eller en afklaring af de nugældende indskrænkningers anvendelighed? ............... 77
6.5 Delkonklusion...................................................................................................................... 78

7 Konklusion .......................................................................................................................... 78

8 Litteraturliste ........................................................................................................................ 80
1 Introduktion

1.1 Indledning

"Som dampmotoren eller elektriciteten gjorde i fortiden, ændrer AI vores verden, samfund og industri".1

Denne profeti, som EU-Kommissionen fremsatte i 2018, synes at være tættere på at gå i opfyldelse end nogensinde. På ganske kort tid har artificial intelligence, eller kunstig intelligens på dansk, (heri benævnt "AI") vundet indpas på globalt plan i et sådant omfang, at indflydelsesrige tech-skikkelser som Bill Gates betragter fænomenet som lige så revolutionerende som internettet.2

Tendensen understøttes af det forhold, at AI-chatbotten ChatGPT i januar 2023 – blot to måneder efter lancering – opnåede 100 millioner månedlige brugere på verdensplan, hvilket har gjort ChatGPT til den hurtigst voksende forbrugerapplikation nogensinde.3 På dets nuværende udviklingsstadie er ChatGPT så "intelligent", at systemet bl.a. er i stand til at bestå den amerikanske advokateksemnen med et resultat, der er placeret blandt de 10 % bedste, samt at bevare overblikket over mere end 25.000 ord i forbindelse med en korrespondance med en bruger.4

ChatGPT tilhører den særlige form for systemer, som er baseret på generativ kunstig intelligens (heri benævnt "generative AI"), som gør dem i stand til at generere nyt, "kreativt" indhold såsom noveller, digte, billeder og musik.5 Generative AI bliver udviklet på baggrund af machine learning (heri benævnt "ML"), hvorved en algoritme bliver trænet til at skabe en bestemt form for indhold ved at analysere og lære fra træningsdata, der består af store mængder af tilsvarende indhold.6

Generative AI-systemer har i høj grad revolutioneret den kreative branche og gjort den kreative skabelsesproces til allemandseje. AI-teknologiens meget hurtige fremmarch har imidlertid også affødt et utal af uafklarede juridiske, herunder en lang række ophavsretlige, problemstillinger. En af de

1 EU-Kommissionens meddelelse om AI (2018), s. 1
2 Liberatore, Stacy: Bill Gates says ChatGPT-like AI is 'as revolutionary as mobile phones and the internet' - and says it could speed up creation of new vaccines, Daily Mail (2023) - https://www.dailymail.co.uk/sciencetech/article-11896051/Bill-Gates-says-AI-revolutionary-mobile-phones-internet.html (lokaliseret d. 27. maj 2023 kl. 10:23)
3 Hu, Krystal: ChatGPT sets record for fastest-growing user base - analyst note, Reuters (2023) - https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/ (lokaliseret d. 27. maj 2023 kl. 11:01)
4 Koetsier, John: GPT-4 Beats 90% Of Lawyers Trying To Pass The Bar, Forbes (2023) - https://www.forbes.com/sites/johnkoetsier/2023/03/14/gpt-4-beats-90-of-lawyers-trying-to-pass-the-bar/?sh=1b5f48a53027 (lokaliseret d. 27. maj kl. 09:48)
6 Chiou (2019), s. 399
væsentligste ophavsretlige problemstillinger udspringer af det forhold, at træningen af en algoritme i praksis ofte bliver gennemført ved brug træningsdata, som består af tusindvis (og undertiden millionvis) af værker. Da den anvendte træningsdata altid bliver kopieret på forskellige tidspunkter ifm. træningen, er der tale om en brug af værkerne, som potentielt er omfattet af den eneret til eksemplarfremsættelse, som indehaverne af ophavsretten til disse værker nyder, jf. OPHL § 2, stk. 1.

Der er imidlertid ingen af de regler om ophavsret, der gælder i EU (og herunder i Danmark), som specifikt regulerer kopieringen af ophavsretligt beskyttede værker ifm. træningen af en ML-algoritme. Derudover har kopieringens forenelighed med de nugældende regler om ophavsret hverken været behandlet af EU-Domstolen (heri benævnt "EUD") eller af de nationale domstole i de enkelte EU-medlemsstater. I skrivende stund er det således uafklaret i praksis, hvornår kopieringen af værker ifm. træningen af en ML-algoritme kræver tilladelse fra indehaverne af ophavsretten til de værker, der er anvendt som træningsdata.

I lyset af den uafklarede ophavsretlige praksis, der hersker i dag, vil formålet med dette speciale være at undersøge, hvordan kopieringen af værker ifm. træningen af en ML-algoritme, der skal implementeres i et generative AI-system, skal behandles efter de nugældende regler om ophavsret, og hvilken betydning de nugældende regler har for adgangen til at kopiere værker ifm. træningen uden tilladelse fra rettighedshaverne. Dette vil blive klarlagt med udgangspunkt i nedenstående problemformulering.

1.2 Problemformulering

Hvornår er kopieringen af træningsdata, der finder sted ifm. træningen af en ML-algoritme mhp. implementering i et generative AI-system, omfattet af eneretten til eksemplarfremsættelse, jf. OPHL § 2, stk. 1, når den anvendte træningsdata består af ophavsretligt beskyttede værker, i hvilket omfang er kopieringen dækket af de lovbestemte indskrænkninger i ophavsretten, og er de nugældende indskrænkninger hensigtsmæssige?

1.3 Specialets struktur

Dette speciale er struktureret som følger:

---

7 Chiou (2019), s. 399. Butterick, Matthew: *Because AI needs to be fair & ethical for everyone*, Stable Diffusion litigation (2023) - [https://stablediffusionlitigation.com/](https://stablediffusionlitigation.com/) (lokalisert d. 7. marts. 2023 kl. 09:28)
8 Lovbekendtgørelse nr. 1144 af 23. oktober 2014 om ophavsret
9 Chiou (2019), s. 402. Axham (2021), s. 9
10 Se bl.a. EU-Kommissionens rapport om AI (2022), s. 21
I afsnit 2 vil der blive redegjort for *generative AI* og *ML*. Herunder vil det blive klart, hvordan kopieringen af træningsdata foregår i praksis. Derefter vil der i afsnit 3 blive redegjort for de regler om ophavsret, der er centrale for besvarelsen af dette speciale problemformulering.

I afsnit 4 vil problemformuleringens første led blive besvaret, og det vil blive undersøgt, hvornår den kopiering af træningsdata, der finder sted ifm. træningen af en *ML*-algoritme mhp. implementering i et *generative AI*-system, er omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1, når den anvendte træningsdata består af værker.

I afsnit 5 vil problemformuleringens andet led blive besvaret. Først vil de relevante lovbestemte indskrænkninger i ophavsretten blive identificeret. Dernæst vil indskrænkningerne blive analyseret hver for sig mhp. at fastslå, i hvilket omfang de dækker kopieringen af værker ifm. træningen.

Problemformuleringens tredje led vil blive besvaret i afsnit 6, hvor de nugældende indskrænkningers hensigtsmæssighed vil blive diskuteret fra en række forskellige perspektiver.

Endelig vil der blive opstillet en samlet konklusion i afsnit 7.

### 1.4 Afgrænsning

*Generative AI* giver anledning til utallige retlige spørgsmål inden for mange forskellige retsområder, herunder flere af de immaterialretlige. Dette speciale behandler dog udelukkende *generative AI* i lyset af ophavsretten.

Dette speciales problemformulering vil blive besvaret med udgangspunkt i dansk ret, hvorfor besvarelsen tager afsæt i de regler om ophavsret, der følger af OPHL. Da disse regler i vidt omfang er EU-harmoniserede, vil de EU-retlige regler om ophavsret ligeledes blive inddraget.11 Nationale regler fra andre lande end Danmark vil dog kun blive inddraget i begrænset omfang og kun mhp. at perspektivere de danske regler hertil.

Specialet har for det første til formål at undersøge, hvornår kopieringen af træningsdata, der finder sted ifm. træningen af en *ML*-algoritme mhp. implementering i et *generative AI*-system, er omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1, når den anvendte træningsdata består af værker.

---

11 Schønning (2021), s. 115ff. Schøvsbo *et al.* (2021), s. 56.
I praksis findes der mange forskellige teknikker, der kan anvendes til træningen af en ML-algoritme.\(^\text{12}\) For at undgå at besvarelsen af dette speciales problemformulering bliver rent hypotetisk, tager besvarelsen afsæt i tre specifikke tilfælde af kopiering af træningsdata, der i praksis almindeligvis finder sted i forbindelse med træningen af en ML-algoritme. Derudover behandler specialet udelukkende den træning af ML-algoritmer, som bliver gennemført af virksomheder (såsom Stability AI, jf. afsnit 2.2.2) til kommercielle formål.


For det andet har specialet til formål at undersøge, i hvilket omfang kopieringen af træningsdata, er dækket af de lovbestemte indskrænkninger i ophavsretten, når den anvendte træningsdata består af værker. Specialet behandler således ikke de ulovbestemte indskrænkninger (såsom parodiundtagelsen, der blev "genintroduceret" i dansk ophavsret i maj 2023\(^\text{13}\)). Det er desuden kun de lovbestemte indskrænkninger, der specifikt er relevante ift. de tre tilfælde af kopiering af træningsdata, der vil blive behandlet.

Endelig har specialet for det tredje til formål at diskutere, om de nugældende indskrænkninger er hensigtsmæssige. Det er alene de lovbestemte indskrænkninger, der specifikt er relevante for de tre tilfælde af kopiering af træningsdata, der er genstand for diskussionen.

Selvom reglernes hensigtsmæssighed selvsagt kan diskuteres fra mange forskellige perspektiver, tager diskussionen afsæt i tre specifikke perspektiver: et rettigheds- og brugerperspektiv, et formålsperspektiv og et globalt konkurrenceperspektiv.

Afslutningsvist skal det bemærkes, at specialet ikke behandler spørgsmålet om, hvorvidt et output, der er blevet skabt af en ML-algoritme, i sig selv er beskyttet af ophavsretten som et værk, jf. OPHL

---

\(^{12}\) Axhamn (2021), s. 39 og s. 80

\(^{13}\) Højesterets dom af 17. maj 2023 i sag BS-24506/2022-HJR, Tom Jensen mod Billedhuggeren Edvard Eriksens Arvinger I/S
§ 1. Derudover behandler specialet alene kopiering af værker ifm. træningen og således ikke den kopiering, der finder sted, efter at ML-algoritmen er blevet implementeret i et generative AI-system.

1.5 Metode

Til besvarelsen af dette speciales problemformulering anvendes den retsdogmatiske metode. Denne metode, som beskæftiger sig med en analyse og systematisering af de foreliggende retskilder, anvendes med det formål at beskrive gældende ret.14 I diskussionen i specialets afsnit 6 vil der dog også blive inddraget mere retspolitiske overvejelser, hvor der er anledning hertil.15

De retskilder, der er centrale for besvarelsen af problemformuleringen, er OPHL, InfoSoc-direktivet16 og DSM-direktivet17. Selvom et EU-direktiv skal implementeres i dansk ret, er dette endnu ikke sket, hvad angår DSM-direktivet.18 Derfor vil der både vil blive henvist til OPHL såvel som til direktiverne.

Da reglerne i OPHL i vidt omfang er EU-harmoniseret gennem direktiverne, vil der løbende blive inddraget afgørelser, hvorv EUD har besvaret præjudicielle spørgsmål fra EU’s medlemsstater om fortolkningen af direktivernes regler. Da EUD er tillagt kompetencen til med endelig og bindende virkning at afgøre præjudicielle spørgsmål om fortolkningen af reglerne i EU-retsakter, udgør sådanne afgørelser væsentlige fortolkningsbidrag ved fastlæggelsen af gældende ret.19 Da EUD dog ikke har kompetence til at fortolke nationale regler, vil der også blive inddraget afgørelser fra danske domstole.20 Der vil ligeledes blive inddraget enkelte administrative afgørelser fra Kulturministeriet, der belyser, hvordan OPHL bliver anvendt i praksis.21

Specialet inddrager derudover danske lovforarbejder til OPHL, da disse afspejler lovgivers hensigt med reglerne og de bagvedliggende hensyn, som reglerne har til formål at varetage.22

På samme vis vil der blive henvist til EU-forarbejderne til de to direktiver, der ligesom de danske lovforarbejder bidrager til fastlæggelsen af baggrunden for reglerne.23 Dog er det alene de indledende

14 Blume (2016), s. 40 f.
15 Ibid., s. 60
16 Europa-Parlamentets og Rådets direktiv 2001/29/EF af 22. maj 2001 om harmonisering af visse aspekter af ophavsrret og beslægtede rettigheder i informationssamfundet
17 Europa-Parlamentets og Rådets direktiv (EU) 2019/790 af 17. april 2019 om ophavsrret og beslægtede rettigheder på det digitale indre marked og om ændring af direktiv 96/9/EF og 2001/29/EF
18 Engsø Sørensen et al. (2019), s. 105
20 Engsø Sørensen et al. (2019), s. 241
21 Blume (2016), s. 248
22 Ibid., s. 185 og 208
23 Ibid., s. 185
præambelbetragtninger, der følger af direktiverne, og som derfor udgør en vedtaget del af disse retsakter, der tillægges en egentlig retskildeværdi i EU-retten.24

Selvom nationale regler og retsprækis fra andre lande ikke kan tillægges nogen retskildeværdi i dansk ret, vil sådanne retskilder dog blive inddraget i specialet.25 Inddragelsen af disse retskilder vil primært ske mhp. at perspektivere de danske regler hertil.

Endelig vil der blive henvist til meddelelser fra EU-Kommissionen, en række rapporter og publikationer fra EU's institutioner samt juridisk litteratur. Selvom der ikke er tale om bindende retskilder, udgør sådanne retskilder ikke desto mindre væsentlige fortolkningsbidrag ved fastlæggelsen af gældende ret.26

2 Generative AI, ML og kopiering af træningsdata

En forudsætning for at behandle kopieringen af træningsdata i lyset af reglerne om ophavsret, er en forståelse af, hvad de to teknologiske fænomener generative AI og ML er, og hvordan de udvikles. Dette vil blive klarlagt i det følgende, hvor det også vil blive fastslået, hvilken rolle træningsdata spiller i den forbindelse.

2.1 Hvad er generative AI?

For at forstå begrebet generative AI er det først nødvendigt at fastslå, hvad det overordnede begreb AI er.

Der findes ikke nogen almengyldig definition af begrebet AI.27 Ikke desto mindre findes der et utal af definitioner, med hvilke det er blevet forsøgt afklaret, hvad det er for et fænomen, begrebet dækker over.28

EU-Kommissionen har fx ved flere lejligheder defineret begrebet.29 Det samme har WIPO og EUIPO.30 Disse definitioner kan sammenfattes således, at begrebet AI omfatter systemer, der ved at kombinere data, algoritmer og computerkraft, udviser en form for intelligent adfærd ved at analysere

24 Ibid., s. 185 og 279
25 Ibid., s. 281
26 Blume (2016), s. 266. Engsig Sørensen et al. (2019), s. 111
27 EU-Kommissionens rapport om AI (2020), s. 21. Veale et al. (2019), s. 24
28 Axhamn (2021), s. 36
30 WIPO (2020), s. 3f. EUIPO (2022), s. 152f.
og lære fra omgivelserne med en lav grad af menneskelig påvirkning for derved at opnå specifikke mål.

Den første egentlige legaldefinition af begrebet AI følger i øvrigt af EU-Kommissionens forslag til AI-forordningen31 fra begyndelsen af 2021, som er verdens første bud på regler, der specifikt regulerer AI.32 AI-systemer bliver i forordningens art. 3, nr. 1, defineret som "software [...] som med henblik på et givet sæt mål, der er fastsat af mennesker, kan generere output såsom indhold, forudsigelser, anbefalinger eller beslutninger, der påvirker de miljøer, de interagerer med". Selvom forordningsforslaget fortsat er under behandling, og forholdet mellem kunstig intelligens og ophavsretten ikke er reguleret af forslaget i dets nuværende udformning33, er det dog relevant for fastlæggelsen af begrebet AI, da det afspejler EU-Kommissionens regulatoriske tilgang til fænomenet.

Som det fremgår af definitionerne, er AI-systemer bl.a. karakteriseret ved, at deres evne til at generere et givent output beror på deres evne til selvstændigt at lære fra deres omgivelser.34 Derved adskiller sådanne systemer sig fra traditionelle computerprogrammer, der kun kan handle på baggrund af de instruktioner, som programmøren har fastsat i computerkoden.35 På trods heraf kan AI ikke betragtes som fuldstændig autonom. På sit nuværende udviklingsstadium er AI kun i stand til at håndtere få opgaver, og det kræver altid en vis grad af menneskelig involvering.36

Med det overordnede begreb AI nogenlunde klarlagt kan det herefter fastslås, at generative AI er et underbegreb, der anvendes om AI-systemer, som (for at bruge termerne fra definitionen i AI-forordningen) er i stand til at generere et output i form af "nyt" indhold som tekst, billeder eller musik; den type indhold, som man almindeligvis ville betegne som "kreativt", hvis det var blevet skabt af et menneske, og som ville kunne henføres til den slags genstande, der kan beskyttes som et værk iht. OPHL, jf. nedenfor i afsnit 3.1.1.37

31 Forslag til Europa-Parlamentets og Rådets forordning om harmoniserede regler for kunstig intelligens (retsakten om kunstig intelligens) og om ændring af visse af unionens lovgivningsmæssige retsakter, COM(2021) 206, 21. april 2021
33 Reglerne i forslaget har til formål at sørge for, at de AI-systemer der bliver bragt i omsætning anvendt i EU er pålidelige og sikre samt i overensstemmelse med grundlæggende rettigheder og EU's værdier – Forslag til AI-forordningen, afsnit 1.1
34 EU-Kommissionens rapport om AI (2018), s. 19.
35 Thomas (2019), s. 2
36 På sit nuværende udviklingsstadium betegnes AI derfor som "Narrow AI" - High-Level Expert Group om AI (2019), s. 5
37 Axhamn (2021), s. 41
Som anført i dette speciales indledning, jf. afsnit 1.1, er et aktuelt eksempel på et generative AI-system ChatGPT, som er i stand til at generere et output i form af tekst, der modsvarer de skriftlige input, som brugeren indsætter i systemet. Fx kan man bede ChatGPT om at generere et digt om arvingerne til den skulptør, der skabte statuen af Den Lille Havfrue (et eksempel der dog sår tvivl om, hvorvidt Chat GPT er bekendt med Højesterets nyligt afsagte dom):

Et andet eksempel på et generative AI-system er Stable Diffusion, som er i stand til at generere et output i form af billeder på baggrund af brugerens skriftlige input. Man kan fx bede Stable Diffusion om at generere billeder, som forestiller statuen af Den Lille Havfrue:

---

38 ChatGPT – [https://openai.com/blog/chatgpt](https://openai.com/blog/chatgpt)
39 Højesterets dom af 17. maj 2023 i sag BS-24506/2022-HJR, Tom Jensen mod Billedhuggeren Edvard Eriksens Arvinger I/S
40 Screenshot af digt genereret ved brug af Chat GPT ([https://chat.openai.com](https://chat.openai.com)) d. 10. februar 2023 kl. 16:15
41 Stable Diffusion Online – [https://stablediffusionweb.com/](https://stablediffusionweb.com/)
2.2 **Udvikling af Generative AI på baggrund af ML**

*Generative AI* bliver som anført i dette speciales indledning, jf. afsnit 1.1, udviklet på baggrund af *ML* som indebærer, at man "træner" en algoritme\(^43\) til at generere det tilsigtede output ved at lade algoritmen analysere og "lære" fra indhold af samme art som det output, algoritmen skal trænes til at generere.\(^44\)

Selvom det skal bemærkes, at der findes mange forskellige teknikker, der i praksis kan anvendes til at træne en *ML*-algoritme\(^45\), er der dog en række elementer, som almindeligtvis går igen.

---

\(^42\) Screenshot af digt genereret ved brug af *Stable Diffusion* ([https://stablediffusionweb.com/#demo](https://stablediffusionweb.com/#demo)) d. 9. februar 2023 kl. 10:36

\(^43\) En algoritme er et sæt af instrukser eller regler, der er baseret på en matematisk formel, som fastslår den trinvise række af specifikke handlinger, en computer skal foretage for at realisere et givent mål – se Schovsbo *et al.* (2021), s. 79 og Drexl *et al.* (2019), s. 4

\(^44\) Chiou (2019), s. 399. *High-Level Expert Group om AI* (2019), s. 3f.

\(^45\) Blandt disse teknikker kan nævnes *supervised learning*, *unsupervised learning* og *reinforcement learning* – *High-Level Expert Group om AI* (2019), s. 3f. Axhamn (2019), s. 39
Treningen af en ML-algoritme bliver indledt ved, at udviklerne indsamler det indhold (fx ophavsretligt beskyttede værker), som algoritmen skal analysere.\textsuperscript{46} I ML-kontekst kaldes indholdet for "træningsdata", som derfor er en term, der løbende vil blive anvendt synonymt dermed.\textsuperscript{47} Den indsamlede træningsdata bliver lagret i det computersystem, som træningen skal gennemføres i.\textsuperscript{48} I den forbindelse bliver den pågældende data undertiden konverteret til et filformat, som kan processeres af algoritmen.\textsuperscript{49}

Som allerede indikeret ovenfor opnår ML-algoritmen sin evne til at generere en bestemt form for indhold ved at analysere eksisterende indhold af samme art.\textsuperscript{50} Derfor er træningsdata det vigtigste element ved gennemførelsen af træningen.\textsuperscript{51} Jo mere data, der er anvendt til træningen, jo mere præcist bliver det output, som algoritmen er i stand til at generere.\textsuperscript{52} Derfor bliver træningen af en ML-algoritme, som skal implementeres i et generative AI-system, gennemført på baggrund af tusindvis – og undertiden millionvis, jf. \textit{Stable Diffusion}-eksemplet i afsnit 2.2.2 nedenfor – af billeder, sange o.l.\textsuperscript{53} Indsamlingen af træningsdata sker undertiden på basis af såkaldt web scraping, hvor udviklerne lader en digital robot scanne et hav af internetsider mhp. at indsamle indhold, som er tilgængeligt på disse.\textsuperscript{54} Dog bliver træningen undertiden også gennemført ved brug af data fra andre end udviklernes allerede eksisterende træningsdatasets (såsom \textit{Laion-5B}, jf. nedenfor).

Algoritmen analyserer herefter den lagrede træningsdata mhp. at identificere bestemte karakteristika (fx kreative træk) eller sammenhænge/fællestræk deri.\textsuperscript{55}

Algoritmen er baseret på en matematisk funktion, og denne funktion indeholder en række parametre/variabler, som er justérbare.\textsuperscript{56} Under algoritmens analyse selvfjunder og optimerer den

\textsuperscript{46} Margoni (2018), s. 2. Chiou (2019), s. 402
\textsuperscript{47} Chiou (2019), s. 399
\textsuperscript{48} \textit{Ibid.}, s. 409
\textsuperscript{49} Margoni (2018), s. 2. Chiou (2019), s. 402
\textsuperscript{50} Chiou (2019), s. 399f. \textit{High-Level Expert Group om AI} (2019), s. 3
\textsuperscript{51} Drexel \textit{et al.} (2019), s. 8
\textsuperscript{52} \textit{Ibid.}
\textsuperscript{53} Chiou (2019), s. 399.
\textsuperscript{55} Chiou (2019), s. 403
\textsuperscript{56} Drexel \textit{et al.} (2019), s. 6f.
løbende disse parametre på baggrund af dens analyse af træningsdata.\textsuperscript{57} Algoritmen gør derved sig selv mere og mere præcis i løbet af træningen.\textsuperscript{58}

På baggrund af analysen generer algoritmen undertiden den form for output, den skal lære at skabe.\textsuperscript{59} Dette er bl.a. en måde, hvorpå udviklerne kan vurdere, om algoritmen har lært det den skulle, eller om der er behov for at finjustere algoritmen.\textsuperscript{60}

Når udviklerne vurderer, at outputtet er af en tilfredsstillende kvalitet, afbryder de træningen, og den trænede \textit{ML}-algoritme (den såkaldte "\textit{ML-model}"	extsuperscript{61}) kan herefter implementeres i et \textit{generative AI}-system.\textsuperscript{62} Det skal i den forbindelse bemærkes, at den analyserede træningsdata ikke bliver lagret i modellen, da denne i bund og grund blot er en matematisk funktion, der fortæller \textit{generative AI}-systemet, hvordan det genererer outputtet.\textsuperscript{63}

\subsection{ML indebærer en kopiering af træningsdata}

Som anført ovenfor findes der mange forskellige teknikker, der kan anvendes til at træne en \textit{ML}-algoritme. Ikke desto mindre vil det indhold (herunder ophavsretligt beskyttede værker), som er anvendt som træningsdata, almindeligvis blive kopieret på tre specifikke tidspunkter i løbet af træningen. Dette speciales problemformulering vil som fastslået i afsnit 4.1 blive besvaret med afsæt i disse tre tilfælde af kopiering.

