

**UiO** • **Institutt for privatrett**  
Det juridiske fakultet

# Tilgang til data for maskinlæring – opphavsrettslige problemstillinger

Professor Inger B. Ørstavik





Foto: Rijksmuseum Nederland

DIGITALISERINGSSTRATEGI

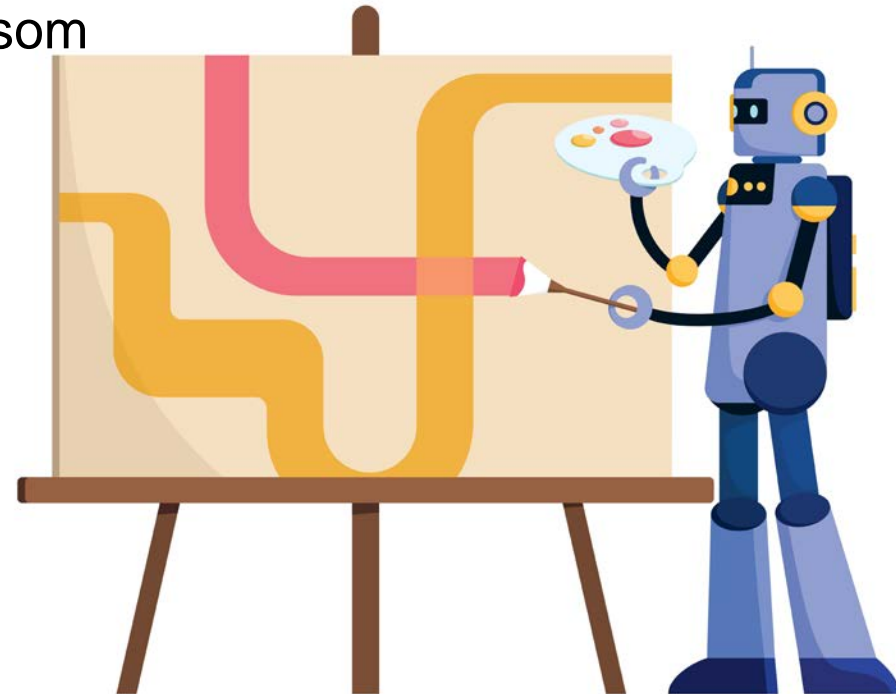
# NorGPT kan bli redningen for det norske språket

En norsk språkmodell har ligget i hardtrenoing i NTNUs supercomputer Idun i vinter. Mye står på spill. På sikt: Det norske språks overlevelse.



# Hva skjer når en maskin «lærer»?

- *Tre trinn:*
  - Definere og programmere en algoritme
  - **Trene opp algoritmen på et sett treningsdata**
  - Bruke algoritmen på nye data som en del av en tjeneste
- **Treningsdata**
  - Tilstrekkelig stor mengde
  - Representative
  - Må kunne generaliseres
- Sentralisert maskinlæring
- Desentralisert maskinlæring



# Rettighetsspørsmål ved tilgang til treningsdata

- Eksemplarfremstillingsretten
  - Åvl. § 3, Infosoc art. 2
- Databaserett, åvl § 24
- Ny enerett for utgivere av pressepublikasjoner
  - DSM (2019/790 Digital Single Market) art. 15
- Nye unntak for tekst- og datautvinning
  - DSM art. 3 og 4

# Eksemplarfremstillingsretten

- Åvl § 3 a jf Infosoc art. 2
  - Digitale kopier, også midlertidige kopier
  - Trenger ikke være leselige for et menneske
  - CJEU C-360/13 *Meltwater*, også RAM og cache memory
  - «kjøre» algoritmer vil stort sett innebære eksemplarfremstilling



# Databaserettigheter

- Åvl § 24, jf. Databasedirektivet art. 7
  - Vesentlig investering i prosessering for treningskorpus
- Gjør ML inngrep i databaserett?
  - «uttrekk» av vesentlige deler
  - «gjentatt og systematisk uttrekk» av uvesentlige deler
- Samtykke til bruk? (åvl § 4 gjelder ikke)
  - CJEU C-203/02, *William Hill* (57-58)
  - samtykke til søk i databasen er ikke samtykke til «uttrekk»/ «gjentatt og systematisk uttrekk» (se også fortalen til databasedirektivet avsn (42))
  - Databaser på åpent nett, (C-202/12, *Innoweb*)

# Ny enerett til utgivere av pressepublikasjoner

- DSM art. 15:

”1. Medlemsstaterne sørger for, at udgiverne af publikationer, der er etableret i en medlemsstat, indrømmes de i artikel 2 og artikel 3, stk. 2, i direktiv 2001/29/EF fastsatte rettigheder med hensyn til udbydere af **informationssamfundstjenesters onlineanvendelse** af deres pressepublikationer.»

