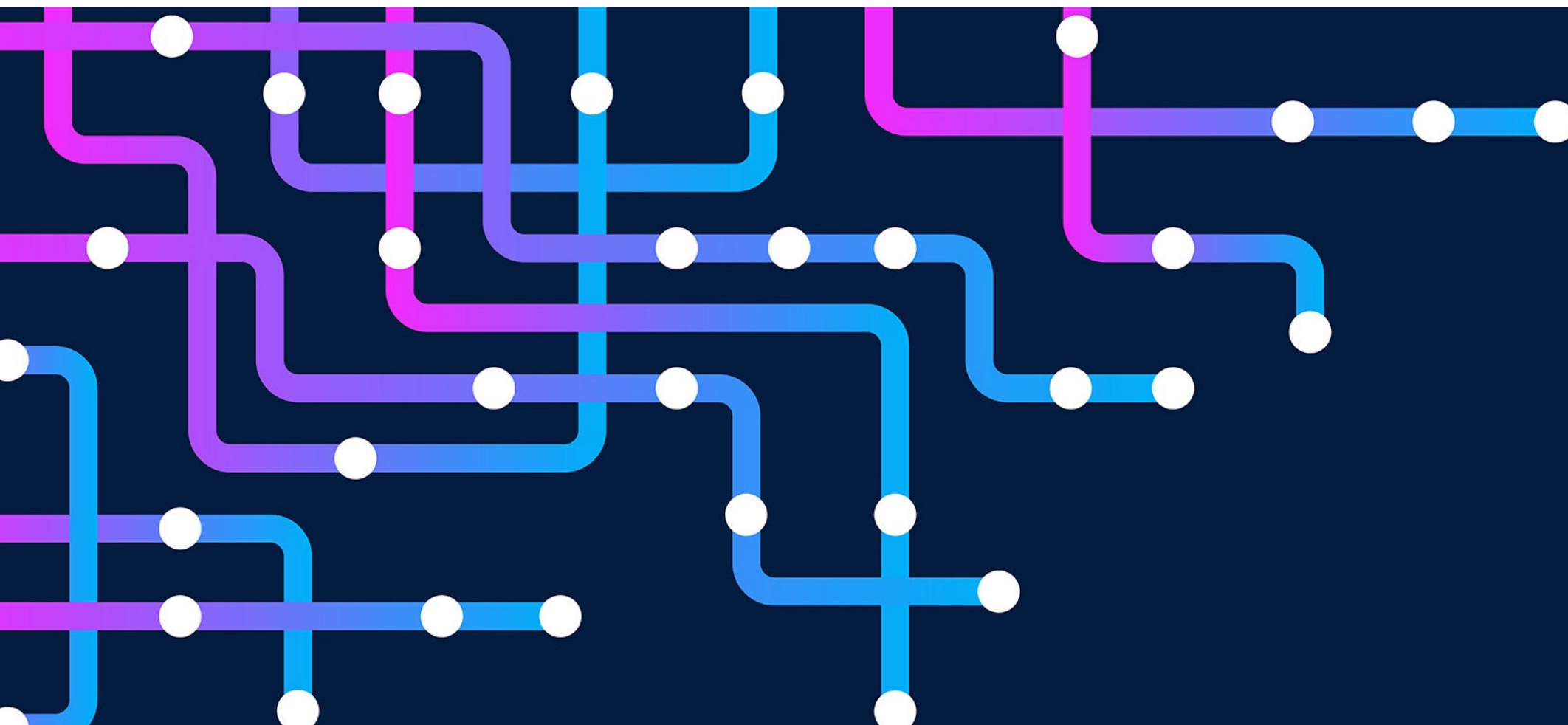


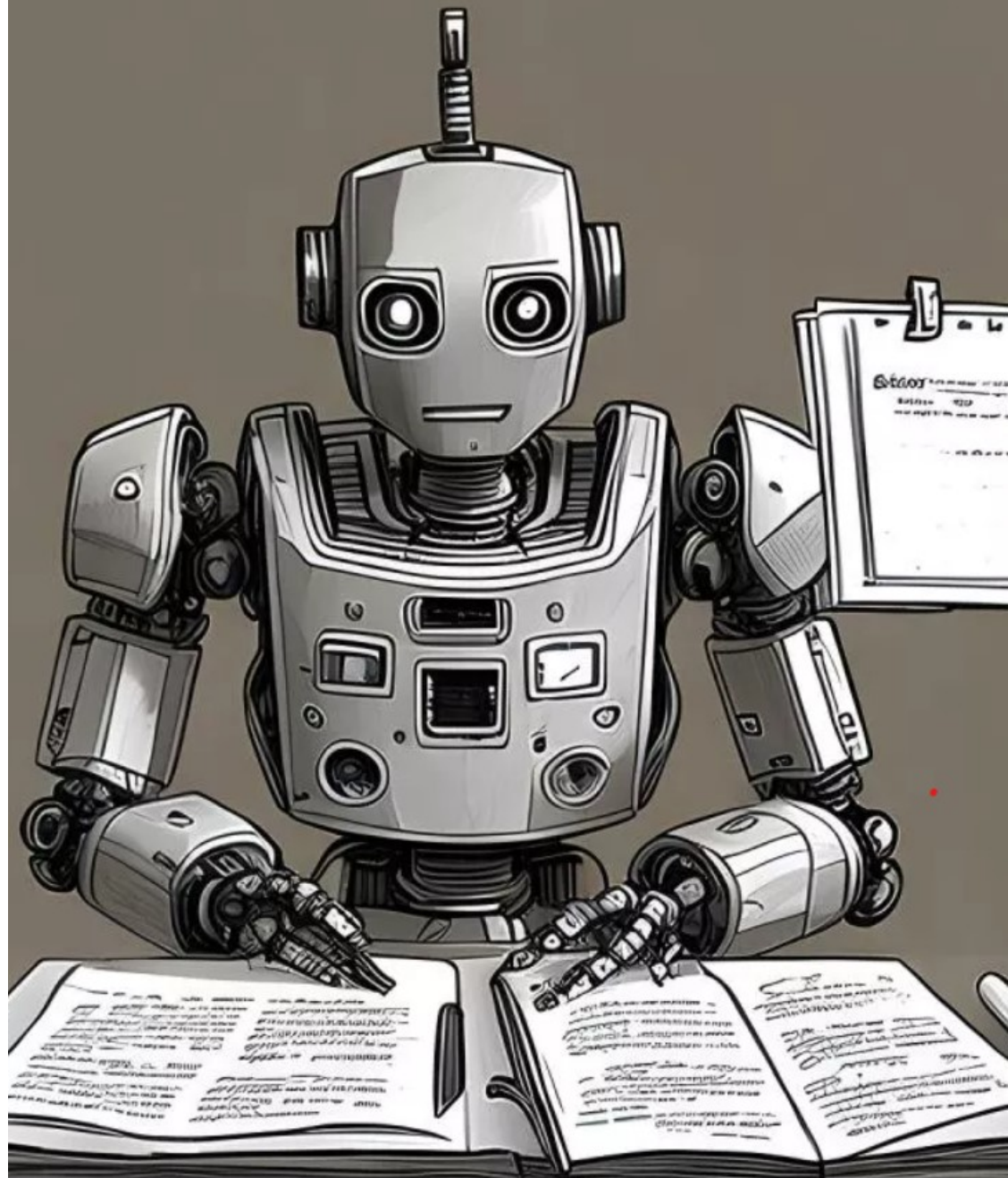
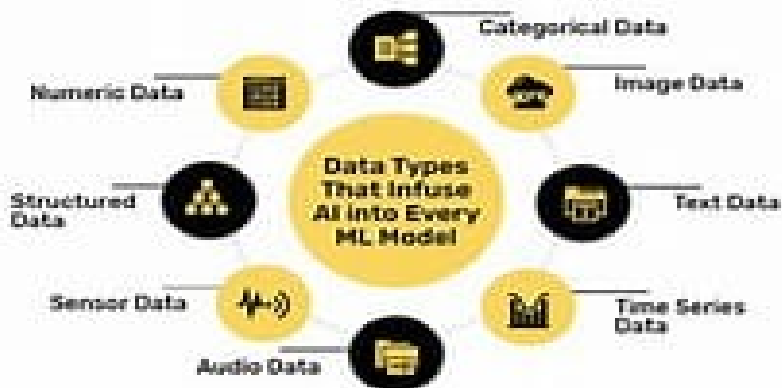
# LOVLIG OG ANSVARLIG BRUG AF DATA

Væsentlige opmærksomhedspunkter



# ”DATA ER DEN NYE OLIE”

- Data er en af de vigtigste faktorer i AI
- Gennem data kan AI systemerne lære, træne og forbedre sig selv over tid
- AI-systemer er afhængige af store mængder data for at kunne identificere mønstre og sammenhænge i dataene og træne sig selv til at udføre specifikke opgaver. Jo mere data, jo mere præcist
- Typer af data - tekst, billede, lyd



# LOVLIGHED OG DATAKVALITET ER AFGØRENDE

“Garbage doesn’t turn into gold,  
no matter how much math you  
throw at it.”