\textbf{Indledningsvist} bliver indholdet kopieret, når det bliver indsamlet og lagret i det computersystem, som træningen af \textit{ML}-algoritmen skal gennemføres i.\textsuperscript{64}

\textbf{Dernæst} bliver indholdet kopieret, når algoritmen analyserer den anvendte træningsdata.\textsuperscript{65} Som anført ovenfor varierer det i praksis, hvilke elementer af det pågældende indhold algoritmen analyserer og så at sige "\textit{leder efter}", hvorfor der både kan være tale om kopiering af dele af indholdet såvel som af indholdet i dets helhed.\textsuperscript{66}

\begin{thebibliography}{99}
\bibitem{Drexl2019} Drexl \textit{et al.} (2019), s. 7
\bibitem{Ginsburg2019} Ginsburg \textit{et al.} (2019), s. 401
\bibitem{Drexl2019b} Kretz (2019)
\bibitem{Drexl2019c} Drexl \textit{et al.} (2019), s. 5. \textit{High-Level Expert Group om AI} (2019), s. 3
\bibitem{Kretz2019b} Kretz (2019)
\bibitem{Drexl2019d} Drexl \textit{et al.} (2019), s. 5 og 8
\bibitem{Chiou2019} Chiou, s. 402
\bibitem{Axhamn2021} Axhamn (2021), s. 50. Chiou (2019), s. 403
\bibitem{Study2014} \textit{Study on the legal framework of TDM} (2014), s. 31
\end{thebibliography}
Endelig bliver indholdet undertiden kopieret, når algoritmen genererer et output ifm. træningen.67

Som det vil fremgå af *Stable Diffusion*-eksemplerne, der løbende vil blive opstillet i dette speciale, kan algoritmens output enten helt eller delvist bestå af indhold, den har analyseret.

2.2.2 Et eksempel på kopieringen af træningsdata ifm. træningen - *Stable Diffusion*

For at illustrere, hvordan kopieringen af træningsdata foregår i praksis, vil en række af de tilfælde af kopiering, der fandt sted ifm. træningen af *ML*-algoritmen, som *Stable Diffusion* er baseret på, blive udfoldet i det følgende.

*Stable Diffusion*-algoritmen blev trænet på baggrund af indhold fra træningsdatasættet *LAION*-5B, som blev udarbejdet af *LAION*, en tysk non-profit-organisation, som er støttet af det britiske selskab *Stability AI*, der står bag udviklingen af *Stable Diffusion*.68 *LAION*-5B består af knap seks milliarder billede, der er *scraped* fra internettet.69 Blandt billede indgik flere millioner beskyttede værker, som var blevet kopieret og lagret i datasættet.70

Under træningen analyserede algoritmen billeder fra *LAION*-5B i disse oprindelige skikkelse samt i en række versioner, hvor billede var tilført forskellige mængder af visuel støj. Disse forskellige versioner kan illustreres som følger:

---

67 Se fx Karpathy *et al.* (2016)
68 Harman (2023)
70 Butterick (2023). Harman (2023)
71 Eksempel på "Forward Diffusion" med kattebillede, der er tilføjet forskellige mængder af visuel støj – https://stable-diffusion-art.com/how-stable-diffusion-work/#!text=Stable%20Diffusion%20is%20a%20latent,why%20it's%20a%20lot%20faster (lokaliseret d. 7. marts 2023 kl. 15:47)
Algoritmen analyserede de pågældende billeder mhp. at forudsige, hvor meget visuel støj, der var tilført hvert enkelt billede, og på baggrund lærte algoritmen at identificere den konkrete mængde af støj, der var tilført ethvert nyt billede, den blev præsenteret for.\(^\text{72}\)

Derfor analyserede algoritmen en række nye billeder, der var tilført varierende mængder af visuel støj. Ved at anvende sin evne til at forudsige mængden af støj genskabte algoritmen de støjfyldte billeder i deres oprindelige skikkelse ved at fjerne den estimerede mængde af støj.\(^\text{73}\) Således genererede algoritmen altså en kopi af billederne i deres oprindelige skikkelse.

I forlængelse af ovenstående skal det bemærkes, at den ovenfor beskrevne proces kun udgjorde en bestanddel af den samlede træning af algoritmen.\(^\text{74}\) Man bliver derfor ikke meget klogere på, hvordan Stable Diffusion fungerer ved at betragte ovenstående proces isoleret fra resten af træningen. Eksemplet giver dog et indblik i, hvordan kopieringen af træningsdata kan foregå i praksis.

Afslutningsvist skal det anføres, at den kopiering af værker, som fandt sted ifm. træningen af Stable Diffusion-algoritmen, er årsagen til, at der i skrivende stund veriserer retssager mod Stability AI i USA og i Storbritannien.\(^\text{75}\) Retssagerne, som blev anlagt i begyndelsen af 2023 af rettighedshaverne til nogle af de værker, der blev anvendt til træningen af algoritmen, angår spørgsmålet om, hvorvidt kopieringen af disse værker ifm. træningen indebar en ophavsretningskrænkelse.\(^\text{76}\)

En af rettighedshaverne, selskabet Getty Images, har i stævningen i den amerikanske sag fremlagt et eksempel på et output, som er genereret af Stable Diffusion-algoritmen. Dette illustrerer, at algoritmen er blevet trænet på baggrund af billeder, som Getty Images ejer rettighederne til. Selv det vandmærke, som er påført billederne i de tilfælde, hvor en bruger ikke har betalt for en licens, går igen i algoritmens output:

---

\(^{72}\) Butterick, Matthew: Because AI needs to be fair & ethical for everyone, Stable Diffusion litigation (2023) – https://stablediffusionlitigation.com/ (lokalisert d. 7. marts. 2023 kl. 09:28)

\(^{73}\) Ibid.

\(^{74}\) Ibid.


\(^{76}\) Ibid.
3 Retsgrundlaget - den ophavsretlige regulering, eneretten og indskrænkningerne heri

I det følgende vil retsgrundlaget, som danner grundlag for besvarelsen af dette speciales problemformulering, blive beskrevet. Herunder vil samspillet mellem den ophavsretlige eneret og indskrænkningerne i denne ret blive klarlagt.

3.1 Ophavsretten til litterære og kunstneriske værker

I dansk ret findes reglerne om ophavsret først og fremmest i OPHL.

Reglerne i OPHL er i vidt omfang EU-harmoniseret ved en efterhånden større mængde direktiver om ophavsret, som er implementeret i OPHL. De harmoniserede regler skal derfor fortolkes i overensstemmelse med EU-retten, som særligt på det ophavsretlige område sjældent tildeler medlemsstaterne nogen fortolkningsfrihed. Dette gælder bl.a. i relation til de regler, der følger af InfoSoc-direktivet og DSM-direktivet.

---

78 I skrivende stund er der gennemført tolv direktiver om ophavsret og beslægtede rettigheder på EU-plan. Se Schønning (2021), s. 115ff. og Schovsbo et al. (2021), s. 56.
79 Rosenmeier et al. (2018), s. 294. Schovsbo et al. (2021), s. 59
Dansk ophavsret er derudover i høj grad influeret af international ret som følge af Danmarks tiltrædelse af en række internationale traktater og aftaler.\(^{81}\)

Endelig er reglerne i OPHL også præget af et historisk tæt nordisk lovsamarbejde, hvorfor ophavsretslovene i de øvrige nordiske lande udgør væsentlige fortolkningsbidrag.\(^{82}\)

### 3.1.1 Beskyttelsens genstand

OPHL beskytter litterære og kunstneriske værker og tildeler den, som har frembragt et sådant værk (ophavsmanden), ophavsretten til værket, jf. OPHL § 1, stk. 1. Selvom ophavsretten alene kan opstå hos fysisk person, kan retten dog overdrages helt eller delvist i visse tilfælde, jf. OPHL § 53, stk. 1, til en fysisk såvel som en juridisk person.\(^{83}\)

For at en frembringelse kan beskyttes som et værk, skal den opfylde to krav.

For det første skal værket efter sin art kunne betegnes som litterært eller kunstnerisk, hvilket vurderes med udgangspunkt i de værkstyper, der er opregnet (ikke udtømmende) i OPHL § 1.\(^{84}\)

For det andet skal frembringelsen være original. Kravet om originalitet er EU-harmoniseret, hvorfor en frembringelse er original i det omfang, at den kan siges at være udtryk for ophavsmandens "egen intellektuelle frembringelse" og altså et resultat af ophavsmandens frie og kreative valg.\(^{85}\) Dette udelukker ophavsretlig beskyttelse af frembringelser, som er skabt på baggrund af ophavsmandens valg, derved noterende et resultat af ophavsmandens frie og kreative valg.\(^{86}\)

### 3.1.2 Beskyttelsens indhold

Ophavsretten til et værk medfører en række økonomiske rettigheder, jf. OPHL § 2, samt en række ideelle rettigheder, jf. § 3. Som anført i dette speciales afgrænsning, jf. afsnit 1.4, beskæftiger dette speciale sig alene med § 2-rettighederne.

Iht. OPHL § 2 er indehaveren af ophavsretten til et værk tildelt en eneret til at fremstille eksemplarer af værket samt til at gøre det tilgængeligt for almenheden ved spredning, visning eller offentlig fremførelse. Det kræver derfor rettighedshaverernes tilladelse, når andre end indehaveren af ophavsretten bruger værket på en af de måder, der er omfattet af OPHL § 2.\(^{87}\) Eneretten omfatter ikke

---

\(^{81}\) Bl.a. TRIPS-aftalen, Bernerkonventionen, WIPO Copyright Treaty – Schovsbo et al. (2021), s. 64
\(^{82}\) Schønning (2021), s. 107. Schovsbo et al. (2021), s. 63.
\(^{83}\) Schønning (2021), s. 126
\(^{84}\) Schønning (2021), s. 131. Schovsbo et al. (2021), s. 73
\(^{85}\) Infopaq I, pr. 35. Painer, pr. 87 og pr. 89. Cofemel, pr. 29f. Schovsbo et al. (2021), s. 69
\(^{86}\) Schovsbo et al. (2021), s. 70
\(^{87}\) Schønning (2021), s. 176
blot værket i oprindeligt men også i ændret skikkelse, jf. ligeledes OPHL § 2, stk. 1, hvorfor indehaveren af ophavsretten er beskyttet mod efterligninger af værket.\textsuperscript{88}

### 3.2 Indskrænkninger i ophavsretten

Der gælder en række indskrænkninger (eller undtagelser, om man vil) i eneretten efter OPHL § 2. Der er altså visse former for brug af et værk, som ikke kræver tilladelse fra indehaveren af ophavsretten, på trods af at der er tale om en brug (en såkaldt ophavsretligt relevant handling\textsuperscript{89}), som kan henføres til de handlinger, der følger af § 2.\textsuperscript{90}

Indskrænkningerne i eneretten til eksemplarafremstilling og overføring til almenheden er EU-harmoniserede, og medlemsstaterne kan derfor kun indføre de typer af indskrænkninger, der følger af den udtømmende opregning i InfoSoc-direktivets art. 5(2) og (3).\textsuperscript{91}

De lovbestemte indskrænkninger, der gælder efter dansk ophavsret, følger af OPHL kap. 2.\textsuperscript{92} Der findes tre typer af lovbestemte indskrænkninger: i) fribrugsregler\textsuperscript{93}, ii) regler om aftalelicens\textsuperscript{94} og iii) regler om tvangslicens\textsuperscript{95}.\textsuperscript{96}

Når en dansk domstol fortolker en indskrænkning i ophavsretten, skal dette ske i overensstemmelse med den såkaldte "tretrinstest", som bl.a. følger af InfoSoc-direktivets art. 5(5) og DSM-direktivets 6. præambelbetragtning.\textsuperscript{97} Denne fastslår, at indskrænkningerne kun kan anvendes i visse specielle tilfælde, som ikke strider mod den normale udnyttelse af værker, og som ikke indebærer urimelig skade for rettighedshaveres legitime interesser.\textsuperscript{98}

Desuden følger der af lovens § 11 en række principper for anvendelsen af bestemmelserne i OPHL kap. 2.

---

88 Schønning (2021), s. 183. Schovsbo et al. (2021), s. 152
89 Schønning (2021), 177
90 Ibid., s. 277
91 Schovsbo et al. (2021), s. 170.
92 Ibid., s. 169
93 Værket må bruges vederlagsfrit uden samtykke fra indehaveren af ophavsretten – Schovsbo et al. (2021), s. 170
94 Værket må bruges mod betaling af vederlag uden samtykke fra indehaveren af ophavsretten – Schovsbo et al. (2021), s. 171
95 Værket må bruges uden at spørge rettighedshaver om tilladelse, hvis brugen er klareret gennem en aftale med en repræsentativ rettighedsorganisation – Schovsbo et al. (2021), s. 171
96 Schovsbo et al. (2021), s. 170f. Rosenmeier (2018), s. 97
97 Schovsbo et al. (2021), s. 172. U 2019.1109 H (Würtz)
98 Jf. bl.a. InfoSoc-direktivets art. 5(5) og DSM-direktivets 6. præambelbetragtning
Det følger af § 11, stk. 1, at de ideelle rettigheder iht. § 3 skal respekteres ved anvendelsen af en bestemmelse i OPHL kap. 2. Dette princip vil dog ikke blive behandlet yderligere, da dette speciale ikke behandler de ideelle rettigheder, jf. afgrænsningen i afsnit 1.4.

Det følger af § 11, stk. 2, at et værk, som anvendes i overensstemmelse med en bestemmelse i kap. 2, "ikke må ændres i videre udstrækning, end den tilladte brug kræver". Denne bestemmelse skal fortolkes restriktivt, hvorfor der således kun er adgang til at foretage begrænsede ændringer (såsom beskæringer), som er nødvendige af praktiske grunde.99

Derudover følger det af § 11, stk. 3, at det ikke er tilladt at fremstille eksemplarer på grundlag af en ulovlig gengivelse af værket, når et værk anvendes i overensstemmelse med en bestemmelse i kap. 2. Der gælder således et krav om "lovligt forlæg", hvilket indebærer, at eksemplarfremstilling kun må ske på grundlag af en gengivelse eller et eksemplar af et værk, som er godkendt af indehaveren af ophavsretten, eller som er omfattet af en indskrænkning i ophavsretten.100

Endelig følger det af § 11, stk. 3, at eksemplarfremstillingen derudover ikke må ske på grundlag af en omgåelse af en teknisk foranstaltning i strid med § 75 c, stk. 1101.

3.3 **Regler (eller mangel på samme?) om brug af værker til træningen af en algoritme**

Selvom flere af reglerne om ophavsret, der gælder i Danmark, er målrettet de udfordringer, som digitale teknologier skaber (hvilket i særdeleshed gælder de regler, der implementerer InfoSoc-direktivet102), er der dog som anført i dette speciales indledning, jf. afsnit 1.1, i skrivende stund ingen af disse regler, der specifikt og udelukkende gælder for den brug af værker, som finder sted ifm. træningen af en *ML*-algoritme.

Det tætteste man kommer på en sådan regel, er den tilføjelse til EU-Kommissionens forslag til AI-forordningen, som EU-Parlamentet efter sigende vil søge at få igennem. Tilføjelsen pålægger en

---

99 Schønning (2021), s. 288. Lund (1961), s. 183
100 Schønning (2021), s. 289. Schovsbo et al. (2021), s. 176.
101 En teknisk foranstaltning, jf. § 75 c, stk. 1, er enhver form for effektiv teknisk foranstaltning, der under deres normale funktion har til formål at beskytte værker, som beskyttes iht. denne lov. Dette kan fx være foranstaltninger såsom adgangskontrol- eller beskyttelsesforanstaltninger såsom kryptering af værket eller en kopikontrolanordning, jf. InfoSoc-direktivets art. 6(3), som er gennemført i OPHL § 75 c, jf. L 19 2002-03.
102 InfoSoc-direktivets 5. præambelbetragtning. Udsen (2021), s. 46

21
udvikler af et generative AI-system at offentliggøre en fortegnelse over de ophavsretligt beskyttede værker, som er blevet brugt til træningen af ML-algoritmen.  

Derudover har brugen af beskyttede værker til træningen af en ML-algoritmes forenelighed med de eksisterende regler om ophavsret på nuværende tidspunkt som anført i dette speciales indledning, jf. i afsnit 1.1, hverken har været behandlet af EUD eller de nationale domstole i EU’s medlemsstater.  

Kopieringen af træningsdata, når denne data består af værker, må derfor behandles i lyset af de eksisterende regler, der ikke relaterer sig specifikt til ML samt retspraksis herom.  

4 Kopiering af træningsdata – en ophavsretligt relevant handling?  

I dette afsnit vil det blive undersøgt, hvornår kopieringen af træningsdata, der finder sted ifm. træningen af en ML-algoritme, som skal implementeres i et generative AI-system, når denne data består af værker, er omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1.  

Som anført i afsnit 1.4 vil besvarelsen af dette speciales problemformulering tage afsæt i de tilfælde af kopiering, der blev beskrevet i afsnit 2.2.1.  

4.1 Eksemplarfremstilling på en hvilken som helst måde og i en hvilken som helst form  

Som beskrevet i afsnit 3.2.1 ovenfor har indehaveren af ophavsretten til et værk en eneret til eksemplarfremstilling, jf. OPHL § 2, stk. 1. Dette er en eneret til at fremstille en fysisk "ting", der gør det muligt at opleve værket.  

Som anført i afsnit 3.2.1 omfatter eneretten både værket i oprindelig og i ændret skikkelse, jf. OPHL § 2, stk. 1, hvorfor rettighedshaver også er beskyttet mod eksemplarfremstilling af efterligninger af værket.  

Det følger af OPHL § 2, stk. 2, at eneretten til eksemplarfremstilling iht. bestemmelsens stk. 1, omfatter "[...] enhver direkte eller indirekte, midlertidig eller permanent og hel eller delvis eksemplarfremstilling på en hvilken som helst måde og i en hvilken som helst form."  

Definitionen i OPHL § 2, stk. 2, blev indført i dansk ophavsret gennem vedtagelsen af lovforslaget L 19 2002-03, der (bl.a.) havde til formål at gennemføre InfoSoc-direktivet i dansk ret.  


104Se bl.a. EU-Kommissionens rapport om AI (2022), s. 21  

105Schønning et al. (2021), s. 120  

106Schønning (2021), s. 180f. Schønning et al. (2021), s. 152  

107Schønning (2021), s. 191
der følger af OPHL § 2, stk. 2, stammer efter sin ordlyd fra direktivets art. 2, som regulerer eneretten til "reproduktion" af et værk.


Det skal i den forbindelse bemærkes, at selvom begrebet "reproduktion" er anvendt i InfoSoc-direktivets art. 2, hvor begrebet "eksemplarfremstilling" derimod er anvendt i OPHL § 2, stk. 1, er der tale om synonyme begreber. 111 Det kan hertil anføres, at selvom ingen af de to bestemmelser efter deres ordlyd udtrykkeligt nævner begrebet "kopiering", er dette begreb også synonymt med de to øvrige begreber. Det fremgår således udtrykkeligt i det danske lovforlæg, at eneretten til eksemplarfremstilling beskytter rettighedshaver mod "enhver form for kopiering". 112 Begreberne eksemplarfremstilling, reproduktion og kopiering vil derfor blive anvendt i flæng i dette speciale.

Det følger af InfoSoc-direktivets 21. præambelbetragsning, at direktivet "[...] definerer de handlinger, der er omfattet af reproduktionsretten". Ifølge EU-lovgiver er det "[...] nødvendigt med en bred definition af disse handlinger for at sikre retssikkerhed i det indre marked". 113 Der er således ingen tvivl om, at EU-lovgivers formål med definitionen var at understrege, at vitterligt enhver tænkelig form for reproduktion af værket er omfattet af eneretten til reproduktion.

Det bemærkes, at EUD tydeligvis også fortolker begrebet "reproduktion" i overensstemmelse med EU-lovgivers hensigt om en vidtrækkende definition af begrebet. I Infopaq I fastslog EUD med

---


109 SENA, pr. 23 og SGAE, pr. 31

110 *Infopaq I*, pr. 27

111 Schønning (2021), s. 179

112 L 19 2002-03, s. 238

113 InfoSoc-direktivets 21. præambelbetragsning
henvisning til bl.a. direktivets 21. præambelbetraktning som gengivet ovenfor, at beskyttelsen efter bestemmelsen skal have "en vid udstrækning". ¹¹⁴


Infopaq's virksomhed bestod bl.a. i udarbejdelsen af resuméer af avisartikler, som blev sendt ud til Infopaq's kunder med e-mail.¹¹⁶ For at finde frem til de relevante publikationer bearbejdede Infopaq artiklerne gennem en såkaldt "datafangstproces".

I første fase blev artiklerne registreret manuelt i Infopaq's elektroniske database.¹¹⁷ I den anden fase blev den del af artiklen, som skulle bearbejdes, indscannet til en computer og omdannet til en såkaldt TIFF-fil.¹¹⁸ I tredje fase blev TIFF-filen konverteret til en tekstfil, som kunne processeres af computeren.¹¹⁹ I den fjerde fase analyserede Infopaq's computersystem tekstfilen for at finde bestemte søgeord. Når systemet fandt et match, lagrede systemet et tekstuddrag bestående af søgeordet samt de fem forudgående og efterfølgende ord i computerens hukommelse (i alt 11 ord), hvorefter tekstfilen blev slettet.¹²⁰ I den femte og sidste fase udprintede systemet en fysisk version af tekstuddraget.¹²¹

Parterne i den danske sag var uenige om, hvorvidt den lagring og efterfølgende udprintning af tekstuddraget på 11 ord, som skete i datafangstprocessens fase 4 og 5, var omfattet af InfoSoc-direktivets art. 2.¹²²

EUD fandt, at lagringen samt det efterfølgende udprint af tekstuddrag på 11 ord fra avisartiklerne kunne henføres under begrebet "delvis reproduktion" i art. 2 i det omfang, at de gengivne elementer i den forbindelse var udtryk for ophavsmandens "egen intellektuelle frembringelse".¹²³ EUD overlod vurderingen af dette forhold til Højesteret, der fandt, at kopieringen var omfattet af eneretten til

¹¹⁴ Infopaq I, pr. 41-43
¹¹⁵ Danske Dagblades Forening
¹¹⁶ U 2013.1782 H (Infopaq)
¹¹⁷ Infopaq I, pr. 17
¹¹⁸ Ibid., pr. 18
¹¹⁹ Ibid., pr. 19
¹²⁰ Ibid., pr. 20
¹²¹ Ibid., pr. 21
¹²² Ibid., pr. 25
¹²³ Ibid., pr. 51
eksemplarfremstilling iht. OPHL § 2, stk. 1, da de pågældende kopier undertiden ville omfatte sætningsled, der var ophavsretligt beskyttet.\footnote{U 2013.1782 H (Infopaq)}

På baggrund af \textit{Infopaq I} kan det fastslås, at den brede definition af eneretten til reproduktion, som følger af art. 2, fører til, at kopiering af beskyttede værker og lagring af disse i en computer udgør eksemplarfremstilling, jf. OPHL § 2, stk. 1, i det omfang, at dø, der er blevet kopieret fra værket, er udtryk for ophavsmandens "\textit{egen intellektuelle frembringelse}". Således skal dø, der bliver kopieret, i sig selv opfylde originalitetskravet, da dette krav som fastslået i afsnit 3.1.1 er EU-harmoniseret og derfor er opfylt i det omfang, at der netop er tale om en ophavsmands "\textit{egen intellektuelle frembringelse}". Dette må derfor utvivlsomt også gøre sig gældende i relation til kopieringen af træningsdata, når denne data består af værker, der som anført i afsnit 2.2.1 finder sted på tre tidspunkter ifm. træningen af en \textit{ML}-algoritme, og som netop indebærer en digital kopiering af værkerne i en computer.

4.1.1 Kopiering af værker i oprindelig skikkelse ifm. træningen

I lyset af ovenstående kan det uden de store vanskeligheder fastslås, at kopieringen af træningsdata, når denne data består af værker, er omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1, og derfor som udgangspunkt kræver tilladelse fra rettighedshaverne, hvis der er tale om en kopiering af værkerne i deres oprindelige skikkelse og helhed.

Foruden at dette følger udtrykkeligt af OPHL § 2, stk. 2, der fastslår, at eneretten til eksemplarfremstilling af et værk i dets oprindelige skikkelse, jf. stk. 1, omfatter "hel eksemplarfremstilling", følger dette ligeledes af \textit{Infopaq I}, hvor EUD fastslog, at eneretten til reproduktion er en eneret til at tillade eller forbyde reproduktion af et "værk".\footnote{\textit{Infopaq I}, pr. 33} Da der er tale om en kopiering af hele værket i dette værks fuldstændige og oprindelige skikkelse, er der selvsagt tale om kopiering af en genstand, der i sin helhed er udtryk for ophavsmandens "\textit{egen intellektuelle frembringelse}".

Derfor vil det fx som udgangspunkt kræve tilladelse fra indehaveren af ophavsretten til et værk, når en \textit{ML}-algoritme som led i træningen genskaber dette værk i dets oprindelige skikkelse ved at fjerne den visuelle støj, der er tilført dette værk, ligesom det gjorde sig gældende i \textit{Stable Diffusion}-eksemplet i afsnit 2.2.2 ovenfor.

\footnote{\textit{Infopaq I}, pr. 33}
Kopieringen kræver dog ikke tilladelse fra rettighedshaverne, hvis den er dækket af en indskrænkning i ophavsretten, hvilket vil blive undersøgt i afsnit 5.