- Dansk Ophavsretslov § 69 a
- Svenska Upphovsrättslagen 48 b §



# Ny enerett til utgivere av pressepublikasjoner

- Informasjonssamfunnstjenester
  - Direktiv 2015/1535 art. 1(1)b) (Information Society Services) «enhver tjeneste som vanligvis ytes mot vederlag, over avstand, og som formidles elektronisk og på individuell anmodning fra en tjenestemottaker»
- 2 år fra publisering
- Hva er «**onlineanvendelse**»?
  - Må «ikke undergrave de investeringer, som utgivere af pressepublikationer foretager i forbindelse med produktion af indhold» (fortalen pkt 58)

# Maskinlæring og tekst- og datautvinning

- DSM art 2(2):
  - «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»
- Omfatter ML metoder



# Det generelle unntaket for TDM, DSM art. 4

1. Medlemsstaterne fastsetter en **undtagelse fra eller en innskærnkning** af de rettigheder, der er fastsat i artikel 5, litra a), og artikel 7, stk. 1, i direktiv 96/9/EF, artikel 2 i direktiv 2001/29/EF, artikel 4, stk. 1, litra a) og b), i direktiv 2009/24/EF og artikel 15, stk. 1, i nærværende direktiv, **for reproduktion og udtræk af lovligt tilgængelige værker eller andre frembringelser med henblik på tekst- og datamining.**
2. Reproduktioner og udtræk foretaget i henhold til stk. 1, **kan opbevares, så længe det er nødvendigt med henblik på tekst- og datamining.**
3. Undtagelsen eller innskærnkningen omhandlet i stk. 1 finder anvendelse på betingelse af, at **rettighedshaverne til værker og andre frembringelser, der er omhandlet i nævnte stykke, ikke udtrykkeligt har forbeholdt sig anvendelsen heraf på passende vis, såsom en maskinlæsbar måde** i tilfælde af indhold, der er stillet til rådighed for almenheden online.
4. Denne artikel berører ikke anvendelsen af artikel 3 i dette direktiv.

# Det generelle unntaket for TDM, DSM art. 4

- fra databaserett, opphavsrett og eneretten til pressepublikasjoner
- Lovlig tilgjengelige verk
- Art. 4(3): individuelle rettighetshavere kan ta forbehold
  - I praksis forlaget
  - Online tilgjengelig – forbeholdet må gjøres i maskinlesbar form
  - Eksisterende kontrakter
  - Kan ikke reservere seg mot art. 3 (forskning)

# Unntaket for TDM i forskning, DSM art. 3

1. Medlemsstaterne fastsætter en **undtagelse** fra de rettigheder, der er fastsat i artikel 5, litra a), og artikel 7, stk. 1, i direktiv 96/9/EF, artikel 2 i direktiv 2001/29/EF og artikel 15, stk. 1, i nærværende direktiv for **reproduktioner og udtræk foretaget af forskningsorganisationer og kulturarvsinstitutioner med henblik på at udføre tekst- og datamining af værker eller andre frembringelser, som de har lovlig adgang til, med henblik på videnskabelig forskning.**
2. Kopier af værker eller andre frembringelser foretaget i overensstemmelse med stk. 1 lagres med et passende sikkerhedsniveau og **kan opbevares med henblik på videnskabelig forskning, herunder med henblik på verifikation af forskningsresultater.**
3. Det er tilladt rettighedshavere at træffe foranstaltninger, der kan garantere sikkerhed og integritet i de netværk og databaser, hvor værkerne eller de andre frembringelser er lagret. Sådanne foranstaltninger må ikke gå videre, end hvad der er nødvendigt for at nå dette mål.
4. Medlemsstaterne opfordrer rettighedshavere, forskningsinstitutioner og kulturarvsinstitutioner til i fællesskab at fastlægge bedste praksis med hensyn til anvendelsen af den forpligtelse og de foranstaltninger, der omhandles i henholdsvis stk. 2 og 3.

# Unntaket for TDM i forskning, DSM art. 3

- Forskningsorganisasjoner og kulturarvinstitusjoner
  - Uten mål om fortjeneste eller utfører allmenntilgitt oppgaver med offentlig finansiering eller på offentlige kontrakter (fortalen 12)
  - Nasjonale arkiv, nasjonalbibliotek (fortalen 13)
- Vitenskapelig forskning
- Lovlig tilgang
  - Open access, abonnement, fritt online
- Begrensninger kun for informasjonssikkerhet
- Kan oppbevare dataene for forskning og verifikasjon av forskningsresultater
- Rettighetshavere kan ikke reservere seg
- Ikke vederlag (fortalen 17)

