Lovlig bruk av data til(t) trening og bruk av maskinlæringsalgoritmer

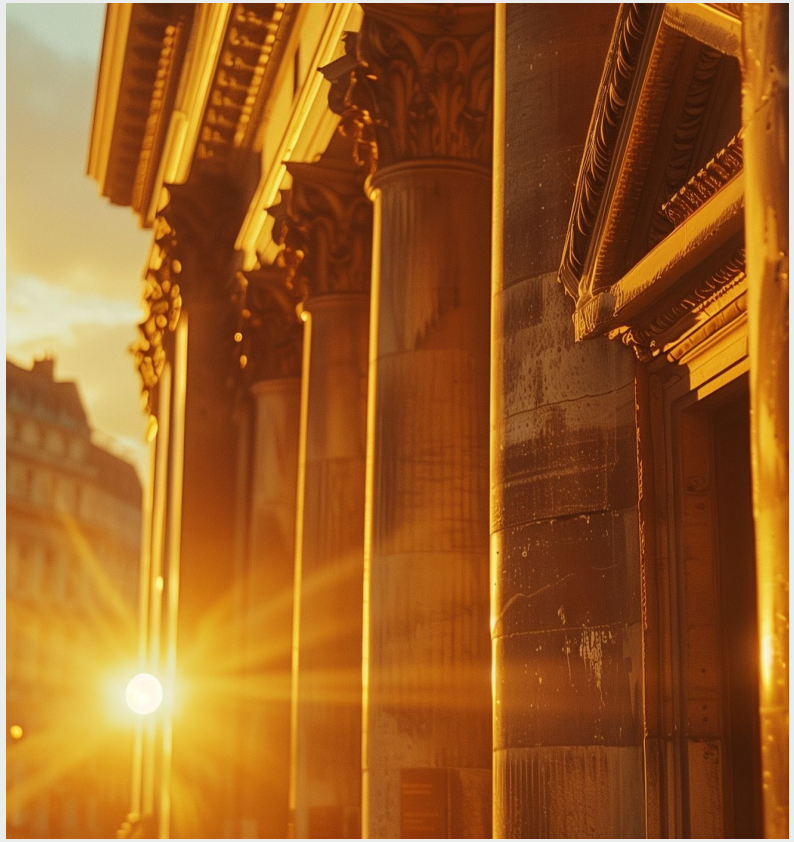
Advokatfullmektig William Eitrem og advokat Ane Rode

Agenda

1. Kort om generative maskinlæringsmodeller
2. Hva mener vi når vi sier data?
3. Immaterialrettslige hindringer
4. Personvernrettslige hindringer
5. Finne relevant data
    a. Skraping
    b. Avtaler
    c. Eksempeltekst i kunde- og leverandøravtaler
6. Use caser
7. Avslutning

Kort om generative maskinlærings-mod eller

- Fem behandlinger
  a) Trening
  b) Fine-tuning
  c) «Lagring»
  d) Ledetekst/prompting
  e) Inferens/generering
- Data – a) og b) (og d?)
- Hvor ligger verdien for selskaper?
  - Applikasjoner
  - Infrastruktur



Hva mener vi når vi sier data?

- Faktisk
  - Tekst
  - Bilder
  - Video
  - Lyd
  - Resonnementer
  - Metadata
- Juridisk
  - Ikke legaldefinert
  - Forbud
  - Påbud
  - Frihet

Immatrialrettslige hindringer

- Opphavsrett, åvl. §§ 2 og 3
  - Enerett til å fremstille eksemplarer og tilgjengeliggjøre for allmennheten
    - Tekster, musikkverk, filmverk, fotografiske verk, malerier, tegninger, grafikk, bygningskunst, kart, datamaskinprogrammer oversettelser, og andre bearbeidelser av verk som er nevnt foran
- Varemerke
- Databasevern, åvl. § 24
  - «*hvor innsamling, kontroll eller presentasjon av innholdet innebærer en vesentlig investering*»