Derimod synes billedet ikke at være lige så klart i de tilfælde som beskrevet i afsnit 2.2.1, hvor der blot er tale om en kopiering af enkelte elementer af et beskyttet værk i dets oprindelige skikkelse.

Som anført ovenfor fastslog EUD i Infopaq I, at lagringen af et tekstuddrag på 11 ord fra ophavsretligt beskyttede aviserartikler (eller litterære værker, om man vil) kunne henføres under begrebet "delvis reproduktion", jf. InfoSoc-direktivets art. 2, hvis de elementer, der derved blev gengivet, var udtryk for ophavsmandens "egen intellektuelle frembringelse".

Ifølge EUD bestod de pågældende aviserartiklers originalitet i "valget, placeringen og kombinationen" af ord, der ikke i sig selv var originale på en måde, som var original. Dog var en bestanddel i form et sætningsled fra en avisartikel original ifølge EUD, hvis bestanddelen var egnet til at meddele læseren den originalitet, der lå i hele teksten. Ifølge EUD øgede det desuden sandsynligheden for, at der var tale om delvis eksemplarfremstilling, at datafangstprocessen muliggjorde udarbejdelsen af flere tekstilltag fra samme artikel, hvilket kunne føre til skabelsen af større afsnit, der var egnet til at afspejle den originalitet, der lå i hele artiklen.

I lyset af Infopaq I kan det således fastslås, at en bestanddel af et værk er udtryk for ophavsmandens "egen intellektuelle frembringelse", hvis den i sig selv afspejler den originalitet, der ligger i det værk, bestanddelen udgør en del af. Det skal dog bemærkes, at der hersker uenighed blandt visse juridiske forfattere om, hvorvidt dette principp også gælder for andre værkstyper end litterære værker.

Ifølge Schønning fører den omstændighed, at EUD lagde "[...] afgørende vægt på, at der i den foreliggende sag var tale om ord, der ikke i sig selv var genstand for beskyttelse", til, at EUD's resultat i Infopaq I ikke med sikkerhed gælder for andre værkstyper end litterære værker. Derudover anfører Schønning, at spørgsmålet om beskyttelse af enkeltdele af værker kun er relevant i de tilfælde, hvor værket består af elementer, som har en sådan karakter, at de ikke i sig selv kan beskyttes.

---

126 Ibid., pr. 44
127 Ibid., pr. 51
128 Ibid., pr. 45
129 Ibid., pr. 47
130 Ibid., pr. 50
131 Schønning (2013)
132 Ibid.
Derfor er spørgsmålet om beskyttelse af bestanddele af værker ikke relevant for værkstyper som film- eller musikværker, hvor alle dele af værket ifølge Schønning er frembragt til indspilningen.133

Heroverfor anfører Rosenmeier, at principippet efter hans mening gælder for alle former for værker, herunder filmværker.134 Rosenmeier begrunder bl.a. sit synspunkt med en henvisning til Premier League135, hvor EUD ifølge forfatteren besvarede et præjudicielt spørgsmål om, hvornår "fortløbende fragmenter af en film" kunne beskyttes, med, at sådanne fragmenter er beskyttet af ophavsretten, når de i sig selv opfylder originalitetskravet.136 Rosenmeier henviser derudover bl.a. til U 2014.888 H (Netto)137, hvor Højesteret fandt, at nogle reklamefilm var blevet krænket af nogle nye reklamefilm, som ifølge retten havde samme filmiske og visuelle udtryk. Dommen støtter ifølge Rosenmeier det synspunkt, at et udsnit på få sekunder af et filmværk i sig selv kan være originalt.138

På den ene side kan det anføres, at EUD's svar i Premier League ikke ordret gengiver spørgsmålet fra den engelske High Court of Justice til EUD om, hvornår "[...] fortløbende fragmenter af en film, et musikalsk værk eller en lydoptagelse" kunne beskyttes.139 Hertil svarede EUD mere generelt, at eneretten til reproduktion "[...] også omfatter flygtige fragmenter af værker [min understregning]", jf. Premier League.140 Således følger det ikke udtrykkeligt af Premier League, hvorvidt enkelte bestanddele af filmværker eller musikværker også kan beskyttes i det omfang, at de opfylder originalitetskravet. Derudover overlod EUD det til den engelske High Court of Justice at afgøre, hvorvidt den konkrete kopiering af fragmenter af beskyttede værker udgjorde en reproduktion.141

Svaret på dette spørgsmål følger imidlertid ikke af den engelske dom.142

På den anden side kan det anføres (som ligeledes bemærket af Rosenmeier143), at præmisserne i både Infopaq I og Premier League synes at være så generelt formuleret, at de må antages at gælde for alle værkstyper. Det følger af Infopaq I, at de forskellige bestanddele af "et værk" er beskyttet, hvis de

---

133 Ibid.
134 Rosenmeier (2014)
135 Ibid.
136 Ibid.
137 U 2014.888 H (Netto)
138 Rosenmeier (2014)
139 Premier League, pr. 54
140 Premier League, pr. 159
141 Premier League, pr. 158
142 Judgement of the High Court of Justice (England and Wales) (Chancery Division), 3rd of February 2012, [2012] EWHC 108 (Ch), Football Association Premier League Ltd. et al. vs. QC Leisure et al., AV Station Plc et al., and Michael Madden et al.
143 "Det fremgår af Infopaq-dommens præmis 37ff. og Premier League-dommens præmis 156, at principippet om, at værksdele i sig selv skal være originale for at kunne beskyttes, gælder for alle former for værker" – Rosenmeier (2014)
opfylder originalitetskravet. På samme vis følger det af Premier League, som gengivet ovenfor, at flygtige fragmenter "af værker" er beskyttet, hvis de opfylder originalitetskravet. På den baggrund synes EUD ikke at have udelukket, at enkelte bestanddele af andre værkstyper end litterære værker er beskyttet, hvis de i sig selv opfylder originalitetskravet.


Sammenfattende kan det derfor fastslås, at den kopiering af enkelte bestanddele af et beskyttet værk, der indgår i den anvendte træningsdata, som finder sted ifm. træningen af en ML-algoritme, er en "delvis eksemplarfremstilling", jf. OPHL § 2, stk. 1, jf. stk. 2, der derfor som udgangspunkt kræver tilladelse fra rettighedshaverne i det omfang, at den pågældende bestanddel i sig selv er udtryk for ophavsmandens "egen intellektuelle frembringelse", jf. Infopaq I. Dette gælder, uanset hvilken værkstype der er tale om. I lyset af Infopaq I kan det anføres, at jo større en bestanddel af værket, der bliver kopieret, jo større er sandsynligheden for, at dét, der er blevet kopieret, er udtryk for ophavsmandens egen intellektuelle frembringelse. Denne kopiering kræver dog heller ikke tilladelse fra rettighedshaverne, i det omfang, at kopieringen er dækket af en indskrænkning i ophavsretten.

4.1.2 Kopiering af værker i ændret skikkelse ifm. træningen
En særlig problemstilling opstår i de tilfælde, hvor algoritmen genererer et output ifm. træningen, der ligner et værk, som indgår i den anvendte træningsdata, og som algoritmen derfor har analyseret, men hvor der ikke er tale om værket i dets oprindelige skikkelse.

144 Infopaq I, pr. 39
145 Premier League, pr. 156
146 U 2014.3428 S (Oister Oi)
Denne problemstilling aktualiseres bl.a. af *Stable Diffusion*-eksemplet som beskrevet i afsnit 2.2.2 ovenfor, hvor algoritmen tydeligvis var blevet inspireret af de billeder, den havde analyseret ifm. træningen, da noget, der mindede om *Getty Images*-vandmærket, fremgik af algoritmens output.

Selvom der var tale om et output, der var genereret i anvendelsesfasen, skal det erindres, at træningen af en *ML*-algoritme som beskrevet i afsnit 2.2 almindeligvis først bliver afsluttet, når udviklerne er tilfredse med kvaliteten af outputtet. Der må således være et tidspunkt i løbet af træningen, hvor outputtet har samme "beskaffenhed" som det output, algoritmen er i stand til at generere efter træningen.

Et andet eksempel på et output, der synes at minde om noget, man har set før, kan frembringes ved at bede *Stable Diffusion*-modellen om et "Painting in the style of Andy Warhol":

Billederne i hhv. øverste venstre hjørne og nederste højre hjørne synes i et vist omfang at minde om de følgende originale Andy Warhol-værker:

---

147 Screenshot af billede genereret ved brug af *Stable Diffusion* (https://stablediffusionweb.com) d. 9. februar 2023 kl. 10:36
Som fastslået ovenfor omfatter eneretten til eksemplarfremstilling af et værk ikke blot
eksemplarfremstilling af værket i oprindelig skikkelse men også i ændret skikkelse, jf. OPHL § 2,
stk. 1. Indehaverne af ophavsretten til et værk er derfor beskyttet mod eksemplarfremstilling af
andres efterligninger af dette værk. Som anført af Schønning er der i et sådant tilfælde tale om en "delvis"
kopiering af værket, jf. § 2, stk. 2. 

148 Campbell's Soup Cans af Andy Warhol (1962) – https://www.moma.org/learn/moma_learning/andy-warhol-
campbells-soup-cans-1962/ (lokaliseret d. 27. februar 2023 kl. 14:29)
27. februar 2023 kl. 14:34)
150 Schønning (2021), s. 181
Om et værk er blevet efterlignet eller ej, afhænger af en konkret vurdering af den påståede efterligning i lyset af de særlige principper, der gælder for denne vurdering. Principperne indebærer bl.a., at døt, der er blevet efterlignet, skal være udtryk for ophavsmandens "eigen intellektuelle frembringelse", jf. fx Infopaq I som beskrevet ovenfor, og at døt, der er overtaget fra det beskyttede værk, ikke blot består i mere abstrakte idéer og principper, jf. fx U 1999.1762 Ø (Kvit eller Dobbelt), hvor fogedretten fandt, at et tv-format ikke nød ophavsretlig beskyttelse, da det bestod af en række idéer og principper, som ikke isoleret set kunne beskyttes af ophavsretten.

Disse principper kunne meget vel føre til, at Stable Diffusion-versionen af Warhols ikoniske Marilyn Monroe-maleri kunne indebære en krænkelse, da flere konkrete træk fra det originale værk synes at gå igen i det AI-genererede værk. Derimod kan det i højere grad diskuteres, om et AI-genereret billede af en samling farverige dåser (i øvrigt uden Campbell-logo) krænker Warhols 32 Campbell's Soup Cans-malerier, da billedet og malerierne adskiller sig fra hinanden i kraft af både farver og antal, og da det eneste, som billedet synes at have overtaget fra malerierne, er en dåse med en label.

Uanset hvad svaret på disse krænkelsesspørgsmål er, kan det dog sammenfattende fastslås, at den kopiering af træningsdata, når denne data består af ophavsretligt beskyttede værker, som finder sted ifm. algoritmens generering af output, der i et eller andet omfang indeholder bestanddele fra et af disse værker, men uden at der er tale om værket i dets oprindelige skikkelse, er omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1 Kopieringen kræver derfor som udgangspunkt tilladelse fra rettighedshaverne, hvis der i lyset af de ovenfor beskrevne principper er tale om en eksemplarfremstilling af værket i ændret skikkelse. Denne eksemplarfremstilling kræver dog ikke tilladelse fra rettighedshaverne, hvis kopieringen er dækket af en indskrænkning i ophavsretten, hvilket vil blive undersøgt i afsnit 5.

Ovenstående gælder naturligvis også for de to øvrige tilfælde af kopiering af træningsdata, der som beskrevet i afsnit 2.2.1 finder sted ifm. træningen af en ML-algoritme.

4.2 Delkonklusion

Kopieringen af træningsdata, der finder sted ifm. træningen af en ML-algoritme, som skal implementeres i et generative AI-system, når denne data består af ophavsretligt beskyttede værker, er omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1, og kræver som udgangspunkt tilladelse fra rettighedshaverne, hvis der er tale om kopiering af værkerne i deres oprindelige skikkelse.

151 Schovsbo et al. (2021), s. 153ff. Schønning (2021), s. 183ff.
152 Schovsbo et al. (2021), s. 154ff. Schønning (2021), s. 187
og i deres helhed. Hvis der derimod er tale om kopiering af enkelte bestanddele af værkerne i oprindelig skikkelse, er kopieringen kun omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1, og kræver således kun tilladelse fra rettighedshaverne, hvis de elementer, der bliver kopieret, i sig selv er udtryk for ophavsmandens "egen intellektuelle frembringelse", jf. Infopaq I, og at elementerne således i sig selv opfylder originalitetskraavet.

Derudover er kopieringen af træningsdata, når denne består af værker, omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1, og kræver derfor tilladelse fra rettighedshaverne, når kopieringen udgør eksemplarfremstilling af værket i ændret skikkelse, dvs. en efterligning af værket. Om dette er tilfældet, må vurderes konkret i lyset af de principper, der gælder for denne vurdering, herunder, om de elementer, der er blevet kopieret, er udtryk for ophavsmandens "egen intellektuelle frembringelse", jf. Infopaq I.

5 Kopieringen af træningsdata i lyset af de lovbestemte indskrænkninger i ophavsretten

I dette afsnit vil det blive undersøgt, i hvilket omfang kopieringen af træningsdata, når denne data består af værker, er dækket af de lovbestemte indskrænkninger i ophavsretten.

I afsnit 5.1 vil det indledningsvist blive vurderet, hvilke lovbestemte indskrænkninger der er relevante for kopieringen af træningsdata, når denne data består af værker. I de efterfølgende afsnit vil de identificerede indskrænkninger blive behandlet hver især mhp. at fastslå, i hvilket omfang de dækker kopieringen.

5.1 Hvilke lovbestemte indskrænkninger er relevante for træningen af en ML-algoritme?

Som fastslået i afsnit 3.2 findes der tre typer af lovbestemte indskrænkninger, der gør det muligt at bruge et værk på en bestemt måde uden tilladelse fra indehaverne af ophavsretten: i) fribrugsregler, ii) regler om aftalelicens og iii) regler om tvangslicens.153

Indledningsvist bemærkes det, at der som følge af det forhold, at dette speciale kun undersøger den træning af ML-algoritmer, der bliver gennemført af virksomheder (såsom Stability AI) som led i disses erhvervsvirksomhed og til kommersielle formål, jf. afgrænsningen i afsnit 1.4, kun er tre lovbestemte indskrænkninger, der synes relevante.

---

153 Schovsbo et al. (2021), s. 170f. Rosenmeier (2018), s. 97
Den første af disse indskrænkninger er undtagelsen for midlertidig eksemplarfremstilling iht. OPHL § 11 a. EU-Kommissionen har ved flere lejligheder (dog uden at konkludere, hvorvidt undtagelsen rent faktisk finder anvendelse) påpeget denne undtagelses relevans i relation til visse reproduktionshandlinger, der finder sted ifm. træningen af en ML-algoritme. Den samme opfattelse af undtagelsens relevans ses også hos flere juridiske forfattere. Indskrænkningens anvendelighed for kopiering af træningsdata har dog, som anført i afsnit 3.3, ikke været prøvet af hverken EUD eller de nationale domstole i EU.

Den anden lovbestemte indskrænkning, der synes relevant, er den obligatoriske undtagelse iht. DSM-direktivets art. 4, der gælder for eksemplarfremstilling af beskyttede værker, som bliver foretaget "med henblik på tekst- og datamining". Den generelle opfattelse hos både EU-Kommissionen og en række juridiske forfattere synes også her at være den, at undtagelsen er relevant for visse af de eksemplarer af værker, der almindelighvis bliver genereret ifm. træningen af en ML-algoritme. Denne indskrænkningens anvendelighed for kopiering af træningsdata har dog, som anført i afsnit 3.3, heller ikke været prøvet af EUD eller de nationale domstole i EU.


154 A literature review (2021), s. 10. EU-Kommissionens rapport om AI (2022), s. 183
157 L 58 2007-08, afsnit 3.2. Schønning (2021), s. 552
158 L 58 2007-08, afsnit 3.3.1. Schønning (2021), s. 553. Schovsbo et al. (2021), s. 215
159 Kulturministeriets oversigt over aftalelicensgodkendelser (se under fanen ’Godkendelser efter den generelle aftalelicens’) – https://kum.dk/kulturomraader/aftalelicensgodkendelser (lokaliseret d. 9. maj 2023 kl. 16:47)
Det er således de tre specifikke lovbestemte indskrænkninger, der er blevet identificeret ovenfor, der vil blive behandlet i det følgende mhp. at fastslå, i hvilket omfang de dækker kopieringen af træningsdata, når denne data består af beskyttede værker.

Først vil OPHL § 11 a blive behandlet i afsnit 5.2, dernæst behandles DSM-direktivets art. 4 i afsnit 5.3, og endelig behandles OPHL § 50, stk. 2, i afsnit 5.4.

5.2 Undtagelsen for midlertidig eksemplarfremstilling, jf. OPHL § 11 a

Undtagelsen for midlertidig eksemplarfremstilling iht. OPHL § 11 a stammer fra InfoSoc-direktivets art. 5(1), der ligesom direktivets øvrige regler blev gennemført i dansk ret på baggrund af vedtagelsen af lovforslaget L 19 2002-03. OPHL § 11 a skal således fortolkes i overensstemmelse med art. 5(1) som følge af den totalharmonisering af undtagelserne til reproduktion, som InfoSoc-direktivet, som anført i afsnit 4.1, havde til formål at sikre.

Det følger af OPHL § 11 a, at det (med visse undtagelser) er tilladt at fremstille midlertidige eksemplarer af ophavsretligt beskyttede værker, hvis de fire kumulative betingelser, der er opregnet i bestemmelsens stk. 1, nr. 1-4, er opfyldt. Det er således tilladt i medfør af OPHL § 11 a, stk. 1, at fremstille eksemplarer, som:

i) er flygtige eller tilfældige  
ii) udgør en integreret og væsentlig del af en teknisk proces  
iii) udelukkende har til formål at muliggøre enten en mellemmands transmission af et værk i et netværk mellem tredjemænd eller en lovlig brug af et værk  
iv) ikke har nogen selvstændig værdi.

Det vil i det følgende blive undersøgt, i hvilket omfang disse betingelser er opfyldt, når der er tale om kopiering af beskyttede værker ifm. træningen af en ML-algoritme. Som fastslået ovenfor, har dette spørgsmål endnu ikke været behandlet af EUD eller nogen national domstol i EU.

---

160 L 19 2002-03, s. 241  
161 InfoSoc-direktivets 31.-33. præambelbetragtning  
162 Undtagelsen i OPHL § 11 a, stk. 1, gælder ikke for eksemplarfremstilling af databaser og edb-programmer, jf. § 11 a, stk. 2.  
163 L 19 2002-03, s. 242
5.2.1 Flygtige eller tilfældige eksempler

Selvom den almindelige sproglige forståelse af begreberne "flygtig" og "tilfældig" lægger op til, at der skal være tale om visse kortvarige eller uplanlagte eksempler, bemærkes det, at det hverken følger af OPHL § 11 a, de danske lovomtiver eller InfoSoc-direktivet, hvad der konkret skal til, for at et midlertidig eksemplar af et værk kan karakteriseres som netop "flygtigt" eller "tilfældigt".

EUD har dog ved flere lejligheder besvaret præjudicielle spørgsmål om fortolkningen af InfoSoc-direktivets art. 5(1).

Den første af de relevante domme er *Infopaq I*, som beskrevet i afsnit 4.1.1.

Indledningsvist bemærkes det, at EUD’s fortolkning af begrebet "flygtigt" er funderet i tre overordnede principper for anvendelsen af art. 5(1).

For det første henviser EUD til det princip, som er fastslået i dens tidligere retspraksis, efter hvilket en undtagelse i et direktiv til et almindeligt princip i samme direktiv skal fortolkes "indskrænkende".

For det andet henviser EUD til, at InfoSoc-direktivets art. 5(5) (den såkaldte *tretrinittest* som beskrevet i afsnit 3.2) skal iagttages ved anvendelsen af art. 5(1).

Endelig henviser EUD for det tredje til, at betingelserne i art. 5(1) skal fortolkes "[...] i lyset af kravet om retssikkerhed for ophavsmændene i relation til beskyttelsen af deres værker". Det bemærkes, at det da også følger af InfoSoc-direktivets 4. præambelbetrætning, at direktivet bl.a. har til formål at øge retssikkerheden og sikre overholdelsen af "et højt beskyttelsesniveau" for intellektuelle ejendomsrettigheder i EU.

De to første principper har ifølge EUD den konkrete betydning, at en midlertidig reproduktionshandling kun er omfattet af undtagelsen i art. 5(1), hvis den ikke går ud over det, som er nødvendigt for, at den tekniske proces, som reproduktionen har til formål at muliggøre, kan udføres korrekt. Det tredje princip indebærer ifølge EUD, at opbevaringen og tilintetgørelsen ikke må

---

165 Kapper, pr. 72. *Kommissionen mod Spanien*, pr. 31
166 *Infopaq I*, pr. 56
167 Ibid., pr. 59
168 Ibid., pr. 61
afhænge af en fysisk persons skøn, da der i så fald ikke er nogen garanti for, at den pågældende person rent faktisk tilintetgør eksemplaret.169

Derfor kan en reproduktion ifølge EUD kun kvalificeres som "flygtig" i overensstemmelse med art. 5(1), "[...] såfremt dens eksistens er begrænset til det tidsrum, der er nødvendigt for en korrekt udførelse af den omhandlede tekniske proces".170 Derudover forudsættes det, at den tekniske proces "[...] forløber automatisk i den forstand, at den automatisk tilintetgør denne handling uden en fysisk persons medvirken, så snart dens funktion bestående i at muliggøre udførelsen af en sådan proces er opfyldt."171

I lyset af de to betingelser kunne EUD ikke udelukke, at reproduktionshandlingerne i datafangstprocessens fase to og tre (dannelsen af TIFF-filer hhv. omdannelsen af disse til tekstfiler) kunne kvalificeres som "flygtige", da reproduktionerne automatisk blev slettet fra computerens hukommelse, når de hver især havde opfyldt deres funktion.172 Desuden kunne EUD ikke fastslå, om reproduktionshandlingen i datafangstprocessens fjerde fase (lagringen af tekstuddraget) var "flygtig". Afgørelsen heraf blev derfor overladt til Højesteret, der dog fandt, at lagringen af tekstuddraget ikke var "flygtig", da filen med uddraget ikke blev slettet "[...] automatisk fra computeren uden en fysisk persons medvirken efter et kort tidsrum".173 Derfor var det kun reproduktionshandlingen i datafangstprocessens femte fase (det fysiske udprint af tekstuddraget), som EUD under de foreliggende oplysninger kunne avise var "flygtig", da tilintetgørelsen af reproductionen afhäng af, at brugeren af datafangstprocessen besluttede sig for at destrukere det fysiske papir, som tekstuddraget fremgik af.174

Heroverfor fandt EUD i den senere dom Meltwater (i øvrigt til stor forundring blandt visse juridiske forfattere, da den traditionelle opfattelse er, at det ikke er eksemplarfremstilling at gengive et værk på en skærm175), at de midlertidige kopier af beskyttede værker, som blev fremstillet på en internetbrugers computerskærm, når denne bruger besøgte en internetside, på hvilken de pågældende værker var tilgængeliggjort, kunne kvalificeres som "flygtige". Dette skyldtes ifølge EUD, at kopierne

169 Ibid., pr. 62
170 Ibid., pr. 64
171 Ibid., pr. 64
172 Ibid., pr. 65
173 U 2013.1782 H (Infopaq)
174 Infopaq I, pr. 69-70
175 Schovsbo et al. (2021), s. 120
på skærmen automatisk blev tilintetgjort af computeren, når brugeren forlod internetsiden og på den måde afsluttede den tekniske proces, som blev anvendt for at besøge den pågældende internetside.176

I øvrigt følger det af Meltwater, at kravet om automatisk tilintetgørelse som fastslået i Infopaq I ikke er til hinder for, at sletningen forudsætter "[...] en menneskelig indgriben med henblik på at afslutte brugen af den tekniske proces", da det ifølge EUD er tilladt, at den tekniske proces både iværksættes og fuldføres manuelt.177 På den baggrund kan det anføres, at selvom tilintetgørelsen af et eksemplar skal ske automatisk uden en fysisk persons medvirken, må tilintetgørelsen af eksemplaret dog godt ske i kraft af en fysisk persons beslutning om at afbryde den tekniske proces.

I forlængelse heraf bemærkes det dog, at der synes at herske modstridende opfattelser hos hhv. Højesteret og EUD af den tilladte længde af det tidsrum, som et midlertidigt eksemplars eksistens iht. Infopaq I skal være begrænset til.

Som beskrevet ovenfor fandt Højesteret i den efterfølgende danske dom, at tekstuddraget på 11 ord, der blev lagret i Infopaqs system, ikke var "flygtigt", da filen med uddraget ikke blev slettet "[...] automatisk fra computeren uden en fysisk persons medvirken efter et kort tidsrum [min understregning]".178 Heroverfor følger det af Meltwater, at det var helt uden betydning for EUD's resultat, at den midlertidige kopi på computerskærmen blev bevaret, lige så længe at internetbrugeren lod sin browser være åben, da den tekniske proces, som blev anvendt for at besøge internetsiden, forblev aktiv i den periode.179 Hvor Højesteret således synes at være af den opfattelse, at det tilladte tidsrum skal være et "kort tidsrum", synes EUD's opfattelse derimod at være den, at det tilladte tidsrums varighed er underordnet, hvis bare den pågældende tekniske proces er aktiv i hele tidsrummet.