## Forholdet til Infosoc direktivet 2001/29

- Åvl § 4, Infosoc art. 5(1) om midlertidige eksemplarer uten selvstendig økonomisk betydning
  - Selvlærende algoritmer der resultatet av tjenesten ikke gjør inngrep
- Tre trinns testen i Infosoc art. 5(5) anvendes på unntakene for TDM
  - Spesielle tilfelle
  - Ikke strider mot normal utnyttelse av verket
  - Ikke urimelig skade rettighetshavers legitime interesser



## Status i Norge?

- Høringsnotat med forslag til endringer bl.a. i åndsverkloven forventes høst 2023
- Merk: DSM Art. 4 (TDM utenfor forskning) var ikke i det opprinnelige forslaget fra EU, og ikke i første høringsrunde
- [Digitalmarkedsdirektivet - regjeringen.no](https://www.regjeringen.no)

# Danmark: ny § 11 b og c i ophavsretsloven

- *Tekst- og datamining*

- **§ 11 b.** Ved tekst- og datamining forstås enhver automatiseret, analytisk fremgangsmåde, der har til formål at analysere tekst og data i digital form med henblik på at generere oplysninger, herunder mønstre, tendenser og korrelationer.
- *Stk. 2.* Den, som har lovlig adgang til et værk, må foretage udtræk og fremstille eksemplarer af værket med henblik på tekst- og datamining, på betingelse af at rettighedshaveren ikke udtrykkeligt har forbeholdt sig anvendelsen af værket på passende vis, jf. dog § 11 c, stk. 1.
- *Stk. 3.* Udtræk foretaget og eksemplarer fremstillet i henhold til stk. 2 kan opbevares, så længe det er nødvendigt med henblik på tekst- og datamining.

- *Tekst- og datamining med henblik på videnskabelige forskning*

- **§ 11 c.** Forskningsorganisationer og kulturarvsinstitutioner, som har lovlig adgang til et værk, må foretage udtræk og fremstille eksemplarer af værket for at udføre tekst- og datamining, jf. § 11 b, stk. 1, med henblik på videnskabelig forskning. 1. pkt. gælder, uanset om den videnskabelige forskning udføres inden for rammerne af offentlig-private partnerskaber. 1. pkt. kan ikke fraviges ved aftale.
- *Stk. 2.* Eksemplarer af værker fremstillet i overensstemmelse med stk. 1 skal lagres med et passende sikkerhedsniveau og kan opbevares med henblik på videnskabelig forskning.
- *Stk. 3.* Det er tilladt for rettighedshavere at træffe foranstaltninger, der kan garantere sikkerhed og integritet i de netværk og de databaser, hvor værkerne er lagret. Sådanne foranstaltninger må ikke gå videre, end hvad der er nødvendigt for at nå dette mål.

# Sverige: ny 15a-c § Upphovsrättslag

## Framställning av exemplar för text- och datautvinning

**15 a §** Den som har lovlig tillgång till ett verk får framställa exemplar av verket för text- och datautvinningsändamål. Exemplaren får inte behållas längre än vad som är nödvändigt för ändamålet och får inte användas för andra ändamål.

Första stycket gäller inte om upphovsmannen på lämpligt sätt har förbehållit sig den rätt som avses där. *Lag (2022:1712)*.

**15 b §** Forskningsorganisationer, sådana bibliotek och museer som är tillgängliga för allmänheten, arkiv samt institutioner för film- eller ljudarvet får framställa exemplar av verk som de har lovlig tillgång till, dock inte datorprogram, för att utföra text- och datautvinning för forskningsändamål. Exemplaren får inte behållas längre än vad som är nödvändigt för ändamålet och får inte användas för andra ändamål. Exemplaren ska lagras på ett sätt som hindrar obehörig användning.

Första stycket hindrar inte att upphovsmannen vidtar proportionerliga åtgärder för att säkerställa integritet och säkerhet i nätverk och databaser som innehåller verk.

Avtalsvillkor som inskränker rätten att använda verk enligt denna paragraf är ogiltiga. *Lag (2022:1712)*.

**15 c §** Med text- och datautvinning avses i 15 a och 15 b § § en automatiserad teknik som används för att analysera text och data i digital form i syfte att generera information.

Med forskningsorganisation avses i 15 b § ett universitet med tillhörande bibliotek, en forskningsinstitution eller en annan enhet vars främsta mål är att bedriva vetenskaplig forskning eller utbildningsverksamhet som omfattar vetenskaplig forskning. Vidare måste verksamheten i organisationen bedrivas på ett sådant sätt att ett företag som utövar ett avgörande inflytande på organisationen inte kan ha prioriterad tillgång till resultaten av forskningen. Forskningen måste bedrivas

1. utan vinstsyfte eller så att all vinst återinvesteras i den vetenskapliga forskningen, eller

**Takk!**

i.b.orstavik@jus.uio.no