Personvernrettslige hindringer

- To perspektiv:
  - Behandlingsansvarlig
  - Databehandler

Personvernrettslige hindringer

Behandlingsansvarlig

- GDPR artikkel 6
  - Samtykke
  - Berettiget interesse
- Husk: fem behandlinger
- Ikke direkte forhold med de registrerte – hvordan sikre lovlig behandling?
- Anonymisering?
- Informasjonsplikten
  - Eks. Slack

Machine learning (ML) and artificial intelligence (AI) are useful tools that we use in limited ways to enhance our product mission. We do not develop LLMs or other generative models using customer data. To develop non-generative AI/ML models for features such as emoji and channel recommendations, our systems analyse Customer Data (e.g. messages, content and files) submitted to Slack as well as Other Information (including usage information) as defined in our privacy policy and in your customer agreement. To ensure the privacy and security of Customer Data in this particular context, we have a few guiding principles:

Contact us to opt out. If you want to exclude your Customer Data from Slack global models, you can opt out. To opt out, please have your org, workspace owners or primary owner contact our Customer Experience team at feedback@slack.com with your workspace/org URL and the subject line 'Slack global model opt-out request'. We will process your request and respond once the opt-out has been completed.

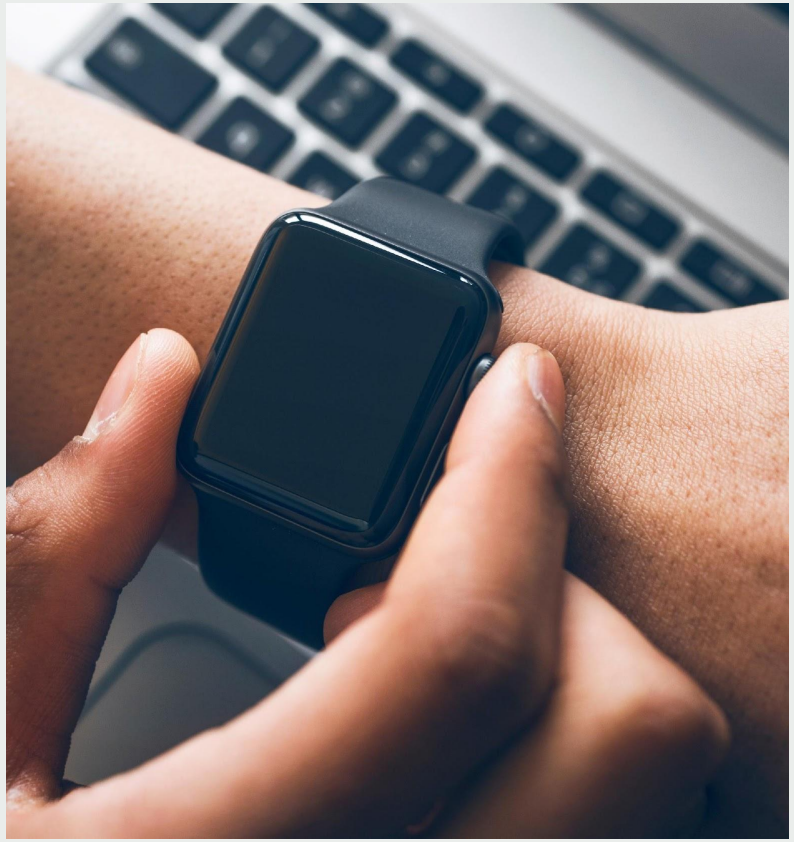
Personvernrettslige hindringer

Databehandler

- Personopplysninger behandles på vegne av kunden i henhold til databehandleravtalen
- Behandling utover instruksene?
  - (i) Brudd på databehandleravtalen
  - (ii) Databehandler identifiseres som behandlingsansvarlig, jf. GDPR art. 28 (10)