Henset til at Højesterets opfattelse af det nødvendige tidsrums varighed ikke synes at kunne finde støtte i præmisserne i Infopaq I (udover det åbenlyse forhold, at eksemplarerne skal være midlertidige), og henset til at Meltwater i øvrigt er afsagt efter U 2013.1782 H (Infopaq), må det være EUD's opfattelse af det nødvendige tidsrums varighed, som denne kommer til udtryk i Meltwater, der skal lægges til grund ved vurderingen af, om et eksemplar er "flygtigt".

176 Meltwater, pr. 44
177 Ibid., pr. 41
178 U 2013.1782 H (Infopaq)
179 Meltwater, pr. 45
Betrager man de ovenstående betingelser, der kan udledes af *Infopaq I og Meltwater*, synes opfyldelsen af samtlige betingelser at bero på den forudsætning, at kopierne er blevet fremstillet ifm. "en teknisk proces", og at kopiernes funktion alene er at muliggøre denne proces. Vurderingen af disse to forhold synes imidlertid at være snævert forbundet med spørgsmålet om, hvorvidt kopierne udgør "en integreret og væsentlig del af en teknisk proces", der vil blive behandlet nedenfor i afsnit 5.2.2. Det må af den årsag først endegyldigt kunne fastslås, om en kopi, der er blevet fremstillet ifm. træningen af en *ML*-algoritme, er "flygtig", når det er blevet fastslået i afsnit 5.2.2, om kopien opfylder betingelsen iht. OPHL § 11 a, stk. 1, nr. 2.

Det synes dog ikke at være udelukket, at visse af de kopier af værker, der bliver fremstillet ifm. træningen af en *ML*-algoritme på de tre tidspunkter, der blev beskrevet i afsnit 2.2.1, vil kunne anses for "flygtige". Denne opfattelse finder støtte hos både Chiou, Axhamn, Margoni og Schönberger. Da tilintetgørelsen af kopierne godt må ske i kraft af en fysisk persons beslutning om at afbryde den tekniske proces, jf. Meltwater, har det således ingen betydning, hvis tilintetgørelsen sker som følge af, at udviklerne afbryder træningen. Derudover kan man argumentere for, at det heller ikke har nogen betydning, at træningen af en *ML*-algoritme undertiden tager flere dage, da der som anført ovenfor formentlig ikke gælder nogen krav til varigheden af det nødvendige tidsrum, under forudsætning af at den tekniske proces, som kopierne bliver fremstillet i forbindelse med, er aktiv i al den tid, som eksemplarerne eksisterer i, jf. Meltwater. Det kan dog først endeligt fastslås, om kopierne kan betragtes som "flygtige", hvis det kan fastslås i afsnit 5.2.2, at kopierne bliver fremstillet ifm. en teknisk proces, og at kopiernes formål er at muliggøre denne proces.

Det skal dog bemærkes, at de kopier, der bliver fremstillet ifm. lagringen af træningsdata i det system, som træningen skal gennemføres i, formentlig undertiden må antages ikke at kunne kvalificeres som "flygtige" i praksis. Dette skyldes, at tilintetgørelsen af disse kopier ofte afhænger af udviklernes aktive beslutning herom, da det hænder, at udviklerne vælger at bibeholde et lagret datasæt til efterfølgende træning af andre algoritmer. Således er der en risiko for, at kopierne eksisterer i en

---

182 Meltwater, pr. 45
længere periode afhængig af brugerens behov, hvilket i *Infopaq I* var en af årsagerne til, at det fysiske udprint af tekstdraget ikke kunne kvalificeres som "flygtigt".\(^{184}\)

I de tilfælde, hvor en kopi af et beskyttet værk, der er blevet fremstillet ifm. træningen af en *ML*-algoritme, ikke kan kvalificeres som "flygtig", er kopien alligevel omfattet af OPHL § 11 a, stk. 1, nr. 1, hvis den i stedet kan kvalificeres som "tilfældig".

Det følger af *Meltwater*, præmis 43, at et eksemplar er "tilfældigt", hvis det "[...] hverken har en selvstændig eksistens eller et selvstændigt formål i forhold til den tekniske proces, som den udgør en del af".\(^{185}\)

I den konkrete sag fandt EUD, at de kopier af beskyttede værker, som blev fremstillet i internetcachen i harddisken på en computer, når indehaveren af computeren besøgte internetsider, på hvilke de pågældende værker var tilgængeliggjort, hverken havde en selvstændig eksistens eller et selvstændigt formål og derfor kunne kvalificeres som “tilfældige”.\(^{186}\) EUD begrundede bl.a. dens resultat med, at den pågældende tekniske proces var fuldstændig bestemmende for de formål, til hvilke kopierne var blevet frembragt og anvendt.\(^{187}\) Derudover henviste EUD til det forhold, at internetbrugerne ikke kunne frembringe kopierne i cachen uden for den tekniske proces.\(^{188}\)

Som anført ovenfor vil kopierne kunne genbruges til efterfølgende træning af andre algoritmer. At det undertiden hænder, at udviklerne bevarer den anvendte træningsdata, understreges bl.a. af *LAION-5B*, da dette datasets fortsat er offentligt tilgængeligt til brug for træningen af andre *ML*-algoritmer.\(^{189}\)

Selvom man kan argumentere for, at det overordnede formål, som kopierne tjener, fortsat vil være det samme (træningen af en *ML*-algoritme), vil der imidlertid være tale om efterfølgende træning af en helt anden algoritme, hvilket taler for, at kopierne har en selvstændig eksistens ift. den tekniske proces, de udgør en del af (i fald at træningen kan betegnes som en sådan proces, hvilket bliver undersøgt i afsnit 5.2.2). Heroverfor var kopierne i internetcatchen, som EUD tog stilling til i *Meltwater*, "tilfældige", fordi de ikke kunne bruges til noget som helst andet end at muliggøre surfing

---

\(^{184}\) *Infopaq I*, pr. 69

\(^{185}\) *Meltwater*, pr. 43

\(^{186}\) *Ibid.*, pr. 50

\(^{187}\) *Ibid.*, pr. 49

\(^{188}\) *Ibid.*, pr. 49

\(^{189}\) Beaumont, Romain: *LAION-5B – A New Era of Open Large-Scale Multi-Modal Datasets* (2022) – [https://www.laion.ai/blog/laion-5b/](https://www.laion.ai/blog/laion-5b/) (lokalisert d. 2. marts 2023 kl. 10:12)
på internettet, og fordi brugeren ikke kunne frembringe kopierne i cachen uden for den tekniske proces.\textsuperscript{190}

Det må på den baggrund afvises, at de kopier, der bliver fremstillet ifm. lagring af træningsdata, vil kunne kvalificeres som "tilfældige". Kopierne opfylder således kun betingelsen iht. OPHL § 11 a, stk. 1, nr. 1, hvis de kan betegnes som "flygtige" i overensstemmelse med betingelserne som beskrevet ovenfor.

5.2.2 Integreret og væsentlig del af en teknisk proces

Indledningsvist skal det bemærkes, at det ikke udtrykkeligt følger af InfoSoc-direktivet eller de danske lovmotiver, hvad der konkret skal forstås ved begrebet "\textit{en teknisk proces}". De nævnte retskilder indeholder desuden kun ganske få eksempler på processer, der er omfattet af indskrænkningen. Et specifikt eksempel på en teknisk proces, der går igen i begge retskilder (og som desuden følger af Meltwater\textsuperscript{191}), er den række af bestemte handlinger, som bliver foretaget af en computer (herunder i internetcatchen på denne), når brugeren af computeren surfer på internettet.\textsuperscript{192}

På trods af denne begrænsede udfoldning af begrebet "\textit{en teknisk proces}" i de nævnte retskilder synes EUD's kendelse i \textit{Infopaq II} dog at understøtte, at træningen af en ML-algoritme også vil kunne betragtes som en sådan proces.

\textit{Infopaq II} angår den samme datafangstproces, som EUD tog stilling til i \textit{Infopaq I}. Baggrunden for \textit{Infopaq II} var, at Højesteret var af den opfattelse, at det kunne blive aktuelt (hvilket det dog aldrig blev\textsuperscript{193}) i den verserende danske sag mellem Infopaq og DDF, at retten skulle tage stilling til, om datafangstprocessen var i overensstemmelse med art. 5(1), hvis Infopaq undlod at foretage det fysiske udprint af tekstuddraget i datafangstprocessens femte fase.\textsuperscript{194}

Det kan udledes af \textit{Infopaq II}, at en proces, der "[...] består i at foretage elektroniske og automatiske søgninger i avisartikler og i at identificere og udvælge forud fastlagte søgeord fra disse artikler", som bliver indledt med dannelsen af en digital fil på baggrund af en indscanning af en avisartikel, og

\textsuperscript{190} Meltwater, pr. 16
\textsuperscript{191} Ibid., pr. 63
\textsuperscript{192} InfoSoc-direktivets 33. præambelbetræftning. L 19 2002-03, s. 242
\textsuperscript{193} Højesteret fandt, at de lagrede tekstuddrag ikke var "flygtige", hvorfor der således ikke var behov for at behandle de resterende betingelser – se U 2013.1782 H (Infopaq)
\textsuperscript{194} Infopaq II, pr. 23
som bliver afsluttet med lagringen af en række tekstuddrag på 11 ord fra artiklen, udgør en teknisk proces.¹⁹⁵

I lyset af Infopaq II kan man argumentere for, at træningen af en ML-algoritme også kan betegnes som "en teknisk proces". Ligesom med Infopaq's datafangstproces indebærer algoritmens analyse netop en elektronisk og automatisk processering af data i form af værker i digital form i et computersystem mhp. at identificere på forhånd fastlagte karakteristika eller sammenhænge i den pågældende data. Derudover kan der påpeges den lighed mellem de to processer, at algoritmens analyse tilsvarende bliver indledt med en lagring af insamlet data i det system, som træningen skal gennemføres i, og efterfulgt af udarbejdelsen og lagringen af kopier af den analyserede data (på baggrund af algoritmens analyse samt ifm. algoritmens generering af output).

For at kopieringen af værker, der finder sted i forbindelse med træningen af en ML-algoritme kan siges at udgøre en "integreret og væsentlig del af en teknisk proces", jf. OPHL § 11 a, stk. 1, nr. 2, skal eksemplarfremstillingen ifølge EUD opfyldte to betingelser, jf. Infopaq II.

For det første er det en betingelse, at "[…] de midlertidige reproduktionshandlinger helt og holdent udføres inden for rammerne af en teknisk proces, og at de således ikke helt eller delvist udføres uden for en sådan proces".¹⁹⁶ Ifølge EUD udelukker InfoSoc-direktivets art. 5(1) ikke, at den midlertidige reproduktionshandling indleder eller afslutter den tekniske proces.¹⁹⁷ Desuden udelukker art. 5(1) ikke, at den tekniske proces indebærer en eller anden form for menneskelig indgriben og herunder, at processen "[…] iværksættes manuelt med det formål at fremstille en første midlertidig reproduktion".¹⁹⁸

For det andet er det en betingelse, at reproduktionshandlingen er "[…] nødvendig i den forstand, at den pågældende tekniske proces ikke ville kunne udføres korrekt og effektivt uden denne handling".¹⁹⁹

EUD fandt, at de reproduktionshandlinger, der blev foretaget i den reviderede version af datafangstprocessen (dannelsen af en TIFF-fil, dannelsen af en tekstfil og dannelsen af en fil med et tekstuddrag på 11 ord), udgjorde en integreret og væsentlig del af en teknisk proces i overensstemmelse med InfoSoc-direktivets art. 5(1), da handlingerne opfyldte de to betingelser.²⁰⁰

¹⁹⁵ Ibid., pr. 33 og pr. 36
¹⁹⁶ Ibid., pr. 30
¹⁹⁷ Ibid., pr. 31
¹⁹⁸ Ibid., pr. 32
¹⁹⁹ Ibid., pr. 30
²⁰⁰ Ibid., pr. 39
Det skal i den forbindelse bemærkes, at datafangstprocessen ifølge EUD ikke ville kunne udføres korrekt og effektivt uden de pågældende reproduktionshandlinger, da en elektronisk søgning i avisartikler kræver, at artiklerne konverteres til data, da dette er nødvendigt for "[...] at genkende de pågældende data, for at identificere søgeordene og for at udvælge disse".\textsuperscript{201}


Hvad angår den anden betingelse, skal det først og fremmest anføres, at træningen som beskrevet i afsnit 2.1 bliver gennemført mhp. at lære algoritmen, hvordan den skaber det bestemte "kreativt" output. Som ligeledes beskrevet kan algoritmen kun lære at generere et sådant output ved at analysere og lære fra indhold af samme type som den type indhold, algoritmen skal lære at skabe. Således er træningsdata det vigtigste element i træningen.\textsuperscript{202} Da algoritmens analyse desuden kun kan gennemføres ved, at det indhold, som er genstand for analysen, bliver lagret i det computersystem, som træningen skal gennemføres i, er kopieringen, der finder sted ifm. lagringen, utvivlsomt nødvendig for den korrekte og effektive gennemførelse af algoritmens analyse. På den baggrund kan det anføres, at den kopiering, der finder sted ifm. lagringen af træningsdata, vil kunne betragtes som "nødvendig" i overensstemmelse med den anden betingelse, der følger af Infopaq II, præmis 30.

Ift. kopieringen af værker, der finder sted på baggrund af algoritmens analyse af den pågældende træningsdata, danner kopieringen som beskrevet i afsnit 2.1 grundlag for algoritmens selvjustering og optimering af de justérbare parametre i dens bagvedliggende matematiske funktion, hvilket som anført er den måde, hvorpå en algoritme lærer at generere et præcist og korrekt output. Som ligeledes beskrevet gør algoritmens generering af output ifm. træningen det muligt for udviklerne at vurdere, om der er behov for at finjustere algoritmen, før den bliver implementeret i et generative AI-system.

\textsuperscript{201} Ibid., pr. 37
\textsuperscript{202} Drexl \textit{et al.} (2019), s. 8
På den baggrund kan man argumentere for, at den kopiering af værker, der er anvendt som træningsdata, som finder sted på baggrund af algoritmens analyse af denne data samt ifm. algoritmens generering af output, også kan betragtes som "nødvendig" i overensstemmelse med den anden betingelse, der følger af Infopaq II.

Sammenfattende kan det derfor anføres, at det ikke er udelukket, at de kopier, der bliver fremstillet i de tre tilfælde ifm. træningen af en ML-algoritme, udgør "en integreret og væsentlig del af en teknisk proces" i overensstemmelse med OPHL § 11 a, stk. 1, nr. 2. Denne opfattelse finder i øvrigt støtte hos både Chiou og Margoni, der dog ikke begrunder deres synspunkt nærmere.203

Som følge heraf kan det retlige hængeparti fra afsnit 5.2.1 desuden endeligt afklares. Kopierne vil således også kunne betegnes som "flygtige" i overensstemmelse med OPHL § 11 a, stk. 1, nr. 1, i det omfang, at de kun bliver opbevaret så længe, det er nødvendigt, og at de bliver slettet automatisk, når de har opfyldt deres funktion.

5.2.3 Muliggørelse af mellemmands transmission eller en lovlig brug af et værk

Et eksemplar er kun omfattet af undtagelsen, hvis det udelukkende har til formål at muliggøre "en mellemmands transmission af et værk i et netværk mellem tredjemænd" eller "en lovlig brug af et værk", jf. OPHL § 11 a, stk. 1, nr. 3.

Hvor betingelsens første led ifølge de danske lovmotiver er målrettet den eksemplarfremstilling, der bliver foretaget af netværksoperatører og internetudbydere som led i disse mellemmandsvirksomhed på internettet, vedrører bestemmelsens andet led om "en lovlig brug" forholdet til "slutbrugeren" af værket.204 Det er derfor bestemmelsens andet led, der er relevant i relation til den kopiering, der finder sted ifm. træningen af en ML-algoritme.

Det følger af InfoSoc-direktivets 33. præambelbetragtning, at en brug af et værk anses for at være "lovlig" i overensstemmelse med betingelsen iht. art. 5(1), "[...] når den er godkendt af rettighedshaveren, eller når den ikke er undergivet lovmaessige restriktioner".205

Selvom det kan hævdes, at det krav, at brugen ikke må være undergivet "lovmaessige restriktioner", kan forstås bredt og således, at brugen af værket ikke må være undergivet begrensninger fra nogen som helst retsregel og ikke kun fra reglerne om ophavsret206, følger det dog af de danske lovmotiver,

203 Chiou (2019), s. 402ff. Margoni (2018), s. 19
204 L 19 2002-03, s. 242. InfoSoc-direktivets 33. præambelbetragtning
205 InfoSoc-direktivets 33. præambelbetragtning
206 Synodinou (2019), s. 23
at det antages, at en brug af et værk ikke er undergivet lovmæssige restriktioner, "[...] hvis brugen enten falder helt uden for de ophavsretlige eneretsbøjelser eller er dækket af en undtagelse". På den baggrund kan det derfor anføres, at kopieringen af værker, der finder sted i de tre tilfælde ifm. træningen af en ML-algoritme, kun opfylder betingelsen i det omfang, at kopieringen udelukkende har til formål at muliggøre en brug af et værk, som a) er godkendt af rettighedshaver eller dækket af en indskrænkning i ophavsretten, hvis brugen er omfattet af OPHL § 2, stk. 1 (der opregner de ophavsretlige eneretsbøjelser), eller som b) slet ikke er omfattet af OPHL § 2, stk. 1.

Før det skal vurderes, hvorvidt den brug af beskyttede værker, som de kopier, der bliver fremstillet ifm. træningen, har til formål at muliggøre, er omfattet af OPHL § 2, stk. 1, skal det dog indledningsvist bemærkes, at det kun er de kopier, der bliver fremstillet ifm. lagringen af træningsdata, som træningen skal gennemføres i, samt ifm. algoritmens generering af output, der synes at have til formål at muliggøre en brug af et værk.

Både Margoni og Axhamn anfører således, at de kopier, der bliver fremstillet ifm. lagringen af træningsdata, har til formål at muliggøre en brug af værker, som består i algoritmens analyse og kopiering af disse værker. Da kopieringen ifm. analysen således udgør selve brugen af et værk, kan det derfor udelukkes, at de kopier, som algoritmen fremstiller ifm. analysen, opfylder betingelsen i OPHL § 11 a, stk. 1, nr. 3.

Hvad angår den brug af et værk, som består i algoritmens analyse og kopiering af værket, er denne brug som fastslået i afsnit 4 omfattet af OPHL § 2, stk. 1, når kopieringen udgør en hel eller delvis eksemplarfremstilling af det pågældende værk. Hvorvidt dette er tilfældet, afhænger som ligeledes fastslået i afsnit 4.1.1 af, om de elementer fra værket, der er blevet kopieret, er udtryk for ophavsmandens "egen intellektuelle frembringelse", jf. bl.a. Infopaq I. Som beskrevet i afsnit 2.2.1 varierer det, hvilke elementer i de analyserede værker, algoritmen kopierer, hvorfor det således kan fastslås, at der i praksis både vil være tilfælde, hvor algoritmens kopiering af de analyserede værker er omfattet af OPHL § 2, stk. 1, og tilfælde, hvor kopieringen slet ikke er omfattet af de ophavsretlige eneretsbøjelser, da de elementer, som algoritmen har gengivet, ikke er udtryk ophavsmandens egen intellektuelle frembringelse.

207 L 19 2002-03, s. 242. Schønning (2021), s. 293
208 Axhamn (2021), s. 55. Margoni (2018), s. 19
209 Infopaq I, pr. 51
I de tilfælde, hvor algoritmens kopiering ikke er omfattet af OPHL § 2, stk. 1, kan man i lyset af de danske lovmotiver argumentere for, at algoritmens kopiering af de analyserede værker udgør en "lovlig" brug af disse værker.\textsuperscript{210} I disse tilfælde vil de kopier, der er blevet fremstillet ifm. lagringen af træningsdata i det system, som træningen skal gennemføres i, og som har til formål at muliggøre denne brug, opfylde betingelsen i OPHL § 11 a, stk. 1, nr. 3.

Derimod må det i lyset af InfoSoc-direktivets 33. præambelbetræftning samt de danske lovmotiver som gengivet ovenfor antages, at algoritmens kopiering af de analyserede værker i de tilfælde, hvor denne kopiering udgør en ophavsretligt relevant eksemplarfremstilling, jf. OPHL § 2, stk. 1, kun udgør en "lovlig brug" af værkerne, hvis indehaveren af ophavsretten har givet samtykke til eksemplarfremstillingen, eller hvis eksemplarfremstillingen er omfattet af en indskrænkning i ophavsretten. I disse tilfælde opfylde de kopier, der bliver fremstillet ifm. lagringen af træningsdata i systemet, som træningen skal gennemføres i, således kun betingelsen iht. OPHL § 11 a, stk. 1, nr. 3, hvis eksemplarfremstillingen, som kopierne har til formål at muliggøre, er godkendt af indehaverne af ophavsretten eller omfattet af en relevant undtagelse til ophavsretten.

Selvom man i lyset af ovenstående betragtninger kan argumentere for, at visse af de kopier, der bliver fremstillet ifm. lagringen af træningsdata i det system, som træningen af ML-algoritmen skal gennemføres i, formentlig opfylde betingelsen i OPHL § 11 a, stk. 1, nr. 3, bemærkes det imidlertid, at Axhamn generelt finder det tvivlsomt, hvorvidt den brug, som består i algoritmens kopiering af de analyserede værker, nogensinde vil kunne betegnes som "lovlig". Ifølge Axhamn er formålet med algoritmens kopiering, at algoritmen skal fremstille nye "værker", der ligner de værker, som algoritmen har analyseret, eller som i et eller andet omfang er inspireret af disse værker. Derfor er der ifølge Axhamn tale om en brug, som potentielt eller aktuelt konkurrerer med rettighedshaverne interesse.\textsuperscript{211} Da Axhamn henviser til rettighedshaverne interesse, synes hans synspunkt således at være funderet i en vurdering af den pågældende brug i lyset af "tretrinstesten", jf. InfoSoc-direktivets art. 5(5), der som påvist i afsnit 3.2 netop fastslår, at undtagelsen i art. 5(1) ikke må anvendes i visse specielle tilfælde, som ikke må indebære urimelig skade for rettighedshaverne legitime interess.

Imod Axhamns synspunkt kan det anføres, at det følger af Infopaq II, at hvis en reproduktionshandling opfylde alle betingelserne i InfoSoc-direktivets art. 5(1), "således som de er fortolket i Domstolens praksis", må det lægges til grund, at reproduktionshandlingerne er i overensstemmelse med

\textsuperscript{210} L 19 2002-03, s. 242
\textsuperscript{211} Axhamn (2021), s. 55
I den forbindelse følger det både af *Premier League* og *Infopaq II*, at en brug af et værk ifølge EUD anses for at være "lovlig", jf. art. 5(1), hvis den i overensstemmelse med direktivets 33. præambelbetrætning er godkendt af rettighedshaveren eller ikke er undergivet restriktioner iht. gældende lovgivning. I lyset deraf kan man argumentere for, at den brug af et værk, som består i algoritmens kopiering af dette værk ifm. dens analyse, når denne kopiering er omfattet af OPHL § 2, stk. 1, er i overensstemmelse med art. 5(5) og således rettighedshaverens legitime interesser, hvis indehaveren af ophavsretten har givet samtykke til brugen, eller hvis brugen er omfattet af en indskrænkning i ophavsretten. Man kan ligeledes argumentere for, at dette også er tilfældet, når algoritmens kopiering af det analyserede værk ikke er omfattet af OPHL § 2, stk. 1, som følge af at de gengivne elementer ikke er udtryk for ophavsmandens egen intellektuelle frembringelse. Dette understøttes af *Infopaq II*, hvor de eksemplarer, der blev fremstillet ifm. Infopaq’s datafangstproces, havde til formål "[...] at muliggøre en mere effektiv udarbejdelse af resuméer af avisartikler og dermed en brug heraf". Denne brug var ifølge EUD ikke underlagt restriktioner i dansk ret, hvorfor den derfor var "lovlig" i overensstemmelse med art. 5(1). På samme måde som med udarbejdelsen af resuméer af avisartikler kan man argumentere for, at en algoritmes kopiering af ikke-originale elementer er lovlig.