- Cassie Kozyrkov

# ANSVARLIGHED OG DATA ETIK BØR VÆRE GRUNDLÆGGENDE MINDSET

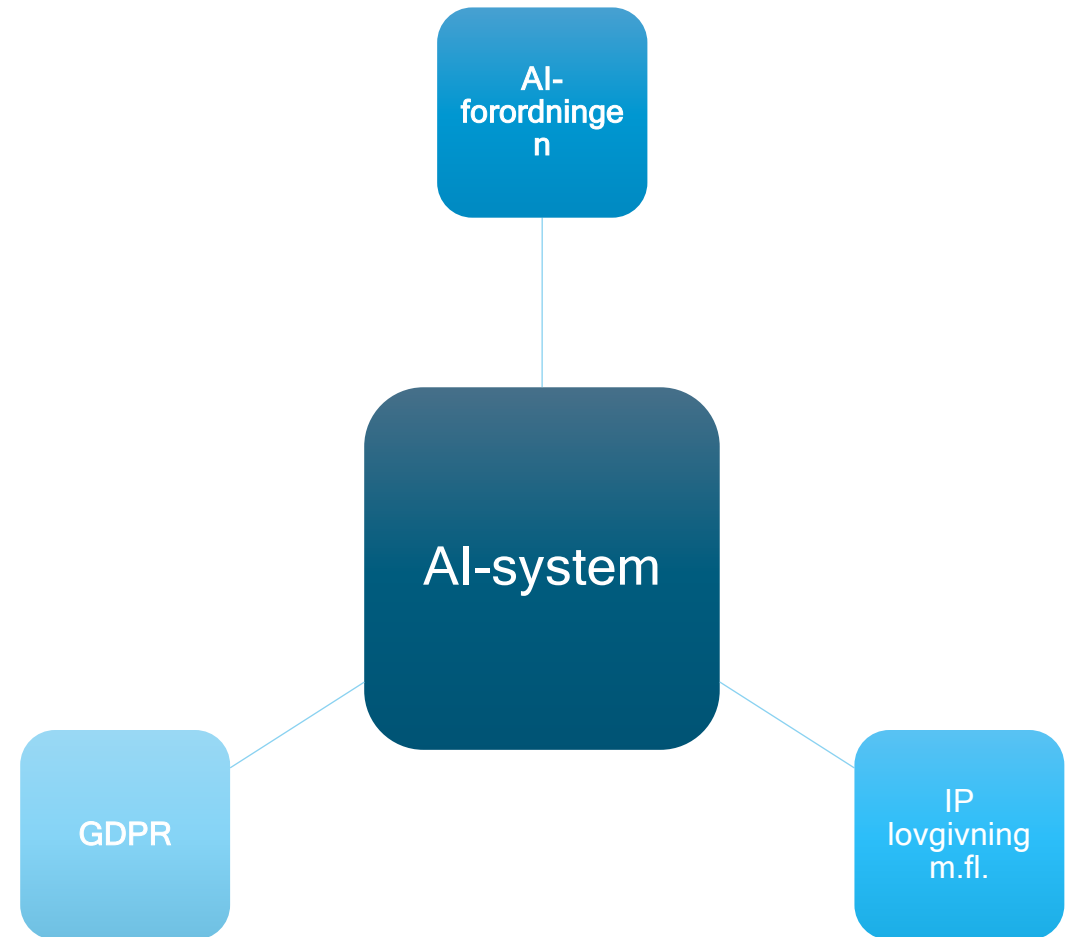
- Hvad er data etik? ”*Dataetik handler generelt om god praksis, når man indsamler, bruger og deler data.*”
- Husk, at:

*“At gøre hvad du efter loven har ret til, er ikke altid det samme, som at gøre hvad der er godt eller rigtigt.”*

- Dataetisk Råd (<https://nationaltcenterforetik.dk/etiske-temaer/dataetik>)


# LOVLIGHED

- Lovlig brug af data i AI kræver overholdelse af bl.a.
  - Databeskyttelsesforordningen
  - IP lovgivning
  - Særlovgivning - hvidvask mv.
- AI-forordningen - vil gælde side om side med og/eller supplerer eksisterende regler.