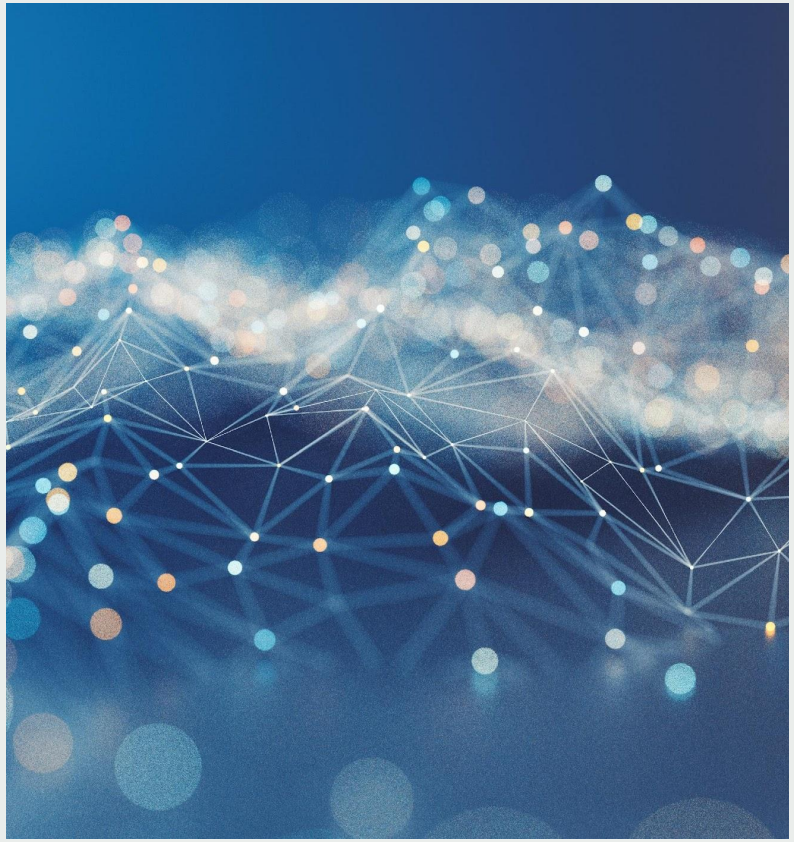
Finne relevant data
Skraping

- Automatisert innsamling av data
  - Nettsider
  - Dokumenter
  - API-er (ish)
- Omstridt
- Digitalmarkedsdirektivet
  - Generell rett til tekst- og datautvinning så lenge ikke rettighetshaveren har «uttrykkelig» sagt nei
- Robot.txt
- **Utvis forsiktighet med enkelte typer data**



Finne relevant data
Avtaleforhold

- Data som hovedytelse (*prime clause*)
  - Ad-tech
  - Markedsinformasjon
  - Meningsmålinger
- Bruksdata
  - Tekst
  - Videodata – Zoom
  - Musikkdata
- Metadata
  - Tidspunkt for interaksjoner



Eksempeltekst i kunde- og leverandøravtaler

Eksempel 1
[Company] *may use customer data (including personal data) for machine learning to support and develop its features and functionality within the subscription service and similar products and services. Customers can opt-out by sending an email to privacy@[company].com.*

Eksempel 2
Service Improvement: [Company] *may use Customer Data to enhance and provide* [Company] *services, including the use of machine learning algorithms to improve functionality and performance.*

Eksempel 3
Research and Development: We may use the information and data we have about you to conduct research and development for the furtherance of knowledge, development of new technologies, and creation of new products and services. This includes using the data for machine learning and artificial intelligence purposes.

SSA-L

7. EIENDOMS- OG DISPOSISJONSRETT

7.1 PARTENES RETTIGHETER

Denne avtalen endrer ikke de opphavs-, disposisjons- eller eiendomsrettighetene partene hadde før avtalen, og som de beholder under gjennomføringen av avtalen, med mindre annet fremgår av bilag 1 eller 2.