Til støtte for Axhamns synspunkt kan det anføres, at det modsætningsvist kan udledes af *Infopaq II*, at en reproduktionshandling, som ikke opfylder samtlige betingelser i art. 5(1), heller ikke er i overensstemmelse med art. 5(5). Hvis den brug, som en reproduktionshandling har til formål at muliggøre, ikke kan betegnes som "lovlig" iht. art. 5(1), er den pågældende brug således heller ikke i overensstemmelse med art. 5(5). Det kan på den baggrund anføres, at i de tilfælde, hvor algoritmens kopiering af de værker, den har analyseret, udgør en brug, som er omfattet af OPHL § 2, stk. 1, men hvor brugen hverken er godkendt af rettighedshaveren eller omfattet af en indskrænkning i ophavsretten, og derfor ikke kan betegnes som "lovlig", er der tale om en brug, som ligeledes er i strid med art. 5(5) og rettighedshaverens legitime interesser. I lyset af *Filmspeler* gør det samme sig formentlig gældende, hvor kravet om lovligt forlæg, jf. OPHL § 11, stk. 3, som beskrevet i afsnit 3.2 ikke er iagttaget. I *Filmspeler* fandt EUD, at den midlertidige eksemplarfremstilling, som fandt sted i en medieafspiller, når denne blev anvendt til streaming af beskyttede værker, som var

---

212 *Infopaq II*, pr. 55-57
214 *Infopaq II*, pr. 43
215 *Infopaq II*, pr. 44
216 Margoni (2018), pr. 19
217 *Infopaq II*, pr. 55-57
tilgængeliggjort på internettet uden rettighedshaveres samtykke, ikke opfyldte kravet om lovlig brug iht. art. 5(1). Ifølge EUD skyldtes dette bl.a. den omstændighed, at streamingen af de ulovligt tilgængelige værker udgjorde en brug af værket, der var i strid med art. 5(5) og herunder rettighedshaveres legitime interesser. Selvom den pågældende sag specifikt angik streaming af ulovligt tilgængelige værker, kan det anføres, at EUD's fortolkning af begrebet "lovlig brug" formentlig kan anvendes generelt om enhver brug. På den baggrund kan det anføres, at den brug af et værk, som består i algoritmens kopiering af værket, ikke kan betegnes som "lovlig" i overensstemmelse med OPHL § 11 a, stk. 1, nr. 3, uanset om kopieringen indebærer en ophavsretligt relevant eksemplarfremstilling, jf. OPHL § 2, stk. 1, eller ej, hvis den midlertidige kopi, der er blevet fremstillet ifm. lagringen af træningsdata, ikke er blevet fremstillet på baggrund af et lovligt forlæg, jf. OPHL § 11, stk. 3, da en sådan brug i lyset af Filmspeler må antages at stride imod rettighedshaveres legitime interesser, jf. art. 5(5).

Hvad angår den brug, som de kopier, der bliver fremstillet ifm. algoritmens generering af et output, har til formål at muliggøre, synes denne at bestå i udviklernes eftersyn og validering af outputtet, som bliver foretaget mhp. at vurdere, om der er behov for at finjustere algoritmen. Dette synes ikke at være en brug, der kan henføres til de ophavsretlige eneretsbeføjelser, jf. OPHL § 2, stk. 1, da der hverken er tale om eksemplarfremstilling eller tilgængeliggørelse for almenheden. Som følge heraf må det i lyset af de danske lovmotiver antages, at denne brug er "lovlig" i overensstemmelse med OPHL § 11 a, stk. 1, nr. 3.

Det kan derfor anføres, at de kopier, der bliver fremstillet ifm. algoritmens generering af output ifm. træningen, opfylder betingelsen iht. OPHL § 11 a, stk. 1, nr. 3. Dette taler desuden for, at kopieringen heller ikke er i strid med InfoSoc-direktivets art. 5(5).

Det skal endelig bemærkes, at da de kopier, som bliver fremstillet ifm. algoritmens generering af output, som beskrevet i afsnit 2.2 bliver fremstillet på baggrund af algoritmens anvendelse af den fremsgangsmåde, den har fastsat ved justeringen af dens bagvedliggende parametre, bliver kopieringen således ikke foretaget på baggrund af en gengivelse af værket, da kopierne af værkerne som anført ikke er lagret i algoritmen, hvorfor kravet om lovligt forlæg, jf. OPHL § 11, stk. 3, formentlig ikke er relevant i relation til denne kopiering.

218 Filmspeler, pr. 71-72
219 Filmspeler, pr. 70-71. Rosenmeier et al. (2018), s. 323. Pontén (2015), s. 580
220 Rosenmeier et al. (2018), s. 323
5.2.4 Ingen selvstændig økonomisk værdi

Endelig følger det af OPHL § 11 a, stk. 1, nr. 4, at et midlertidigt eksemplar kun er omfattet af undtagelsen, hvis eksemplaret "ikke har selvstændig økonomisk værdi".


I Premier League tog EUD's fortolkning af betingelsen afsæt i den overordnede betragtning, at da ophavsretligt beskyttede værker "[…] har en økonomisk værdi, må adgangen til dem således nødvendigvis have en økonomisk værdi".222

Med henvisning til Premier League fastslog EUD efterfølgende i Infopaq II, at det ikke kun er adgangen til ophavsretligt beskyttede værker men også "brugen" af sådanne værker, der har en økonomisk værdi.223 En midlertidig reproduktionshandling opfylder derfor kun betingelsen, hvis den ikke giver anledning til en yderligere økonomisk fordel, der går ud over den fordel, der opnås ved den lovlige brug af et beskyttet værk, som handlingen har til formål at muliggøre.224 Derudover må den økonomiske fordel, der opnås ved iværksættelsen af de midlertidige reproduktionshandlinger, "[…] ikke være forskellig fra eller kunne adskilles fra den økonomiske fordel, der opnås ved lovlig brug af det pågældende værk".225

Ifølge EUD er fordelen, der opnås ved midlertidige reproduktionshandlinger, forskellig eller adskillelig, "[…] hvis ophavsmanden til denne handling må antages at kunne opnå en fortjeneste ved en økonomisk udnyttelse af de midlertidige reproduktioner".226 Det samme gælder ifølge EUD, hvis reproduktionshandlingerne "[…] leder frem til en ændring af genstanden for reproduktionen, således som den forefindes på det tidspunkt, hvor den pågældende tekniske proces igangsættes".227

---

221 L 19 2002-03, s. 242
222 Premier League, pr. 174
223 Infopaq II, pr. 48
224 Ibid., pr. 50
225 Ibid., pr. 50
226 Ibid., pr. 52
227 Ibid., pr. 53
Hvad angår de kopier, som bliver fremstillet ifm. lagringen af træningsdata i det system, som træningen af ML-algoritmen skal gennemføres i, bliver disse kopier i praksis ofte solgt eller licenseret som en del af et samlet træningsdatsæt af den, som har udarbejdet datashættet, til andre fysiske eller juridske personer til brug for disses egne AI-udviklingsprojekter. På den baggrund kan man ifølge Chiou argumentere for, at de kopier, der bliver fremstillet ifm. lagringen af træningsdata, har en selvstændig økonomisk værdi og derfor ikke opfylder betingelsen iht. OPHL § 11 a, stk. 1, nr. 4. Chious opfattelse understøttes af Infopaq II, hvor EUD som beskrevet ovenfor fastslag, at den økonomiske værdi af et eksemplar er forskellig fra den økonomiske værdi af den lovlige brug, som eksemplaret har til formål at muliggøre, hvis den, der har fremstillet eksemplarerne, må antages at kunne opnå en økonomisk fortjeneste ved økonomisk udnyttelse af eksemplarerne. I modsætning til Chiou er Margoni imidlertid af den opfattelse, at den eventuelle økonomiske fordel, der vil kunne opnås, formentlig ligger i den lovlige bruge af værket (algoritmens analyse) og ikke i selve kopien. I lyset af de to modstridende opfattelser må det formentlig lægges til grund, at det afhænger af de pågældende kopiers beskaffenhed samt egnethed til at blive genbrugt til træningen af andre ML-algoritmer, om de opfylder betingelsen i OPHL § 11 a, stk. 1, nr. 4. Hvis kopierne ikke er egnet hertil, opfylder de formentlig betingelsen.


I den forbindelse skal det i øvrigt bemærkes, at selvom det som beskrevet i afsnit 3.2 følger af OPHL § 11, stk. 2, at et værk må ændres i den udstrækning, som den tilladte brug kræver, må det antages, at Infopaq II således udelukker, at der gælder nogen som helst adgang til at ændre værket, når dette

228 Chiou (2019), s. 406.
229 Ibid., s. 406
230 Infopaq II, pr. 52
231 Margoni (2018), s. 19
232 Axhamn (2021), s. 55
233 Infopaq II, pr. 53. Axhamn (2021), s. 55.
anvendes i overensstemmelse med OPHL § 11 a, stk. 1, da en ændring ifølge EUD indebærer, at kopien har en selvstændig økonomisk værdi.

Hvis algoritmens output derimod indebærer en eksemplarfremstilling af et værk, som algoritmen har analyseret, i dets oprindelige skikkelse (som det fx gjorde sig gældende ifm. Stable Diffusion-algoritmens genskabelse af billeder i deres oprindelige skikkelse ved fjernelsen af visuel støj som beskrevet i afsnit 2.2.2), jf. OPHL § 2, stk. 1, kan man dog argumentere for, at eksemplaret opfylder betingelsen iht. OPHL § 11 a, stk. 1, nr. 4. Sådanne eksemplarer bliver som anført i afsnit 2.2 kun brugt til at vurdere, om der er behov for at finjustere algoritmen, hvilket taler imod, at det har nogen som helst selvstændig økonomisk værdi, herunder da værket ikke er blevet ændret.

Sammenfattende kan det anføres, at visse af de kopier, der bliver fremstillet ifm. lagringen af træningsdata i det system, som træningen skal gennemføres i, samt ifm. algoritmens generering af output, opfylder betingelsen iht. OPHL § 11 a, stk. 1, nr. 4.

5.2.5 Delkonklusion

Den kopiering af træningsdata, der finder sted ifm. træningen af en ML-algoritme, som skal implementeres i et generative AI-system, når denne data består af ophavsretligt beskyttede værker, er dækket af undtagelsen for midlertidig eksemplarfremstilling, jf. OPHL § 11 a, stk. 1, og kræver derfor ikke tilladelse fra rettighedshaverne, hvor de fire kumulative betingelser, som er opregnet i bestemmelsen, er opfyldt. EUD's fortolkning af disse betingelser indebærer, at det imidlertid kun er visse af de kopier, der bliver fremstillet ifm. lagringen af træningsdata samt ifm. algoritmens generering af et output under træningen, der er dækket af indskrænkningen. Dette forudsætter bl.a., at kopieringen ikke indebærer en ændring af værket, jf. Infopaq II.234 OPHL § 11 a, stk. 1, dækker derimod ikke den kopiering af værker, der finder sted på baggrund af algoritmens analyse, hvorfor kopieringen i dette tilfælde kræver tilladelse fra rettighedshaverne i det omfang, at kopieringen er omfattet af OPHL § 2, stk. 1.

5.3 Undtagelsen for reproduktion af værker mhp. tekst- og datamining, jf. DSM-direktivets art. 4

Iht. DSM-direktivets art. 4(1) skal EU's medlemsstater fastsætte en obligatorisk indskrænkning i den eneret til reproduktion, der er hjemlet i en række af EU-direktiverne om ophavsret (herunder i

---

234 Infopaq II, pr. 53
InfoSoc-direktivets art. 2), som skal gælde for "[...] reproduktion og udtræk af lovligt tilgængelige værker eller andre frembringelser med henblik på tekst- og datamining".

Det skal indledningsvist bemærkes, at det følger af DSM-direktivets 18. præambelbetraktning, at formålet med undtagelsen bl.a. er at "skabe større retssikkerhed" i de tilfælde, hvor en bruger af tekst- og datamining (heri benævnt ‘TDM’) (hvad dette fænomen er, vil blive uddybet nedenfor i afsnit 5.3.1.1) foretager reproduktioner eller udtræk af beskyttede værker "med tekst- og datamining for øje", men hvor disse ophavsretligt relevante handlinger ikke er omfattet af undtagelsen for midlertidige reproduktioner i InfoSoc-direktivets art. 5(1).235 Det kan på den baggrund anføres, at DSM-direktivets art. 4(1) således potentielt vil kunne gribes visse af de kopier, der som fastslået i afsnit 5.2 ikke er omfattet af OPHL § 11 a, stk. 1. Dette kunne fx være de kopier, der bliver fremstillet i forbindelse med lagringen af træningsdata, når disse ikke kan betegnes som hverken flygtige eller tilfældige i overensstemmelse med § 11 a, stk. 1, nr. 1.

Derudover skal det bemærkes, at art. 4 er blandt de af DSM-direktivets bestemmelser, der endnu ikke er gennemført i dansk ret på trods af direktivets implementeringsfrist d. 7. juni 2021.236 I skrivende stund indeholder OPHL derfor ikke en dansk ækvivalent til art. 4, og Sverige og Finland er således de eneste af de nordiske EU-lande, der rent faktisk har gennemført DSM-direktivets art. 4 i deres respektive nationale ophavsretslove.237 Den 3. maj 2023 fremsatte den danske kulturminister dog et lovforslag238 om gennemførelsen af den resterende del af direktivet (herunder art. 4, som foreslås gennemført i den nye bestemmelse i OPHL § 11 b239), som må forventes at blive vedtaget i juli 2023.240

Indtil lovforslaget bliver vedtaget, og art. 4 derved bliver gennemført i dansk ret, vil indskrænkningen ikke kunne påberåbes af en bruger af et værk over for indehaveren af ophavsretten til dette værk, hvis brugen af værket finder sted i Danmark. Selvom EUD har fastslået, at et direktiv, der endnu ikke er blevet gennemført i en EU-medlemsstat, på trods af at implementeringsfristen er overskredet,

---

235 DSM-direktivets 18. præambelbetraktning
238 L 125 2022-23
239 Ibid., s. 49
240 Ibid., s. 4
undertiden har direkte virkning\textsuperscript{241}, kan denne virkning ifølge EUD aldrig være horisontal.\textsuperscript{242} En bestemmelse i et direktiv vil derfor aldrig kunne påberåbes over for en privatperson, før direktivet er blevet gennemført i national ret.\textsuperscript{243}

I det følgende skal det undersøges, i hvilket omfang undtagelsen i DSM-direktivets art. 4 dækker de tre tilfælde af kopiering af træningsdata, når denne data består af værker, der som beskrevet i afsnit 2.2.1 finder sted ifm. træningen af en \textit{ML}-algoritme.

Da art. 4 som fastslået ovenfor endnu ikke har været behandlet af EUD, tager undersøgelsen udgangspunkt i en fortolkning af undtagelsen iht. art. 4 i lyset af publikationer fra en række af EU’s institutioner samt relevant juridisk litteratur. Derudover vil det nyligt fremsatte danske lovforslag blive inddraget, hvor der er anledning til dette. Da det følger af lovforslaget, at Kulturministeriet har foreslået en "direktivnær implementering" af art. 4\textsuperscript{244}, må det dog antages, at lovforslaget kun i begrenset omfang vil kunne bidrage med betragtninger, som ikke allerede følger af direktivet. Dette understøttes af det forhold, at direktivet er udtryk for en totalharmonisering.\textsuperscript{245}

5.3.1 Betingelser for anvendelsen af DSM-direktivets art. 4(1)

Som undtagelsen eller indskrænkningen iht. DSM-direktivets art. 4(1) som gengivet ovenfor er formuleret, synes der at kunne opstilles to overordnede betingelser, der skal være opfyldt, for at en reproduktionshandling er omfattet af undtagelsen eller indskrænkningen.

For det første skal reproduktionshandlingen være foretaget "\textit{med henblik på} TDM.

For det andet skal der være tale om en reproduktion af "\textit{lovligt tilgængelige værker}".

Disse to betingelser vil blive behandlet hver for sig i det følgende.

5.3.1.1 Betingelse 1: "\textit{Med henblik på} TDM"

For at vurdere, om en reproduktion bliver genereret "\textit{med henblik på} TDM, er det først og fremmest nødvendigt at fastslå, hvad TDM i det hele taget er for et fænomen.

\textsuperscript{241} "Hvis et direktivs bestemmelser er retligt klare og utvetydige og kun overlader gennemførelsen til de nationale myndigheder, må de have en tilsvarende virkning som de traktatbestemmelser, som Domstolen har anerkendt som umiddelbart gældende"; jf. van Duyn
\textsuperscript{242} Marshall, pr. 48. Engsig Sørensen \textit{et al.} (2019), s. 175
\textsuperscript{243} Marshall, pr. 48
\textsuperscript{244} L 125 2022-23, s. 49
\textsuperscript{245} Se bl.a. DSM-direktivets 1. og 2. præambelbetrægtning
Det følger af DSM-direktivets 8. præambelbetragtning, at TDM almindeligvis forstås som et begreb, der dækker over "[...] en automatisk computeranalyse af information i digital form, fx tekst, lyd, billeder og data [...]". Ifølge EU-lovgiver gør TDM det muligt at behandle store mængder af information "[...] med det formål at opnå ny viden og finde frem til nye tendenser". På denne baggrund er begrebet TDM defineret i DSM-direktivets art. 2, nr. 2, som "enhver automatiseret analytisk fremgangsmåde, der har til formål at analysere tekst og data i digital form med henblik på at generere oplysninger, herunder, men ikke begrænset til, mønstre, tendenser og korrelationer".

På trods af at de klare ligheder, der synes at kunne påpeges mellem TDM som beskrevet ovenfor og ML som beskrevet i afsnit 2.2, (herunder det forhold, at de begge indebærer en automatiseret analyse af en større mængde data mhp. at identificere bestemte informationer i denne data), er der imidlertid tale om to forskellige teknologiske fænomener.

Hvor den automatiserede analyse af en given mængde data ifm. TDM (som det ligeledes følger af definitionen i DSM-direktivets art. 2, nr. 2) bliver foretaget mhp. at generere oplysninger i form af bl.a. sammenhænge, mønstre o.l., er en sådan analyse blot en af de teknikker, der anvendes til at realisere formålet med træningen af en ML-algoritme, der som fastslået i afsnit 2.2 ovenfor er at lære algoritmen at generere en bestemt form for output, så den kan implementeres i et AI-system. Med andre ord er det den trænede algoritme, der er resultatet af ML, hvorimod resultatet af TDM er de genererede oplysninger i form af mønstre, sammenhænge o.l.

Da ML og TDM altså er to forskellige teknologiske fænomener, kunne man allerede på den baggrund argumentere for, at art. 4 ikke dækker den kopiering af værker, der bliver foretaget ifm. træningen af en ML-algoritme, da bestemmelsen efter sin ordlyd kun nævner reproduktioner af værker "med henblik på TDM" og ikke mhp. ML.

Imidlertid synes der at være bred enighed blandt de juridiske forfattere om, at DSM-direktivets art. 4 også har en vis relevant i relation til den kopiering af værker, der bliver foretaget ifm. træningen af en ML-algoritme.

Dette skyldes for det første, at ML og TDM gør brug af algoritmer, der har en lang række fællestræk.

---

246 DSM-direktivet, 8. præambelbetragtning
247 Keller (2023)
248 The exception for text and data mining (TDM) (2018), s. 2. Brooks (2017)
249 Brooks (2017)
250 The exception for text and data mining (TDM) (2018), s. 2. Axhamn (2021), s. 12
For det andet skyldes det, at begge teknologiske fænomener langt hen ad vejen indebærer de samme rutineprægede handlinger. Både TDM og ML indebærer i praksis en indsamling, konvertering og lagring af en mængde data, som herefter bliver gjort til genstand for en automatiseret analyse.


Hertil bemærkes det, at definitionen i DSM-direktivets art. 2, nr. 2, efter sin ordlyd omfatter "enhver automatiseret analytisk fremgangsmåde [min understregning]", der har til formål at analysere tekst og data i digital form mhp. at generere oplysninger i form af bl.a. mønstre, sammenhænge o.l. Som følge af denne særdeles bredt formulerede definition kan man argumentere for, at en ML-algoritmes analyse af træningsdata, der som anført i afsnit 2.2 netop er en analyse af værker, der forløber automatisk, og som i visse tilfælde bliver gennemført mhp., at algoritmen skal identificere bestemte sammenhænge mellem elementer i disse værker, også er omfattet af definitionen i det omfang, at analysen kan henføres under det formål, der følger af DSM-direktivets art. 2, nr. 2.

I lyset af ovenstående betragtninger synes DSM-direktivets art. 4 derfor også relevant ift. den kopiering, som finder sted ifm. treningen en ML-algoritme. Dette understøttes af den omstændighed som anført af EU-Kommissionen, at den brede definition i direktivets art. 2, nr. 2, medfører, at art. 4 ligeledes får et bredt anvendelsesområde.

Som anført ovenfor skal en reproduktion af et værk være foretaget "med henblik på" TDM (eller med TDM "for øje" for at gengive EU-lovgivers ord) for at opfylde den første betingelse, der kan udledes af art. 4(1). Som følge af definitionen på TDM i direktivets art. 2, nr. 2, skal reproduktionen af et ophavsretligt beskyttet værk derfor være foretaget "med henblik på" en automatiseret fremgangsmåde, der har til formål at analysere tekst og data i digital form mhp. at generere oplysninger i form af sammenhænge, mønstre o.l., for at reproduktionen er omfattet af undtagelsen i art. 4(1). På den baggrund kan det anføres, at den kopiering, der finder sted ifm. træningen af en

251 A literature review (2021), s. 10  
252 Study on the legal framework of TDM (2014), s. 28. Margoni (2018), s. 20. Geiger et al. (2018)., s. 817  
253 Ducato et al. (2018), s. 2. Study on the legal framework of TDM (2014), s. 7.  
254 EU-Kommissionens rapport om AI (2022), s. 196  
255 Ibid., s. 195  
256 DSM-direktivets 18. præambelbetragtning  
257 EU-Kommissionens rapport om AI (2022), s. 195
**ML-algoritme** opfylder den første betingelse iht. art. 4(1) i det omfang, at kopieringen bliver foretaget mhp. et formål, der kan henføres til definitionen i art. 2, nr. 2.  

Hvad angår de kopier af værker, der er blevet fremstillet ifm. lagringen af disse værker i det computersystem, som træningen af **ML-algoritmen** skal gennemføres i, danner disse som beskrevet i afsnit 2.2 grundlag for algoritmens analyse. I det omfang at algoritmen analyserer kopierne mhp. at identificere sammenhænge, mønstre o.l. i overensstemmelse med formålet iht. art. 2, nr. 2, kan man argumentere for, at kopieringen bliver foretaget "mhp." **TDM**, og at kopieringen derfor er omfattet af DSM-direktivets art. 4(1). I et sådant tilfælde synes kopierne netop at kunne siges at være blevet fremstillet "med henblik på" gennemførelsen af en fremgangsmåde, der er omfattet af definitionen af **TDM** iht. DSM-direktivets art. 2, nr. 2. I den forbindelse skal det bemærkes, at det ikke har nogen betydning for anvendelsen af undtagelsen eller indskrænkningen, at algoritmen analyserer de pågældende værker mhp. at lære, hvordan den selvstændigt skaber et "kreativt" output, der i et eller andet omfang er inspireret af de analyserede værker; en brug af værkerne, der som beskrevet i afsnit 5.2.3 ifølge **Axhamn** potentielt eller aktuelt konkurrerer med rettighedshaverernes interesser.  

Det følger således af de danske lovomtiver, at der ikke stilles krav om, at **TDM** "[…] skal ske til et specifikt formål", for at en reproduktion er omfattet af undtagelsen.  

Ifølge det danske lovforslag er adgangen til eksemplarfremstilling derfor "[…] alment gældende, så længe det er med henblik på datamining". Denne opfattelse af undtagelsen anvendelsesområde hersker også på EU-plan.  

Hvad angår de kopier, der bliver fremstillet ifm., at algoritmen analyserer den anvendte træningsdata samt ifm., at algoritmen genererer et output, synes det derimod udelukket, at kopieringen i disse tilfælde er omfattet af art. 4(1). Da kopierne bliver fremstillet i lyset af de informationer, som algoritmen har identificeret ifm. dens analyse af den anvendte træningsdata, synes der snarere at være tale om kopiering "på baggrund af" **TDM** end reproduktion "med henblik på" **TDM**. Da art. 2, nr. 2, kun nævner reproduktioner mhp. en analyse af data, synes det derfor tvivlsomt, at art. 4 også finder anvendelse på kopier, som bliver fremstillet på baggrund af analysen. Da det output, som er resultatet af gennemførelsen af **TDM**, skal bestå i oplysninger i form af sammenhænge, mønstre o.l., jf. DSM-direktivets art. 2, nr. 2, kan det desuden anføres, at definitionen således synes at udelukke,

---

---

258 Chiou (2019), s. 409. **Axhamn** (2021), s. 56  
259 **Axhamn** (2021), s. 55  
260 **Ibid.**, s. 50  
261 **Ibid.**, s. 205
262 **EU-Kommissionens rapport om AI** (2022), s. 195.  
263 **Ibid.**, s. 205
at outputtet er et ophavsretligt beskyttet værk. Det må derfor antages, at de kopier af værker, der bliver fremstillet på baggrund af algoritmens analyse og ifm. algoritmens generering af output under træningen, ikke er omfattet af DSM-direktivets art. 4.

Sammenfattende kan det derfor anføres, at det kun er de kopier, der bliver fremstillet ifm. lagringen af beskyttede værker i det computersystem, som træningen af algoritmen skal gennemføres i, der opfylder den første betingelse, der kan udledes af DSM-direktivets art. 4(1). Dette forudsætter dog, at algoritmens analyse kan henføres til definitionen i DSM-direktivets art. 2, nr. 2.

5.3.1.2 Betingelse 2: "Lovligt tilgængelige værker"

Som fastslået ovenfor kan det som den anden betingelse udledes af DSM-direktivets art. 4(1), at der skal være tale om reproduktion af "lovligt tilgængelige værker".

Det følger af DSM-direktivets 18. præambelbetractning, at den, der udfører TDM, bl.a. har tilgået et værk lovligt, når dette værk 

\[\ldots\] stilles til rådighed for almenheden online, og i det omfang rettighedshaverne ikke på passende vis har forbeholdt sig retten til at foretage reproduktioner og udtækt til tekst- og dataminingformål\].

264 Således giver DSM-direktivets art. 4 bl.a. adgang til reproduktion af værker, der er lovligt tilgængelige på internettet.

Dog skal det bemærkes, at betingelsen om "lovligt tilgængelige værker" næppe bringer noget nyt til dansk ophavsret, da det allerede følger af OPHL § 11, stk. 3, som beskrevet i afsnit 3.2, at eksemplarfremstilling, som foretages i overensstemmelse med en af bestemmelserne i OPHL kap. 2, kun må foretages på grundlag af et lovligt forlæg (eksemplarer eller gengivelser af værket, som er godkendt af rettighedshaveren eller omfattet af en indskrænkning i ophavsretten). Det fremgår derfor ikke overraskende af de danske lovmotiver, at anvendelsen af den foreslåede bestemmelse i OPHL § 11 b, forudsætter et lovligt forlæg, jf. OPHL § 11, stk. 3.

Dog understreger DSM-direktivets 18. præambelbetractning imidlertid en anden væsentlig omstændighed af betydning for adgangen til at reproducerere beskyttede værker i overensstemmelse med DSM-direktivets art. 4(1), nemlig den, at denne adgang begrænser af rettighedshavers ret til at forbeholde sig sine rettigheder. Iht. DSM-direktivets art. 4(3) er en reproduktion af et værk, der bliver foretaget mhp. TDM, nemlig kun omfattet af indskrænkningen eller undtagelsen i art. 4(1), under forudsætning af at rettighedshaveren"[\ldots] ikke udtrykkeligt har forbeholdt sig anvendelsen heraf på

264 DSM-direktivets 18. Præambelbetractning
passende vis, såsom en maskinlæsbar måde i tilfælde af indhold, der er stillet til rådighed for almenheden online".

Rettighedshaverens ret til at forbeholde sig brugen af sit værk til TDM-formål, er en såkaldt "opt-out-løsning", som indebærer, at et værk frit kan anvendes af andre end rettighedshaveren til TDM-formål, jf. DSM-direktivets art. 4(1), hvis det er "lovligt tilgængeligt", medmindre rettighedshaver – på baggrund af et forbehold – har undtaget brugen af det pågældende værk til dette formål i overensstemmelse med art. 4(3).265

Det bemærkes, at et sådant forbehold ifølge EU-Kommissionen godt kan være delvist og således kun omfatte nærmere afgrensende TDM-formål. Rettighedshaver kan fx gyldigt forbeholde sig sine rettigheder i stand til at skabe konkurrerende kreationer, men fortsat lade tilsvarende reproducitioner være tilladt i medfør af art. 4(1), hvis analysen derimod har til formål at træne en algoritme til at identificere manipulationer af digitalt indhold (såkaldte deepfakes).266

Hvad angår rettighedshavers udøvelse af sin opt-out-ret, kan det for det første udledes af DSM-direktivets art. 4(3) som gengivet ovenfor, at rettighedshaver skal forbeholde sig sine rettigheder "på passende vis". I den forbindelse sondrer EU-lovgiver iht. DSM-direktivets 18. præambelbetragtning imellem tilfælde, hvor værket er stillet til rådighed for almenheden online og "andre tilfælde".267

I de førstnævnte tilfælde skal rettighedshaver, som det ligeledes udtrykkeligt følger af DSM-direktivets art. 4(3), forbeholde sig sine rettigheder på en "maskinlæsbar måde". Dette kan ifølge EU-lovgiver bl.a. være "[…] i metadata og i vilkår og betingelser for brug af et websted eller en tjeneste".268 Et eksempel på et forbehold i metadata er brugen af den såkaldte "robot.txt"-standard på en hjemmeside.269 Denne standard fortæller digitale søgemaskineroboter o.l., hvilke dele af indholdet på en hjemmeside, robotten må undersøge.270 Dette er således bl.a. et værn mod web scraping. I alle andre tilfælde kan det ifølge EU-lovgiver "[…] være passende at forbeholde sig rettighederne på anden måde, såsom i kontraktmæssige aftaler eller ved en ensidig erklæring".271

265 Chiou (2019), s. 409. EU-Kommissionens rapport om AI (2022), s. 195.
266 EU-Kommissionens rapport om AI (2022), s. 196.
267 Axhamn (2021), s. 12. EU-Kommissionens rapport om AI (2022), s. 196.
268 DSM-direktivets 18. præambelbetragtning
271 DSM-direktivets 18. præambelbetragtning.
Foruden at fastslå, at forbeholdet skal tages "på passende vis", følger det for det andet af DSM-direktivets art. 4(3) som gengivet ovenfor, at rettighedshaver "udtrykkeligt" skal have forbeholdt sig sine rettigheder. Det synes således at kunne udledes af bestemmelsens ordlyd, at uklare eller tvetydige forbehold ikke har virkning over for en brug af værker til TDM, der ellers opfylder kriterierne iht. DSM-direktivets art. 4(1). Hverken DSM-direktivets præambelbetrægninger eller bestemmelser uddyber imidlertid begrebet "udtrykkeligt", og herunder hvad og hvor meget der konkret skal til, for at rettighedshavers forbehold kan siges at være "udtrykkeligt".

Det bemærkes, at netop denne uklarhed vedrørende betydningen af begrebet "udtrykkeligt", som det er anvendt i DSM-direktivets art. 4(3), tilsyneladende har givet anledning til en vis retlig uenighed på EU-plan. I et studie af ophavsret og nye teknologier (herunder ML og TDM), som blev gennemført af EU-Kommissionen i 2022, herskede der vidt forskellige opfattelser af begrebet "udtrykkeligt" blandt en lang række adspurgte juridiske ekspert.272 Hvor nogle eksperter mente, at begrebet skulle forstås således, at rettighedshaver er tvunget til at skulle forholde sig konkret til enhver brug af værket mhp. TDM, mente andre ekspert, at rettighedshavers udøvelse af sin forbeholdsret kun kræver en minimal indsat, og at det således er i overensstemmelse med art. 4(3), hvis rettighedshaver forbeholder sig sine rettigheder generelt ifm. en tilrådighedsstillelse af værket.273

Som følge af disse modstridende opfattelser blandt ekspetere udtrykte EU-Kommissionen stor bekymring for gennemførelsen af art. 4(3) i medlemsstaterne nationale ret, da der hersker en risiko for en uensartet implementering af bestemmelsen.274 Det kan da interessant nok også konstateres, at det alene er begrebet "på passende vis", der har fundet vej til den svenske ophavsretsløvs § 15 a, i hvilken DSM-direktivets art. 4(3) er gennemført.275 Heroverfor følger det af forslaget til den nye danske bestemmelse i OPHL § 11 b, at rettighedshaver "udtrykkeligt" skal have forbeholdt sig anvendelsen af værket til TDM på passende vis.276

Uanset hvordan kravet om udtrykkelighed skal fortolkes, kan det dog overordnet set anføres, at undtagelsen eller indskrænkningen i DSM-direktivets art. 4(1) ikke finder anvendelse på en kopiering af et værk, som bliver foretaget ifm. lagringen af træningsdata i det system, som træningen af en ML-

272 EU-Kommissionens rapport om AI (2022), s. 200
273 Ibid., s. 200
274 Ibid., s. 200
275 "Första stycket gäller inte om upphovsmannen på lämpligt sätt har förbehållit sig den rätt som avses där" – den svenske ophavsretsløvs § 15 a
276 L 125 2022-23, s. 82
algoritme skal gennemføres i, hvis rettighedshaver har forbeholdt sig sine rettigheder i overensstemmelse med art. 4(3).

Foruden rettighedshavers ret til at forbeholde sig sine rettigheder, jf. art. 4(3), følger det derudover af DSM-direktivets art. 7(1), at alle aftalebestemmelser, der er i strid med direktivets art. 3, 5 og 6 er uden retskraft. Det gælder således modætningsvist, at DSM-direktivets art. 4(1) kan fraviges ved en aftale mellem en bruger af TDM og rettighedshaver, jf. art. 7(1).277 En rettighedshaver er således ikke forhindret i at udstede licenser til brug af sine værker.278 Dette udgør således endnu en væsentlig begrensning i en brugers adgang til at påberåbe sig DSM-direktivets art. 4(1).

I lyset af ovenstående kan det sammenfattende fastslås, at den kopiering af træningsdata, når denne data består af beskyttede værker, der finder sted ifm. lagringen af træningsdata i det system, som træningen af en ML-algoritme skal gennemføres i, opfylder den anden betingelse, der kan udledes af DSM-direktivets art. 4(1), i det omfang, at der er tale om værker, som den, der forestår træningen, har tilgået på lovlig vis. Det er bl.a. tilfældet, hvis de pågældende værker er offentligt tilgængelige på internettet med rettighedshavers samtykke. De pågældende værker må imidlertid ikke anvendes til træningen af en ML-algoritme, hvis rettighedshaver har forbeholdt sig sine rettigheder i overensstemmelse med art. 4(3), eller hvis rettighedshaver og den, der forestår træningen, har fraveget art. 4(1) ved aftale, jf. art. 7(1). Om betingelsen er opfyldt i praksis, afhænger dog i sagens natur af en konkret vurdering af de værker, der indgår i den anvendte træningsdata i det enkelte ML-træningstilfælde.

5.3.2 Opbevaring så længe det er nødvendigt, jf. DSM-direktivets art. 4(2)

Iht. DSM-direktivets art. 4(2) kan reproductioner, der opfylder kravene for anvendelsen af undtagelsen eller indskrænkningen iht. art. 4(1), "[...] opbevares, så længe det er nødvendigt med henblik på TDM". Det følger imidlertid ikke af direktivet, hvordan den tilladte opbevaringsperiode konkret skal afgrænses, andet end at opbevaringen skal være nødvendig. Bestemmelsen er heller ikke uddybet i særlig stor grad i det danske lovforslag. Det følger blot heraf, at der er tale om en begrensning i opbevaringsperioden, hvorfor den, der foretager dataminingen, "[...] skal gøre sig overvejelser om – og sikre – at dataet ikke opbevares længere, end hvad der er nødvendigt til brug for dataminingen".279

277 Axhamn (2021), s. 12
278 L 125 2022-23, s. 50. Lizarralde (2020)
279 L 125 2022-23, s. 50f.
Hvad angår tidsrummets længde, kan det på den ene side anføres, at da DSM-direktivets art. 4(1) kun finder anvendelse på reproductioner mhp. en automatiseret analyse af data i digital form, jf. art. 2, nr. 2, taler dette for, at opbevaringen kun er "nødvendig" indtil det tidspunkt, hvor reproduktionerne er blevet analyseret.280

På den anden side kan man dog argumentere for, at opbevaringen fortsat synes at kunne betragtes som "nødvendig" mhp. TDM, på trods af at analysen allerede er blevet gennemført, hvis algoritmen har udledt de forkerte oplysninger af den analyserede data, og analysen af denne data derfor må gennemføres igen.

Indtil der foreligger et svar fra EUD om fortolkningen af DSM-direktivets art. 4(2), er det uklaart, hvor længe de eksemplarer, der er blevet fremstillet ifm. lagringen af træningsdata i det system, som træningen skal gennemføres i, må opbevares.

5.3.3 Delkonklusion
Kopieringen af træningsdata, når denne data består af beskyttede værker, der finder sted i forbindelse med træningen af en ML-algoritme, som skal implementeres i et generative AI-system, er dækket af undtagelsen i DSM-direktivets art. 4, og kræver således ikke tilladelse fra rettighedshaverne, i det omfang, at algoritmens analyse af de pågældende værker kan henføres under den definition af TDM, der følger af direktivets art. 2, nr. 2. Som følge heraf er det kun den kopiering af værker, der finder sted i ifm. lagringen af træningsdata i det system, som træningen af ML-algoritmen skal gennemføres i, der er dækket af indskrænkningen, da denne kopiering er det eneste af de tre tilfælde af kopiering, der opfylder betingelsen, efter hvilken kopieringen skal foretages "med henblik på" TDM. Derudover er kopieringen kun dækket af indskrænkningen i det omfang, at de øvrige betingelser for indskrænkningens anvendelse er opfyldt, herunder at der er tale om kopiering af lovlige tilgængelige værker, og at indehaverne af ophavsretten til værkerne ikke har forbeholdt sig deres rettigheder til TDM-formål, jf. art. 4(3).

5.4 Den generelle aftalelicens, jf. OPHL § 50, stk. 2
Indledningsvist skal der påpeges den væsentlige forskel fra de to andre indskrænkninger, der blev behandlet ovenfor, at indskrænkningen i OPHL § 50, stk. 2, er en aftalelicensregel og ikke en fribrugsregel. Derfor giver OPHL § 50, stk. 2, (i modsætning til OPHL § 11 a, stk. 1, og DSM-direktivets art. 4(1)) ikke hjemmel til, at en bruger kan bruge beskyttede værker på en bestemt måde

280 Gerrish et al. (2019), s. 61
helt uden tilladelse hertil.\textsuperscript{281} Dog giver indskrænkningen i OPHL § 50, stk. 2, hjemmel til, at brugeren kan klarere den pågældende brug af værkerne gennem en aftale med en rettighedsorganisation, hvorved brugeren således undgår at skulle indhente individuelle tilladelser fra samtlige rettighedshavere.\textsuperscript{282}


Spørgsmålet er herefter, i hvilket omfang OPHL § 50, stk. 2, dækker den kopiering af træningsdata, når denne data består af værker, der finder sted ifm. træningen af en \textit{ML}-algoritme. Dette vil blive undersøgt i det følgende. Da der som tidligere anført ikke er indgået nogen aftaler om kopiering af værker ifm. træningen af en \textit{ML}-algoritme\textsuperscript{283}, vil indskrænkningens omfang blive fastslået med afsæt i lovmotiverne samt allerede eksisterende licensaftaler om kopiering i andre sammenhengen.

5.4.1 En aftale om en brug af et værk inden for et nærmere defineret område, jf. § 50, stk. 2

Det følger af OPHL § 50, stk. 2, at aftalelicens kan påberåbes, hvis en bruger har indgået en aftale om udnyttelelsen af værker ”inden for et nærmere defineret område” med en rettighedsorganisation, men det fremgår ikke af bestemmelsen, hvordan dette udnyttesområde konkret skal afgrænses.

Iht. lovmotiverne er § 50, stk. 2, dog bl.a. relevant for helt nye udnyttesområder, der ikke er omfattet af de eksisterende aftalelicenser i OPHL.\textsuperscript{284} Derudover følger det af lovmotiverne, at OPHL § 50, stk. 2, udgør en generel hjemmel til aftalelicens på ”[… områder, hvor både rettighedshavere og brugere finder behov for det og kan blive enige om det”\textsuperscript{285}.

Selvom det bemærkes, at den generelle aftalelicens ifølge lovmotiverne bl.a. var funderet i et politisk ønske om at imødegå de praktiske vanskeligheder forbundet med klarering af rettigheder

\textsuperscript{281} Schovsbo \textit{et al.} (2021), s. 170f.
\textsuperscript{282} Ibid., s. 171 og 215
\textsuperscript{283} Kulturministeriets oversigt over aftalelicensgodkendelser (se under fanen ”Godkendelser efter den generelle aftalelicens”) – \url{https://kum.dk/kulturomraader/aftalelicensgodkendelser} (lokaliseret d. 9. maj 2023 kl. 16:47)
\textsuperscript{284} L 58 2007-08, afsnit 3.2. Et eksempel på en sådan aftalelicens er den, der følger af OPHL § 13, stk. 1, som bl.a. tillader eksemplarfremsstilling af udgivne værker, hvis betingelserne for aftalelicens iht. OPHL § 50, stk. 1, er opfyldt.
\textsuperscript{285} Ibid., afsnit 3.3.1
(herunder rettigheder til såkaldte "orphan works" som vil blive beskrevet nærmere nedenfor) ifm. digitalisering af kulturarven286, er anvendelsesområdet for OPHL § 50, stk. 2, i lyset af lovmotiverne bredere end som så. Aftalelicens kan således påberåbes i medfør af OPHL § 50, stk. 2, for enhver værkstype og enhver udnuttelsesform i det omfang, at betingelserne for anvendelsen af bestemmelsen er opfyldt og herunder, at brugeren og rettighedsorganisationen kan blive enige om udnuttelsen samt vilkårene og vederlaget for denne.287


Det følger imidlertid af lovmotiverne, at aftalelicens "[...] er en alternativ løsning, der kan benyttes, hvis individuelle aftaler direkte med ophavsmandene ikke er praktisk muligt [min understregning]".289 Den generelle aftalelicens iht. OPHL § 50, stk. 2, er således subsidiær ift. individuelle aftaler med rettighedshavere.290

Spørgsmålet er herefter, hvor meget der skal til, før individuelle aftaler ikke er "praktisk muligt".

Iht. lovmotiverne er individuelle aftaler ikke praktisk muligt, hvis der er tale om en masseudnyttelse af "omfattende materiale med mange rettighedshavere".291 Det samme gælder ifm. udnuttelsen af "orphan works".292 "Orphan works" er typisk ældre værker, hvor beskyttelsestiden ikke er udløbet, og hvor det "[...] efter en rimelig indsats ikke er muligt at identificere eller lokaliserre rettighedshaverne via kendte registre, forlag og lignende".293 I lyset af lovmotiverne synes OPHL §

---

286 Ibid., afsnit 3.2. Rapport om digitalisering af kulturarven (2007), s. 60
287 L 58 2007-08, afsnit 3.3.1. Kyst (2009), s. 47. Schovsbo et al. (2021), s. 215.
288 Kyst (2009), s. 48
289 L 58 2007-08, afsnit 11
290 Kyst (2009), s. 50
291 L 58 2007-08, afsnit 3.3.1
292 L 58 2007-08, afsnit 3.3.2.
293 L 58 2007-08, afsnit 3.3.2. Rapport om digitalisering af kulturarven (2007), s. 56. Schønning (2021), s. 553.
50, stk. 2, således kun at være relevant, hvor der er tale om en masseudnyttelse af værker, og hvor det i den forbindelse er umuligt at identificere indehaverne af ophavsretten til de pågældende værker.

Skæver man derimod til Kulturministeriets administrative afgørelser om godkendelse af representative rettighedsorganisationer, synes virkeligheden imidlertid ikke at være helt så snæver, som den kommer til udtryk i lovmotiverne.

En rettighedsorganisation, som har indgået en aftale med en bruger af værker iht. OPHL § 50, stk. 2, skal godkendes af kulturministeren til at indgå aftaler på det pågældende område, jf. OPHL § 50, stk. 4, og ligeledes nedenfor i afsnit 5.4.2. Kulturministeriets afgørelser om godkendelsen af rettighedsorganisationer er offentligt tilgængelige på Kulturministeriets hjemmeside.294

I en offentliggjort godkendelse fra 31. marts 2023 blev forvaltningsorganisationen Copydan Tekst & Node godkendt, jf. OPHL § 50, stk. 4, jf. § 50, stk. 2, i relation til en aftale med Foreningen lex.dk om eksempelarafremstilling af en række bogudgivelser fra Gyldendal i digital form.295 Det følger af begrendelsen for godkendelsen, at den konkrete udnyttelse 

"[...] vedrører et stort antal værker og rettighedshavere, og det vurderes, at individuel klarering vil være forbundet med betydelige administrative byrder [min understregning]."296 Netop denne formulering ses også i andre afgørelser. På den baggrund kan det derfor anføres, at OPHL § 50, stk. 2, også er relevant i tilfælde, hvor rettighedshaverne er identificeret, men hvor det er for administrativt tungt at indgå individuelle aftaler med disse rettighedshavere om udnyttelsen, da der er tale om et stort antal værker og rettighedshavere. Her synes formålet med aftalelicensen, som iht. lovmotiverne er at sikre betaling for en udnyttelse, som ellers ville have fundet sted uden betaling, i øvrigt netop at komme til sin magt.297

Som anført i afsnit 2.2 bliver træningen af en ML-algoritme i praksis ofte gennemført ved brug af træningsdata, der består af tusindvis (og undertiden millionvis) af beskyttede værker. I et sådant tilfælde er der utvivlsomt tale om masseudnyttelse af disse værker. Da de pågældende værker desuden ofte bliver *scraped* fra et stort antal kilder på internettet, må det antages, at det formentlig undertiden vil kræve særdeles store administrative ressourcer at finde frem til samtlige af indehaverne af ophavsretten til den store mængde værker, hvis det i det hele taget er muligt at identificere de pågældende rettighedshavere. Man kan derfor argumentere for, at det ikke altid kan siges at være

---

294 Kulturministeriets oversigt over aftalelicensgodkendelser (se under fanen “Godkendelser efter den generelle aftalelicens”) – [https://kum.dk/kulturomraader/aftalelicensgodkendelser](https://kum.dk/kulturomraader/aftalelicensgodkendelser) (lokaliseret d. 9. maj 2023 kl. 16:47)
295 Kulturministeriets afgørelse af 31. marts 2023
297 L 58 2007-08, afsnit 3.3.1.
"praktisk muligt" at indgå individuelle aftaler med indehaverne af ophavsretten, hvilket taler for, at kopieringen, der finder sted ifm. træningen, vil kunne klareres gennem en aftale med en rettighedsorganisation i medfør af OPHL § 50, stk. 2. Dette vil derimod i mindre grad være tilfældet, hvis der er tale om et væsentligt lavere antal beskyttede værker, og hvor samtlige værker kommer fra klart identificerede kilder.

Om individuelle aftaler med rettighedshaverne er "praktisk muligt" eller ej afhænger dog selvsagt af omstændighederne, herunder oprindelsen og mængden af træningsdata. I lyset af ovenstående betragtninger kan det dog sammenfattende anføres, at indskrænkningen iht. OPHL § 50, stk. 2, dækker kopieringen af træningsdata bestående af værker, der finder sted i tre tilfælde ifm. træningen af en ML-algoritme, i det omfang, det kan fastslås, at det ikke er "praktisk muligt" at indgå aftaler med rettighedshaverne til værkerne, og at udviklinger af ML-algoritmen kan nå til enighed med en rettighedsorganisation om en aftale. Der er dog en række yderligere forhold af betydning for indskrænkningens omfang, der vil blive uddybet i det følgende.

5.4.2 En repræsentativ rettighedsorganisation som er godkendt af Kulturministeriet, jf. § 50, stk. 4

Derudover skal organisationen som fastslået ovenfor være godkendt af kulturministeren, jf. OPHL § 50, stk. 4, til at indgå den pågældende aftale. Godkendelsen har ifølge lovomtiverne til formål at sikre, at der kun bliver godkendt én repræsentativ organisation inden for hver værksart. Det kan på den baggrund anføres, at det er den konkrete værkstype, der skal anvendes til træningen af algoritmen, som er afgørende for, hvilken repræsentativ organisation uddybetes skal klareres hos.

298 Schønning (2021), s. 550
299 Ibid.
300 L 58 2007-08, afsnit 11
Hvis de to betingelser ikke er opfyldt, kan en aftale med en rettighedsorganisation ikke tillægges aftalelicensvirkning, jf. OPHL § 50, stk. 2, og aftalen giver således kun brugeren ret til at udnytte værker, som er skabt af ophavsmænd, der er repræsenteret af rettighedsorganisationen.\(^{301}\)

5.4.3 **Begrensningen: Rettighedshavers forbud efter OPHL § 50, stk. 2, 2. pkt.**

Derudover følger det af OPHL § 50, stk. 2, 2. pkt., at aftalelicens ikke kan påberåbes af en bruger for en udnyttelse af et værk, hvis indehaveren af ophavsretten til det pågældende værk har nedlagt forbud mod værkets udnyttelse over for enten brugeren eller den repræsentative organisation. Iht. lovmotiverne skal et sådant forbud nedlægges personligt og individuelt over for brugeren eller organisationen med konkret angivelse af de værker, der skal være omfattet af forbuddet.\(^{302}\)

Selvom det følger af lovmotiverne, at generelle tilkendegivelser om, at værket ikke må bruges i en bestemt sammenhæng, ikke har virkning som et forbud\(^{303}\), sætter rettighedshaverens forbudsret selvsagt en begrænsning på adgangen til at bruge beskyttede værker i de tilfælde, hvor brugen er klareret i overensstemmelse med OPHL § 50, stk. 2.

5.4.4 **Delkonklusion**

Den kopiering af træningsdata, der finder sted i ifm. træningen af en *ML*-algoritme, som skal implementeres i et *generative AI*-system, når denne data består af ophavsretligt beskyttede værker, er dækket af indskrænkningen iht. OPHL § 50, stk. 2, i det omfang, at det ikke er praktisk muligt at indgå individuelle aftaler med rettighedsøverne til de værker, der indgår i den anvendte træningsdata. Dette er bl.a. tilfældet, hvis det er forbundet med betydelige administrative byrder at indgå sådanne aftaler. Derudover skal udviklerne af algoritmen kunne nå til enighed om en aftale om udnyttelsen (herunder vilkårene og vederlaget for udnyttelsen) med en repræsentativ rettighedsorganisation, der er godkendt af Kulturministeriet, jf. OPHL § 50, stk. 4. Da den generelle aftalelicens, jf. OPHL § 50, stk. 2, er fleksibel og således omfatter enhver udnyttelsesform, giver indskrænkningen adgang til at klarere **alle tre tilfælde** af kopiering gennem en generel licensaftale med en rettighedsorganisation. Indskrænkningen dækker dog ikke kopieringen af et værk i forbindelse med træningen, hvis rettighedshaveren til værket har nedlagt forbud mod denne brug, jf. OPHL § 50, stk. 2, 2. pkt.