Tilgangen til tjenesten omfatter alle de beføyelser som er nødvendig for å benytte tjenesten i henhold til avtalens formål. Med mindre annet er særskilt avtalt, overdras ingen immaterielle rettigheter til Kunden. Kunden har heller ikke eksklusiv tilgang til tjenesten med mindre dette avtales særskilt.

7.2 EIENDOMSRETT TIL DATA

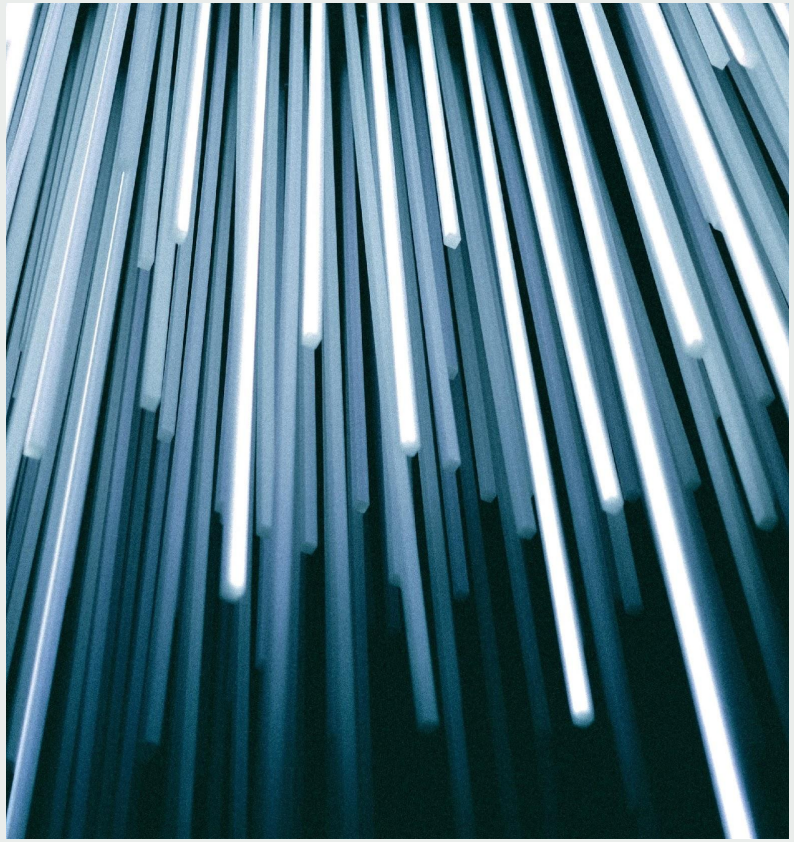
Kunden beholder eiendomsrett til alle data som overlates til Leverandøren for behandling, og som lagres eller prosesseres ved hjelp av tjenestene under denne avtalen. Det samme gjelder resultatet av Leverandørens behandling av slike data.

Leverandøren har tilgang til data som nevnt ovenfor utelukkende i den utstrekning som er nødvendig for at Leverandøren skal kunne oppfylle sine forpliktelser i henhold til avtalen.

Leverandøren har ikke under noen omstendighet rett til å utøve tilbakeholdsrett i Kundens data.

Use Caser

- Programmering
- Kundeservice
- Etterspørselsplanlegging
- Grafisk design
- Juridisk
- Journalistikk
- Ed-tech



Avslutning

- Data er i utgangspunktet fritt vilt
- Pass på opphavsrett – verden følger med
- Vanskelig å bruke personopplysninger

Takk for oss!

William Eitrem
Associate

william.eitrem@schjodt.com
+47 469 60 452
+47 23 01 17 22
Oslo

Ane Victoria Rode
Senior Associate

ane.rode@schjodt.com
+47 23 01 18 77
+47 464 66 946
Oslo