\(301\) Schønning (2021), s. 550

\(302\) L 58 2007-08, afsnit 11

\(303\) Ibid., afsnit 11
6 Er de nugældende indskrænkninger hensigtsmæssige?

I afsnit 4 blev det først fastslået, hvornår kopieringen af beskyttede værker, der finder sted ifm. træningen af en *ML*-algoritme, som skal implementeres i et *generative AI*-system, er omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1, og derfor som udgangspunkt kræver tilladelse fra indehaverne af ophavsretten. Derefter blev det klarlagt i afsnit 5, i hvilket omfang kopieringen af træningsdata er dækket af de lovbestemte indskrænkninger i ophavsretten, og derfor ikke kræver rettighedshaverernes tilladelse.

I lyset af delkonklusionerne fra afsnit 4 og 5 vil det i dette afsnit blive diskuteret, om de nugældende indskrænkninger er hensigtsmæssige, når man betragter dem fra en række forskellige perspektiver. Som fastslået i afgrænsningen i afsnit 1.4, er det kun hensigtsmæssigheden af de tre lovbestemte indskrænkninger, der blev behandlet i afsnit 5, der vil blive diskuteret.

6.1 Indskrænkningernes hensigtsmæssighed fra hhv. et bruger- og et rettighedshaverperspektiv

En diskussion af de nugældende indskrænkningers hensigtsmæssighed fra hhv. et bruger- og et rettighedshaverperspektiv kan passende indledes med en henvisning til det forhold, at det i skrivende stund er helt og aldeles *uafklaret* i praksis, i hvilket omfang de tre indskrænkninger dækker kopieringen af træningsdata, der finder sted ifm. træningen af en *ML*-algoritme. Dette forhold synes at give anledning til en række konkrete udfordringer for en udvikler af en algoritme, hvis man betragter indskrænkningerne fra et brugerperspektiv.

Som tidligere fastslået har anvendeligheden af OPHL § 11 a, stk. 1, og DSM-direktivets art. 4(1) hverken været prøvet af EUD eller af de nationale domstole i de enkelte af EU’s medlemsstater. Om end begge indskrænkninger som konkluderet i afsnit 5.2.5 og 5.3.3 synes at omfatte visse af de kopier, der bliver fremstillet ifm. træningen af en *ML*-algoritme, er denne konklusion dog udelukkende funderet i en række afgørelser fra EUD, som ikke specifikt omhandler *ML*, samt i den juridiske litteratur. Som følge heraf må det antages, at en udvikler af en *ML*-algoritme ikke på nuværende tidspunkt kan handle i blind tillid til de to indskrænkninger, da der er en vis risiko for, at kopieringen af værkerne indebærer en ophavsretskrænkelse. Da indskrænkningernes anvendelighed ikke er afklaret i praksis, udgør de to indskrænkninger derfor også et i praksis *usikkert* grundlag for kopiering af værker ifm. træningen uden tilladelse fra rettighedshaverne.

På samme vis findes der som tidligere anført heller ikke nogen licensaftaler efter OPHL § 50, stk. 2, om kopiering af beskyttede værker ifm. træningen, som er godkendt af Kulturministeriet i overensstemmelse med OPHL § 50, stk. 4, som vil kunne afklare, i hvilket omfang kopieringen af
værker ifm. træningen vil kunne dækkes af en sådan aftalelicens i praksis.\textsuperscript{304} Det er desuden uafklaret, hvorvidt en udvikler af en algoritme rent faktisk vil kunne nå til enighed om en aftale om kopieringen med en repræsentativ rettighedsorganisation i henhold til OPHL § 50, stk. 2, eller, hvis det lykkes at nå til enighed med organisationen, om Kulturministeriet rent faktisk vil godkende rettighedsorganisationen, jf. § 50, stk. 4.

I det omfang OPHL § 50, stk. 2, ikke kan påberåbes, vil en udvikler af en \textit{ML}-algoritme således være forpligtet til at klarere kopieringen af værkerne på eneretsbasis og således gennem individuelle aftaler med rettighedsavere (enten samtlige rettighedsavere, hvis udvikleren ikke kan nå til enighed om en aftale, eller alle ikke-representerede rettighedsavere, hvis aftalen ikke kan tillægges aftalelicensvirkning). Dette er selvsagt en omfattende opgave for en udvikler, henset til at træningen i praksis bliver gennemført ved brug af tusindvis (og undertiden millionvis) af værker.

Det kan dog afhængigt af indskrænkningen næppe indebære den samme krænkelserisiko som de to øvrige indskrænkninger, da OPHL § 50, stk. 2, ikke er en fribrugsregel. Da OPHL § 50, stk. 2, forudsetter, at brugeren kan nå til enighed om en aftale og om vilkårene for denne med en rettighedsorganisation, må det antages, at der ikke vil være nogen risiko for at krænke ophavsretten, så længe brugen af værkerne sker i overensstemmelse med vilkårene for aftalen.

Foruden at indskrænkningernes anvendelsesområde på nuværende tidspunkt er uafklaret i praksis, skal det derudover i lyset af delkonklusionerne fra afsnit 5 bemærkes, at det kun er OPHL § 50, stk. 2, der synes at dække alle tre tilfælde af kopiering af værker ifm. træningen af en \textit{ML}-algoritme.

Som konkluderet i afsnit 5.2.5 er det kun visse af de kopier, der bliver fremstillet ifm. lagringen af træningsdata i det system, som træningen skal gennemføres i, samt ifm. algoritmens generering af output, der synes at være omfattet af OPHL § 11 a, stk. 1. Anvendelsesområdet af DSM-direktivets art. 4 er endnu snævrere, da denne indskrænkning som konkluderet i afsnit 5.3.3 kun synes at finde anvendelse på visse af de kopier, der bliver fremstillet ifm. lagringen af træningsdata.

På den baggrund kan der ligeledes påpeges den udfordring for en udvikler af en algoritme, hvis man betragter indskrænkningerne fra et brugerperspektiv, at det efter de nugældende indskrænkninger slet ikke er muligt at bruge værker som træningsdata helt uden tilladelse. Da den kopiering af værker, der finder sted ifm. algoritmens analyse, ikke er dækket af fribrugsreglerne, kræver denne kopiering

\textsuperscript{304} Kulturministeriets oversigt over aftalelicensgodkendelser (se under fanen ”Godkendelser efter den generelle aftalelicens”) – \url{https://kum.dk/kulturomraader/aftalelicensgodkendelser} (lokaliseret d. 9. maj 2023 kl. 16:47)
derfor altid tilladelse i det omfang, at dét, som algoritmen kopierer i den forbindelse, er udtryk for ophavsmandens egen intellektuelle frembringelse, jf. Infopaq I.\textsuperscript{305}

Endelig kan der påpeges den udfordring, hvis man anskuer indskrænkningerne fra et brugerperspektiv, at der er en række forhold, der synes at tale for, at indskrænkningernes respektive anvendelsesområder er væsentligt snævrere i praksis end i teorien, hvad angår kopiering af værker ifm. træningen af en ML-algoritme.

Som fastslået i afsnit 5.2 er et eksemplar kun dækket af OPHL § 11 a, stk. 1, hvis de fire kumulative betingelser, som disse er fortolket af EUD, er opfyldt. Foruden det forhold, at EUD ifølge Axhamn i høj grad har fortolket undtagelsen til fordel for rettighedshaverne\textsuperscript{306}, kan der som anført af Margoni, Ducato et al. og Chiou påpeges den udfordring, at betingelserne er formuleret så snævert og uklart, at de formentlig ekskluderer de fleste tilfælde af kopiering af værker ifm. træningen af en ML-algoritme i praksis.\textsuperscript{307} Dette skyldes bl.a. det forhold, at der ikke findes én universel teknik men mange forskellige teknikker, som kan anvendes til træningen af en ML-algoritme.\textsuperscript{308} Derfor er der ikke tale om en indskrænkning, som generelt dækker kopieringen af værker ifm. træningen, men en indskrænkning, som kun dækker kopieringen, hvis det på baggrund af en særdeles grundig vurdering i det enkelte tilfælde kan fastslås, at de kumulative betingelser er opfyldt.\textsuperscript{309} Af samme årsag er Axhamn i modsætning til de øvrige juridiske forfattere af den opfattelse, at kopieringen normalt slet ikke vil være dækket af indskrænkningen.\textsuperscript{310}

Som anført af Schönberger skal det bemærkes, at EU-lovgiver slet ikke havde ML i tankerne, da InfoSoc-direktivet blev vedtaget i 2001.\textsuperscript{311} Dette kunne være en forklaring på, hvorfor betingelserne umiddelbart synes særdeles snævre og uklare, hvad angår de kopier, der bliver fremstillet ifm. træningen af en ML-algoritme. Det bemærkes dog, at det som anført i afsnit 5.3 følger af DSM-direktivets 9. præambelbetragtning, at der ifølge EU-lovgiver vil være tilfælde af TDM, som er omfattet af InfoSoc-direktivets art. 5(1). Da den brede definition af TDM iht. DSM-direktivets art. 2, nr. 2, som konkluderet i afsnit 5.3.3 omfatter visse af de kopier, der bliver fremstillet ifm. lagringen af træningsdata, kan man argumentere for, at EU-lovgiver har bekræftet, at visse af de kopier

\textsuperscript{305} Infopaq I, pr. 51
\textsuperscript{306} Axhamn (2010), s. 353
\textsuperscript{308} Chiou (2019), s. 410
\textsuperscript{309} Margoni (2018), s. 20. Chiou (2019) s. 406
\textsuperscript{310} Axhamn (2021), s. 55
\textsuperscript{311} Schönberger (2018), s. 407
undertiden er omfattet af InfoSoc-direktivets art. 5(1). Dette taler således imod Axhamns synpunkt. Imidlertid løser det ikke de konkrete udfordringer, der som anført af de juridiske forfattere, er forbundet med opfyldelsen af betingelserne. På baggrund af de juridiske forfatteres synspunkter kan det derfor anføres, at selvom OPHL § 11 a, stk. 1, som konkluderet i afsnit 5.2.5 synes at omfatte visse af de kopier, der bliver fremstillet ifm. træningen af en ML-algoritme, udgør indskrænkningen et for en bruger snævert grundlag for kopiering af værker ifm. træningen af en ML-algoritme uden tilladelse fra rettighedshaverne i praksis.

Hvad angår DSM-direktivets art. 4, synes de udfordringer, der ifølge de juridiske forfattere kan påpeges i relation til denne indskrænkning, at bero på rettighedshavers ret til at forbeholde sig sine rettigheder i relation til TDM-formål, jf. art. 4(3). Denne ret har ifølge en lang række juridiske forfattere potentialet til at indsnævre den effektive anvendelse af indskrænkningen i praksis, da det forhold, at et værk er lovligt tilgængeligt, ikke er ensbetydende med, at det må reproduceres, da det kan være omfattet af et forbehold. For at afklare, hvorvidt et værk er omfattet af et forbehold, er en udvikler af en ML-algoritme nødt til at foretage en forudgående undersøgelse af samtlige værker, hvilket ifølge Chiou både er tidskrævende og omkostningstungt.


I lyset af de juridiske forfatteres synspunkter kan det anføres, at selvom DSM-direktivets art. 4 som konkluderet i afsnit 5.3.3 synes at finde anvendelse på visse af de kopier, der bliver fremstillet ifm. lagringen af træningsdata i det system, som træningen skal gennemføres i, synes rettighedshavers forbeholdser relativt iht. DSM-direktivets art. 4(3) at indebære, at indskrænkningen i art. 4(1) også

---

313 Chiou (2019), s. 409
314 Rosati (2019), s. 23
315 Gerrish (2019), s. 66
indeholder en i praksis snæver adgang til kopiering af værker ifm. træningen af en algoritme uden tilladelse.


I forlængelse af ovenstående skal det desuden bemærkes, at principperne iht. OPHL § 11, der som fastslået i afsnit 3.2 skal iagttages af en bruger, når denne anvender en indskrænkning iht. OPHL kap. 2, indsnævrer adgangen til at anvende de tre indskrænkninger ifm. træningen af en *ML*-algoritme yderligere.

For det første skal OPHL § 11, stk. 2, iagttages, hvilket indebærer, at værket ikke må ændres i videre udstræknelse, end den tilladte brug kræver. Som anført i afsnit 3.2 tillader bestemmelsen kun mindre, praktisk nødvendige ændringer. På den baggrund kan man fx argumentere for, at rettighedshavers tilladelse er påkrævet, selvom den kopiering af værket, som finder sted ifm. algoritmens generering af output, er omfattet af en indskrænkning, hvis algoritmens output fører til en ændring af værket. Som fastslået i afsnit 5.2.4 er ændringer dog *helt udelukket* ved anvendelsen af OPHL § 11 a, stk. 1, jf. *Infopaq II*.316

For det andet skal kopieringen ske på baggrund af et lovligt forlæg, jf. OPHL § 11, stk. 3, hvilket selvsagt sætter en begrundelse på udvalget af de eksemplarer og gengivelser af værket, som lovligt kan anvendes til træningen. Desuden kan det anføres, at en udvikler af en *ML*-algoritme skal have fuldstændigt overblik over lovligheden af samtlige værker, der er anvendt som træningsdata. Dette kan påpeges som en udfordring, hvad angår det forhold, at en *ML*-algoritme som anført i afsnit 2.1 ofte bliver trænet på baggrund af værker, som er *scraped* fra internettet, hvorfor en udvikler af en algoritme således må foretage en vis retlig due diligence mhp. at afklare samtlige værkers lovlighed.317

---
316 *Infopaq II*, pr. 53
317 Chiou (2019), s. 410
På baggrund af de mange udfordringer, der er forbundet med anvendelsen af indskrænkningerne i praksis, kan man argumentere for, at de nugældende indskrænkninger ikke er hensigtsmæssige, hvis man betragter dem fra et brugerperspektiv.

Derimod kan man argumentere for, at de nugældende indskrænkninger er hensigtsmæssige, hvis man betragter dem fra et rettighedshaverperspektiv. Selvom indskrænkningernes anvendelighed i relation til kopiering af værker ifm. træningen af en ML-algoritme er uafklaret i praksis, må rettighedshavers forbeholdser, jf. DSM-direktivets art. 4(3), og adgang til at nedlægge forbud mod værkets anvendelse, jf. OPHL § 50, stk. 2, 2. pkt., indebære, at det alene er i tilfælde, hvor kopieringen opfylder de usikre og snævre betingelser i OPHL § 11 a, stk. 1, at rettighedshaver reelt må tåle, at værket bliver kopieret ifm. træningen – dog ikke i større omfang, end at OPHL § 11 a, stk. 1, som anført ovenfor aldrig omfatter kopieringen ifm. algoritmens analyse, hvorfor denne kopiering således altid kræver tilladelse.

6.2 Indskrænkningernes hensigtsmæssighed fra et formåls perspektiv

Betrætter man indskrænkningerne i lyset af de formål, som følger af lovmotiverne, og således lov-givers hensigt med indskrænkningerne, kan det for det første anføres om OPHL § 50, stk. 2, at denne indskrænkning som anført i afsnit 5.4 har til formål at gøre det muligt for en bruger at påberåbe sig aftalelicens, hvor det ikke er "praktisk muligt" at indgå individuelle aftaler med indehaverne af ophavsretten.\textsuperscript{318} Indskrænkningen er derfor en alternativ løsning til sådanne aftaler.\textsuperscript{319} Om end det er uafklaret i praksis, hvorvidt OPHL § 50, stk. 2, dækker kopieringen af værker ifm. træningen, kan man dog argumentere for, at reglen er hensigtsmæssig i lyset af formålet, som dette kommer til udtryk i lovmotiverne.

På trods af at indskrænkningen blev foreslået for at imødekomme de ønsker om nye aftalelicenser, som opstod i takt med den teknologiske udvikling\textsuperscript{320}, er OPHL § 50, stk. 2, ikke en garanti for, at en brug af værker i digital form kan klareres gennem en generel licensaftale, da indskrænkningen iht. lovmotiverne kun finder anvendelse, hvor en bruger og en repræsentativ rettighedsorganisation kan nå til enighed om brugen, herunder om vilkårene og vederlaget.\textsuperscript{321} Lovgiver har således ladet det være op til brugerne og rettighedsorganisationerne at finde frem til en løsning.

\textsuperscript{318} L 58 2007-08, afsnit 3.3.1
\textsuperscript{319} Ibid., afsnit 3.3.1
\textsuperscript{320} Ibid., afsnit 3.2.
\textsuperscript{321} L 58 2007-08, afsnit 3.3.1. Kyst (2009), s. 46.
Hvad angår undtagelsen for midlertidig eksemplarfremstilling, følger det af InfoSoc-direktivets 4. præambelbetrægtnings, at direktivet har til formål at sikre et højt beskyttelsesniveau for intellektuelle ejendomsrettigheder.\textsuperscript{322} Dog følger det af InfoSoc-direktivets 5. præambelbetrægtnings, at direktivet samtidig har til formål at tilpasse den ophavsretlige regulering, så denne tager højde for den teknologiske udvikling.\textsuperscript{323} I lyset af disse formål kan man argumentere for, at indskrænkningen ikke er hensigtsmæssig, da den snævre og usikre adgang til kopieringen af beskyttede værker uden tilladelse fra indehaverne af ophavsretten ifm. træningen af en \textit{ML}-algoritme netop ikke synes at tage højde for den teknologiske udvikling, men derimod udelukkende rettighedshaverernes interesser.

På den anden side skal det dog påpeges, som ligeledes anført af Schönberger, at lovgiver ikke havde \textit{ML} i tankerne, da indskrænkningen blev indført i 2001.\textsuperscript{324} Selvom \textit{Infopaq I og Infopaq II} bekræfter, at indskrænkningen også finder anvendelse på andre tekniske processer end dem, som EU-lovgiver formentlig havde i tankerne (som tidligere fastslået blev indskrænkningen bl.a. indført for at muliggøre surfing på internettet\textsuperscript{325}), og selvom DSM-direktivets 9. præambelbetrægtnings desuden bekræfter, at visse kopier af værker, der bliver fremstillet ifm. træningen af en \textit{ML}-algoritme, er omfattet af OPHL § 11 a, stk. 1, kan man argumentere for, at træningen af en \textit{ML}-algoritme er langt væk fra det anvendelsesområde, som EU-lovgiver oprindeligt havde påtænkt.\textsuperscript{326} Således synes EU-lovgiver ikke at have lovet mere, end hvad EU-lovgiver har kunnet holde, da indskrænkningen blev vedtaget, hvorfor mest synes at tale for, at indskrænkningen er hensigtsmæssig, hvis man betragter den fra et formålsperspektiv.

Hvad endelig angår DSM-direktivets art. 4(1), følger det af direktivets 3. præambelbetrægtnings, at direktivet har til formål at fremtidssikre reglerne om ophavsret i lyset af den hurtige teknologiske udvikling.\textsuperscript{327} Herunder har indskrænkningen for \textit{TDM} i art. 4 til formål at skabe retssikkerhed og fremme innovation i den private sektor, jf. direktivets 18. præambelbetrægtnings.

Som følge af rettighedshavers forbeholdsret, jf. art. 4(3), der som fastslået ovenfor indsnævrer adgangen til at anvende indskrænkningen i praksis, har en række juridiske forfattere imidlertid sat spørgsmålstegn ved, om EU-lovgiver rent faktisk har realiseret formålet med direktivet ved at vedtage art. 4 i indskrænkningens nuværende udformning. \textit{Lizarralde} anfører således, at formålet med art. 4

\begin{footnotesize}
\begin{itemize}
\item[322] InfoSoc-direktivets 4. præambelbetrægtnings
\item[323] InfoSoc-direktivets 5. præambelbetrægtnings
\item[324] Schönberger (2018), s. 407
\item[325] InfoSoc-direktivets 33. præambelbetrægtnings
\item[326] Chiou (2019), s. 497
\item[327] DSM-direktivets 3. præambelbetrægtnings
\end{itemize}
\end{footnotesize}
efters hendes mening er mislykket, da forbeholdsretten i art. 4(3) begrænser den effektive anvendelse af indskrænkningen.\textsuperscript{328} I tråd hermed anfører Rosati, at art. 4 snarere udgør en forhindring for reproduktion af beskyttede værker end en adgang til reproduktion.\textsuperscript{329} På den anden side anfører Gerrish et al. dog, at det er for tidligt at konkludere, hvorvidt art. 4 fremmer eller hindrer innovation, da indskrænkningen fortsat er så ny, at vi endnu ikke har set dens endelige effekt.\textsuperscript{330}

Uanset hvad konklusionen er eller bliver, skal det under alle omstændigheder bemærkes, at DSM-direktivet ifølge Chiou synes at være blevet overhalet af den teknologiske udvikling, da verden er gået ind i en ny digital æra siden vedtagelsen af direktivet.\textsuperscript{331} Chiou’s synspunkt understøttedes i høj grad af den omstændighed, at ML slet ikke er behandlet i den meddelelse fra EU-Kommissionen\textsuperscript{332} fra 2015, der i et vist omfang danner baggrund for reglerne i DSM-direktivet\textsuperscript{333}. Selvom indskrænkningen som følge af rettighedshavers forbeholdsret efter art. 4(3) udgør et i praksis snævert og usikkert grundlag for fremstillingen af kopier af værker ifm. træningen af en ML-algoritme, kan man dog i lyset af lovgivers formål argumentere for, at indskrænkningen er hensigtsmæssig.

Ligesom det gør sig gældende i relation til OPHL § 11 a, stk. 1, synes der i lyset af Chiou’s betragtninger at være tale om en indskrænkning, der blev vedtaget, uden at EU-lovgiver havde ML for øje, hvorfor EU-lovgiver således ikke synes at have forudsat indskrænkningens anvendelighed for kopieringen af værker ifm. træningen af en algoritme (der som fastslået i afsnit 5.3.1.1 er et andet teknisk fænomen end TDM). Hertil kommer det forhold, at DSM-direktivets regler ligesom InfoSoc-direktivet har til formål at sikre et højt niveau af beskyttelse for rettighedshavere, jf. direktivets 2. præambelbetraktning.\textsuperscript{334}

Sammenfattende kan man derfor argumentere for, at de nugældende indskrænkninger er hensigtsmæssige, hvis man betragter dem fra et formålsperspektiv og således i lyset lovgivers formål med indskrænkningerne. Selvom OPHL § 11 a, stk. 1, og DSM-direktivets art. 4 som anført i afsnit 6.2 indeholder en i praksis usikker, snæver og uafklaret adgang til at kopiere værker ifm. træningen uden tilladelse, har EU-lovgiver netop forudsat et højt beskyttelsesniveau for rettighedshavere. Ligeledes har ML ikke været tænkt ind i indskrænkningerne, da disse blev vedtaget. På samme vis

\textsuperscript{328} Lizarralde (2020)
\textsuperscript{329} Rosati (2020), s. 23
\textsuperscript{330} Gerrish et al. (2019), s. 67
\textsuperscript{331} Chiou (2019), s. 411
\textsuperscript{332} EU-Kommissionens meddelelse om en tidssvarende ramme for ophavsret (2015)
\textsuperscript{333} DSM-direktivets 3. præambelbetraktning
\textsuperscript{334} DSM-direktivets 2. præambelbetraktning

6.3 Indskrænkningernes hensigtsmæssighed fra et globalt konkurrenceperspektiv

Det kan endelig diskuteres, om de nugældende indskrænkninger (herunder særligt de EU-harmoniserede indskrænkninger iht. OPHL § 11 a, stk. 1, og DSM-direktivets art. 4) er hensigtsmæssige, hvis man betragter dem fra et globalt konkurrenceperspektiv.

Man kan argumentere for, at det forhold, at indskrænkningerne som fastslået ovenfor indeholder en i praksis både uafklaret og nødvendig adgang til kopiering af beskyttede værker uden tilladelse fra indehaverne af ophavsretten, vil kunne svække EU’s (og herunder Danmarks) position i den globale konkurrence på markedet for AI. En sådan opfattelse ses hos både Chiou og Geiger et al.335

Dette argument understøttes bl.a. af det forhold, som anført af både Schönberger, Lizarralde og Gerrish et al., at små og mellemstore virksomheder (særligt når disse befinder sig i start-up-fason) ofte kun råder over ganske begrænsede midler, hvilket indebærer, at de hverken har råd til at klarerere brugen af tusindvis af værker gennem licensaftaler med rettighedshaverne, eller, hvis de sætter deres lid til de usikre indskrænkninger, råd til at risikere at krænke rettighederne til de værker, de har anvendt som træningsdata, med et vederlagskrav fra rettighedshaverne til følge.336 På den baggrund kan man argumentere for, at uklarheden om indskrænkningernes anvendelse i praksis og det forhold, at kopieringen ifm. algoritmens analyse altid kræver tilladelse (og en vis betaling af vederlag for udnyttelsen) formentlig vil afholde visse små og mellemstore virksomheder fra at udvikle generative AI-systemer, hvilket i så fald vil begrænse udviklingen af sådanne systemer i EU (herunder i Danmark). Det er således kun virksomheder, der rent faktisk har råd til at klarerere rettigheder, der vil kunne udvikle generative AI-systemer i EU (herunder i Danmark).

Perspektiverer man til de regler om ophavsret, der gælder i lande uden for EU, fx Japan og USA, synes de indskrænkninger i ophavsretten, der gælder i disse lande, at være væsentligt mere gunstige, hvad angår adgangen til at bruge værker til træningen uden tilladelse hertil. Dette understøtter ligeledes argumentet om, at de nugældende indskrænkninger vil kunne svække EU’s (herunder Danmarks) position i den globale konkurrence på markedet for AI.

335 Chiou (2019), s. 411, Geiger et al. II (2018), s. 110
Som anført af Gerrish et al. og Lizarralde blev der i 2018 indført nye indskrænkninger i den japanske ophavsretslov, som gør det muligt for enhver bruger af beskyttede værker at gennemføre en analyse af disse værker på en hvilken som helst måde til ML-formål uden tilladelse fra indehaverne af ophavsretten samt at fremstille tilfældige kopier af værkerne i den forbindelse.337 Disse indskrænkninger havde til formål at sikre, at ophavsret aldrig kan være en udfordring for udviklingen af AI i Japan.338

På samme vis henviser Rosati, Margoni, Quintais og Schönberger til, at den amerikanske "fair use"-doktrin339 formentlig indebærer, at adgangen til at bruge beskyttede værker ifm. træningen af en ML-algoritme er mere sikker i USA.340

"Fair use"-doktrinen gør det muligt at udnytte et værk uden tilladelse, hvis det kan fastslås på baggrund af en konkret vurdering i lyset af en række momenter341, at udnyttelsen kan betragtes som "fair".342 Afgørelsen i den amerikanske sag Authors Guild v Google, Inc.343 synes ifølge Rosati og Margoni at bekærte, at "fair use"-doktrinen konkret fører til, at en digital analyse af litterære værker mhp. at udlede specifik information deri, er en aktivitet, som ikke kræver tilladelse fra indehaverne af ophavsretten til disse værker.344 Således har de amerikanske domstole bekræftet – i modsætning til EUD – at en sådan analyse af værker kan gennemføres uden tilladelse.

Som fastslået i afsnit 2.2.2 verserer der dog på nuværende tidspunkt en sag ved de amerikanske domstole, der angår spørgsmålet om, hvorvidt Stability AI har krænket rettighederne til de værker, der er blevet anvendt til træningen af Stable Diffusion-algoritmen. Stability AI har i den forbindelse bl.a. påberåbt sig "fair use"-doktrinen for den pågældende udnyttelse.345 Hvis sagen bliver afgjort til fordel for de rettighedshavere, som har anlagt sagen, vil afgørelsen selvsagt potentielt kunne indsnævre adgangen til at udvikle generative AI-systemer i USA.

339 § 107 i den amerikanske ophavsretslov
341 Herunder formålet med udnyttelsen og udnyttelsens påvirkning af værkets værdi – se den amerikanske ophavsretslovs § 107
342 Rosati (2019), s. 15
344 Margoni (2018), s. 21. Rosati (2019), s. 16
Uanset hvad resultatet af *Stable Diffusion*-sagen bliver, kan man dog argumentere for, at de indskrænkninger, der på nuværende tidspunkt gælder i EU (herunder i Danmark) ikke synes hensigtsmæssige, hvis man betragter dem fra et globalt konkurrenceperspektiv, da adgangen til at gennemføre treningen af en *ML*-algoritme uden tilladelse fra indehaverne af ophavsretten er snævrere i EU (herunder i Danmark) end i lande som Japan og USA. Dette medfører, at det kun er virksomheder, der har råd til at klarere rettighederne til de værker, der er anvendt som træningsdata, der vil kunne udvikle *generative AI*-systemer.

På den anden side bemærkes det dog, at den ophavsretlige regulering ifølge *Gerrish et al.* ikke i sig selv er afgørende for EU's (og Danmarks) position i den globale konkurrence på markedet for *AI*. De udfordringer, som de nugældende indskrænkninger medfører for konkurrencen, kan bl.a. løses ved at sikre virksomheder i EU en adgang til økonomisk støtte af deres *AI*-projekter, som bl.a. kan anvendes til klarering af rettigheder.\(^{346}\) Det skal i den forbindelse bemærkes, at EU allerede har investeret massivt i *AI*-udvikling.\(^{347}\)

Det kan dog anføres, at EU-lovgiver synes tvetydig, når det kommer til tilkendegivelser af, hvilken rolle EU bør spille i den globale konkurrence på markedet for *AI*. EU-Kommissionen har således på den ene side tilkendegivet, at EU skal være en "*global leader in innovation in the data economy and its application*"\(^{348}\), men på den anden side tilkendegivet, at det ikke er et spørgsmål om "*winning or losing a race but of finding the way of embracing the opportunities offered by AI in a way that is human-centered, ethical, secure, and true to our core values*"\(^{349}\). Det bemærkes dog, at begge ønsker er afspejlet i EU-Kommissionens forslag til AI-forordningen, hvilket taler for, at EU-Kommissionen har fået en vis interesse i at facilitere udvikling af *AI*-systemer i EU.\(^{350}\)

Disse omstændigheder synes dog ikke at ændre ved, at de nugældende indskrænkninger ikke synes hensigtsmæssige fra et globalt konkurrenceperspektiv. Selvom udviklingen af *generative AI* vil kunne faciliteres gennem økonomisk støtte og lignende tiltag, ændrer dette ikke på, at indskrænkningerne isoleret set giver en i praksis uafklaret og usikker adgang til at kopiere værker ifm. træningen af en

---

346 *Gerrish et al.* (2019), s. 67

347 Alene i perioden 2014-2017 investerede EU 1,1 milliarder Euro i AI-relateret forskning og innovation – se EU-Kommissionens meddelelse om AI (2018), s. 5

348 EU-Kommissionens White Paper om AI (2020), s. 2f.

349 EU-Kommissionens rapport om AI (2018), s. 17

350 Det følger af forordningen, at den både har til formål at fremme udviklingen af et indre marked for lovlige, sikre og pålidelige *AI*-anvendelser og til formål at styrke Europas konkurrenceevne inden for kunstig intelligens – se EU-Kommissionens Forslag til AI-forordningen, s. 3 og 7
ML-algoritme uden tilladelse, og at indskrænkningerne således er mindre gunstige end de indskrænkninger, der gælder i lande uden for EU som fx Japan og USA.

6.4 Nye regler eller en afklaring af de nugældende indskrænkningers anvendelighed?

I lyset af ovenstående diskussion af de nugældende indskrænkningers hensigtsmæssighed kan det afslutningsvist diskuteres, hvordan de udfordringer, som er forbundet med indskrænkningerne vil kunne imødekommes.

Fra et brugerperspektiv kan man på den ene side argumentere for, at en ønskelig løsning kunne bestå i vedtagelsen af en præcist formuleret indskrænkning, som specifikt gælder for kopiering af værker ifm. træningen af en ML-algoritme. En klart og præcist affattet indskrænkning for brug af værker uden tilladelse fra indehaverne af ophavsretten reducerer selvsagt de risici, som er forbundet med at handle i tillid til indskrænkningen.\textsuperscript{351} Dette vil desuden facilitere adgangen til udvikling af "generative AI" for virksomheder, der ikke har råd til at klarere rettigheder.

På den anden side kan dog påpeges den udfordring, at man formentlig ikke kan forvente at se en sådan indskrænkning i den nærmeste fremtid.\textsuperscript{352} Som anført af Gerrish et al. gik der næsten tyve år fra den første ophavsretsform om digital brug af værker, som blev gennemført af EU-lovgiver med InfoSoc-direktivet i 2001, før endnu en reform blev gennemført med DSM-direktivet i 2019.\textsuperscript{353}

Derudover er det eneste initiativ, der i skrivende stund er taget på EU-plan, hvad angår reguleringen af brugen af beskyttede værker ifm. træningen af en ML-algoritme, EU-Parlamentets forslag om tilføjelsen af en regel til forslaget til AI-forordningen, der som beskrevet i afsnit 3.3 pålægger udviklere af sådanne algoritmer at offentliggøre en fortegnelse over de værker, som de har anvendt som træningsdata. Selvom denne regel vil kunne imødekomme den i praksis aktuelle udfordring, at det ofte er vanskeligt for rettighedshavere at opdage, at deres værker er blevet brugt til træningen\textsuperscript{354}, er der dog ikke tale om en regel, som faciliterer generative AI-udvikleres adgang til at anvende værker til træningen af en ML-algoritme uden tilladelse. Den foreslåede regel afspejler tværtimod det høje beskyttelsesniveau af rettighedshaverne, der hersker på EU-plan, og som bl.a. kommer til udtryk i EU-direktiverne om ophavsret.

\textsuperscript{351} Hilty et al. (2017), s. 1
\textsuperscript{352} Gerrish et al. (2019), s. 66
\textsuperscript{353} Ibid., s. 66
\textsuperscript{354} Study on the legal framework of TDM (2014), s. 87
I lyset af at der allerede findes tre indskrænkninger, der som konkluderet i afsnit 5 synes at omfatte visse af de kopier, som bliver fremstillet ifm. træningen af en ML-algoritme, må løsningen i første omgang være at af vente, at disse indskrænkningers anvendelsesområde bliver endeligt afklaret i praksis. En sådan afklaring vil gøre det muligt at vurdere den reelle effekt af de nugældende indskrænkninger og herunder, om der er behov for at tilpasse reglerne til den nye digitale tidsalder, verden tydeligvis er gået ind i siden DSM-direktivet blev vedtaget. Som følge af den hastighed, hvormed udviklingen af generative AI som beskrevet i afsnit 1.1 sker for tiden, må det antages, at det blot er et spørgsmål om tid, før den nødvendige afklaring kommer.

6.5 Delkonklusion


De nugældende indskrænkninger er ligeledes hensigtsmæssige, hvis man betragter dem fra et formålsperspektiv og således i lyset af lovgivers hensigt med indskrænkningerne, da ML ikke var tænkt ind i indskrænkningerne, da disse blev vedtaget.

Endelig er indskrænkningerne ikke hensigtsmæssige, når man betragter dem fra et globalt konkurrencesynspunkt, da de indskrænkninger, der gælder i lande uden for EU som fx Japan og USA, er væsentligt mere gunstige for adgangen til at bruge værker til træningen af en ML-algoritme uden tilladelse. Dette vil kunne svække EU's (herunder Danmarks) position i den globale konkurrence på markedet for AI.

7 Konklusion

Generative AI har på særdeles kort tid revolutioneret den kreative skabelsesproces på verdensplan. Det er nu blevet muligt at overlade skabelsen af kreativt indhold såsom billeder, musik og tekst til en ML-algoritme, som kan generere et sådant indhold på et splitsekund. Dette teknologiske fænomen har
Imidlertid den slagside, at algoritmen alene er i stand til at generere indhold, fordi den er blevet trænet hertil ved at analysere træningsdata, der ofte udgøres af tusindvis (og undertiden millionvis) af ophavsretligt beskyttede værker, der løbende bliver kopieret ifm. træningen. Der er således tale om en brug, der potentielt er omfattet af rettighedshavernes eneret til eksemplarfremstilling, jf. OPHL § 2, stk. 1.

Formålet med dette speciale har været at undersøge, hvornår kopieringen af træningsdata, der finder sted ifm. træningen af en ML-algoritme mhp. implementering i et generativt AI-system, er omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1, når denne data består af ophavsretligt beskyttede værker, i hvilket omfang kopieringen er dækket af de lovbestemte indskrænkninger i ophavsretten, og om de nugældende indskrænkninger er hensigtsmæssige.

På baggrund af specialets analyser kan det konkluderes, at kopieringen af træningsdata, når denne består af værker, er omfattet af eneretten til eksemplarfremstilling, jf. OPHL § 2, stk. 1, og derfor kræver tilladelse fra rettighedshaverne i det omfang, at der er tale om kopiering af værkerne i deres oprindelige, originale skikkelse og i deres helhed. Er der derimod tale om kopiering af enkelte elementer af værkerne i deres oprindelige skikkelse, er kopieringen kun omfattet af § 2, stk. 1, hvis de pågældende elementer i sig selv opfylder originalitetskravet, jf. Infopaq I. Derudover er kopieringen af værker ifm. træningen omfattet af OPHL § 2, stk. 1, hvis det på baggrund af en konkret vurdering i lyset af de principper, der gælder for denne vurdering, kan fastslås, at der er tale om fremstilling af en kopi af værket i ændret skikkelse.


Kopieringen er dækket af den EU-harmoniserede indskrænkning i OPHL § 11 a, stk. 1, i det omfang, at de fire kumulative betingelser, der følger af bestemmelsen, er opfyldt. Som følge af EUD's fortolkning af betingelserne er det dog kun de kopier, der bliver fremstillet ifm. lagringen af værkerne samt ifm. algoritmens generering af et output, der er dækket af indskrænkningen.

DSM-direktivets art. 4(1) har et snævrere anvendelsesområde, og indskrænkningen dækker alene den kopiering, der finder sted i forbindelse med lagringen af værkerne. Kopieringen er dog kun omfattet
af indskrænkningen, såfremt algoritmens analyse kan henføres under definitionen af TDM iht. direktivets art. 2, nr. 2, og at rettighedshaver ikke har forbeholdt sig sine rettigheder, jf. art. 4(3).

Endelig dækker OPHL § 50, stk. 2, alle tre tilfælde af kopiering i det omfang, at udvikleren af ML-algoritmen kan nå til enighed med en repræsentativ rettighedsorganisation om en aftale, og forudsat, at rettighedsorganisationen bliver godkendt af Kulturministeriet i medfør af § 50, stk. 4.

Hvad angår spørgsmålet om hensigtsmæssighed, kan det konkluderes, at de nugældende indskrænkninger ikke er hensigtsmæssige fra et brugerperspektiv. Dette skyldes, at indskrænkningernes anvendelsesområde synes at være sævertr, hvad angår kopiering af træningsdata, og at det ikke er afklaret i praksis, i hvilket omfang indskrænkningerne dækker kopieringen af træningsdata ifm. træningen. Derudover kræver kopieringen, der finder sted ifm. algoritmens analyse, altid tilladelse. Disse forhold er omvendt hensigtsmæssige fra et rettighedshaverperspektiv, idet rettighedshaver sjældent må tåle, at værkerne bliver kopieret.

Indskrænkningerne er ligeledes hensigtsmæssige set fra et formålsperspektiv og dermed i lyset af lovgivers hensigt med indskrænkningerne. Indskrænkningerne havde til formål at sikre et højt beskyttelsesniveau af rettighedshaverne, og desuden blev indskrænkningerne ikke vedtaget mhp. at skulle dække kopiering ifm. træningen af en ML-algoritme. Indskrænkningerne er dog ikke hensigtsmæssige set fra et globalt konkurrenceperspektiv, idet adgangen til kopiering af værker ifm. træningen af en ML-algoritme uden tilladelse er snævrere i EU (herunder Danmark) end i lande uden for EU såsom Japan og USA, hvilket kan være hæmmende for virksomheder, der udvikler generative AI-systemer og dermed for EU's (herunder Danmarks) konkurrenceposition på det globale marked for AI.

En løsning på de udfordringer, som de nugældende indskrænkninger indebærer for generative AI-udviklere, er en ny indskrænkning, der i videre omfang dækker kopiering af værker ifm. træningen af en ML-algoritme. Da sådanne regler dog næppe bliver vedtaget inden for den nærmeste fremtid, må løsningen indtil da være at afvente, at de nugældende indskrænkningers anvendelsesområde, hvad angår udvikling af generative AI, bliver afklaret i praksis.

8 Litteraturliste

LOVGIVNING OG LOVFORARBEJDER

Dansk lovgivning:
Love:
- Lovbekendtgørelse nr. 1144 af 23. oktober 2014 om ophavsret – "OPHL"

Lovforarbejder:
- Forslag L 19 af 2. oktober 2002 til Lov om ændring af ophavsretsloven (Gennemførelse af InfoSoc-direktivet, nye afdelicenser m.v.), FT 2002-03 – "L 19 2002-03"
- Forslag L 58 af 30. januar 2008 til Lov om ændring af ophavsretsloven (Overdragelse af ophavsret, digitalisering af kulturarven m.v.), FT 2007-08 (.2 samling) – "L 58 2007-08"
- Forslag L 205 2020-21 af 26. marts 2021 til Lov om ændring af lov om ophavsret (Implementering af dele af direktiv om ophavsret og beslægtede rettigheder på det digitale indre marked), FT 2020-21 – "L 205 2020-21"
- Forslag L 125 af 3. maj 2023 til Lov om ændring af lov om ophavsret (gennemførsel af dele af direktiv om ophavsret og beslægtede rettigheder på det digitale indre marked), FT 2022-23 (2. samling) – "L 125 2022-23"

EU-ret:

Traktater:
- Traktaten om den Europæiske Unions Funktionsmåde – "TEUF"

Direktiver:

Forslag til retsakter:
- Forslag til Europa-Parlamentets og Rådets forordning om harmoniserede regler for kunstig intelligens (retsakten om kunstig intelligens) og om ændring af visse af unionens lovgivningsmæssige retsakter, COM(2021) 206, 21. april 2021 – "Forslag til AI-forordningen"
National lovgivning fra andre lande end Danmark:

Love:

- Lag 1960:729 om upphovsrätt till litterära och konstnärliga verk – "Den svenska ophavsretslov"
- Tekijänoikeuslaki (8.7.1961/404) – "Den finske ophavsretslov"
- The Japanese Copyright Act – engelsk version kan tilgås på [https://www.cric.or.jp/english/clj/] - "Den japanske ophavsretslov"
- The U.S. Copyright Act (Title 17 of the United States Code) – "Den amerikanske ophavsretslov"

RETSpraakS og ADMINistrativ PRAKsis

Danske afgørelser:

Retspaksis:


Administrative afgørelser:

- Kulturministeriets afgørelse af 31. marts 2023: Godkendelse af Copydan Tekst & Node i henhold til ophavsretslovens § 50, stk. 4, jf. § 50, stk. 2 (aftale digital eksemplarfremstilling
og tilgængeliggørelse for almenheden af ophavsretningsligt beskyttet materiale mellem
Foreningen lex.dk, VISDA og Copydan Tekst & Node) -

Afgørelser fra EU-Domstolen:

- Domstolens Dom af 4. december 1974 i sag 41-74, Yvonne van Duyn mod Home Office, Indenrigsministeriet – "van Duyn"
- Domstolens Dom af 26. februar 1986 i sag 152/84, M. H. Marshall mod Southampton and South-West Hampshire Area Health Authority (Teaching) – "Marshall"
- Domstolens Dom (Sjette Afdeling) af 6. februar 2003 i sag C-245/00, Stichting ter Exploitatie van Naburige Recthen (SENA) mod Nederlandse Omroep Stichting (NOS) – "SENA"
- Domstolens Dom (Femte Afdeling) af 29. april 2004 i sag C-476/01, Straffesag mod Felix Kapper – "Kapper"
- Domstolens Dom (Tredje Afdeling) af 26. oktober 2006 i sag C-36/05, Kommissionen for De Europeiske Fællesskaber mod Kongeriget Spanien – "Kommissionen mod Spanien"
- Domstolens Dom (Tredje Afdeling) af 7. december 2006 i sag C-306/05, Sociedad General de Autores y Editores de Españ̦a (SGAE) mod Rafael Hotels SA – "SGAE"
- Domstolens Dom (Fjerde Afdeling) af 16. juli 2009 i sag C-5/08, Infopaq International A/S mod Danske Dagblades Forening – "Infopaq I"
- Domstolens Dom (Store Afdeling) af 4. oktober 2011 i de forenede sager C-403/08 og C-429/08, Football Association Premier League Ltd m.fl. mod QC Leisure m.fl. og Karen Murphy mod Media Protection Services Ltd – "Premier League"
- Domstolens Dom (Tredje Afdeling) af 1. december 2011 i sag C-145/10, Eva-Maria Painer mod Standard VerlagsGmbH m.fl. – "Painer"
- Domstolens Kendelse (Tredje Afdeling) af 17. januar 2012 i sag C-302/10, Infopaq International A/S mod Danske Dagblades Forening – "Infopaq II"
- Domstolens Dom (Fjerde Afdeling) af 5. juni 2014 i sag C-360/13, Public Relations Consultants Association Ltd mod Newspaper Licensing Agency Ltd m.fl. – "Meltwater"
- Domstolens Dom (Anden Afdeling) af 26. april 2017 i sag C-527/15, Stichting Brein mod Jack Frederik Wullems – "Filmspeler"

Udenlandsk retspraksis:

• Judgement of the High Court of Justice (England and Wales) (Chancery Division) - 3rd of February 2012, [2012] EWHC 108 (Ch), Football Association Premier League Ltd. et. al. vs. QC Leisure et. al., AV Station Plc et. al., and Michael Madden et. al.

MEDDELELSER, RAPPORTER MV.

Meddelelser fra EU-Kommissionen:


Rapporter og øvrige publikationer fra EU-Kommissionen:

• EU-Kommissionen - Directorate-General for Communications Networks, Content and Technology: *Trends and Developments in Artificial Intelligence – Challenges to the Intellectual Property Framework* (2020) – "EU-Kommissionens rapport om AI (2020)"
• EU-Kommissionen - Directorate-General for Communications Networks, Content and Technology: *Study on copyright and new technologies: Copyright data management and artificial intelligence* (2022) – "EU-Kommissionens rapport om AI (2022)"

Andre rapporter, publikationer mv.:


**LITTERATUR**

Bøger:

• Axhamn, Johan: *Copyright and Artificial Intelligence – with a focus on the area of music*, fra Rosenmeier, Morten et al.: *Festskrift til Jørgen Blomqvist*, Ex Tuto Publishing (2021) – "Axhamn (2021)"


• Sørensen, Karsten Engsig *et al.*: *Uddrag af EU-Retten*, 7. rev. udgave, Djøf Forlag (2019) – "Engsig Sørensen *et al.* (2019)"


**Juridiske tidsskriftsartikler o.l.:**


• Geiger, Christophe *et al.*: *Crafting a Text and Data Mining Exception for Machine Learning and Big Data in the Digital Single Market*, Intellectual Property and Digital Trade in the
Age of Artificial Intelligence and Big Data, Global Perspectives and Challenges for the Intellectual Property System, Issue No. 5 – "Geiger et al. II (2018)"

- Rosati, Eleonora: Copyright as an Obstacle or an Enabler? A European Perspective on Text and Data Mining and Its Role in the Development of AI Creativity, 27(2) Asia Pacific Law Review 198-217 (2020) – "Rosati (2020)"
- Schönberger, Daniel: Deep Copyright: Up – and Downstream Questions Related to Artificial Intelligence (AI) and Machine Learning (ML), Droit D'auteur 4.0/Copyright 4.0, DE WERRA Jacques (ed.) (2018) – "Schönberger (2018)"
INTERNETKILDER

Juridiske webartikler, indlæg o.l.:

- Butterick, Matthew: *Because AI needs to be fair & ethical for everyone*, Stable Diffusion litigation (2023) - [https://stablediffusionlitigation.com/](https://stablediffusionlitigation.com/) (lokalisert d. 7. marts. 2023 kl. 09:28) – "Butterick (2023)"

Webartikler om machine learning:


88


Øvrige internetskilder:


- ChatGPT - https://openai.com/blog/chatgpt


- Den Danske Ordbog om begrebet "flygtig" - https://ordnet.dk/ddo/ordbog?query=flygtig (lokalisert d. 15. marts 2023 kl. 09:34)

- Den Danske Ordbog om begrebet "tilfeldig" - https://ordnet.dk/ddo/ordbog?query=tilf%C3%A6ldig (lokalisert d. 15. marts 2023 kl. 09:38)


- Koetsier, John: GPT-4 Beats 90% Of Lawyers Trying To Pass The Bar, Forbes (2023) - https://www.forbes.com/sites/johnkoetsier/2023/03/14/gpt-4-beats-90-of-lawyers-trying-to-pass-the-bar/?sh=1b5f48a53027 (lokalisert d. 27. maj kl. 09:48)


• Kulturministeriets oversigt over aftalelicensgodkendelser (se under fanen "Godkendelser efter den generelle aftalelicens") – [https://kum.dk/kulturomraader/aftalelicensgodkendelser](https://kum.dk/kulturomraader/aftalelicensgodkendelser) (lokalisert d. 9. maj 2023 kl. 16:47)

• Liberatore, Stacy: *Bill Gates says ChatGPT-like AI is 'as revolutionary as mobile phones and the internet' - and says it could speed up creation of new vaccines*, Daily Mail (2023) - [https://www.dailymail.co.uk/sciencetech/article-11896051/Bill-Gates-says-AI-revolutionary-mobile-phones-Internet.html](https://www.dailymail.co.uk/sciencetech/article-11896051/Bill-Gates-says-AI-revolutionary-mobile-phones-Internet.html) (lokalisert d. 27. maj 2023 kl. 10:23)


• Stable Diffusion Online - [https://stablediffusionweb.com/](https://stablediffusionweb.com/)