# AI-FORORDNINGEN

- En række generelle principper for alle AI systemer, bl.a.
- Privacy and data governance
  - “*Developed and used in compliance with existing privacy and data protection rules, while processing data that meets high standards in terms of quality and integrity*”

 EUROPA-KOMMISSIONEN


Bruxelles, den 21.4.2021  
COM(2021) 206 final  
2021/0106 (COD)

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

{SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final}

DA

 Council of the European Union

Bruxelles, 29 November 2021  
(OR\_en)

14278/21

Interinstitutional File:  
2021/0106(COD)

LIMITE

TELECOM 430  
JAI 1288  
COPEN 412  
CYBER 307  
DATAPROTECT 267  
EJUSTICE 103  
COSI 236  
IXIM 262  
ENFOPOL 465  
FREMP 272  
RELEX 1012  
MI 879  
COMPET 860  
CODEC 1530

**NOTE**

From: Presidency  
To: Delegations  
No. Cion doc.: 8115/20  
Subject: Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts  
- Presidency compromise text

**I. INTRODUCTION**

1. The Commission adopted the proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act, AIA) on 21 April 2021.

14278/21

TREE.2.B RB/ek 1  
LIMITE EN



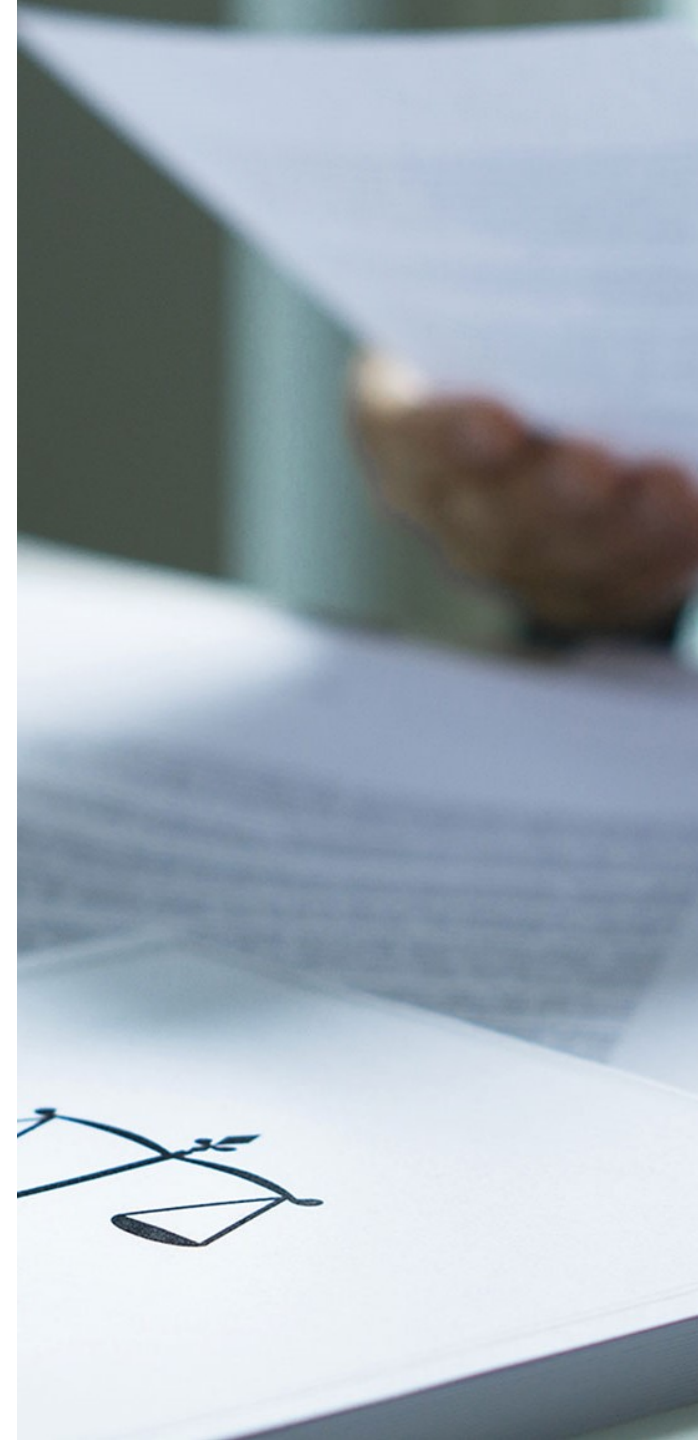
# RISIKO FOR AI BIAS GENNEM DATA

Unfair behandling af minoriteter og etniske grupper

- Udvælgelsesbias - data udvalgt ud fra begrænsede befolkningsgrupper
- Labeling bias - subjektiv og diskriminerende data labeling
- Kulturel bias - data repræsenterer bestemt kultur og værdier
- Dataindsamlingsbias - data er indsamlet ud fra fejlagtige metoder
- Tidsmæssig bias - træning data er ikke opdateret

# AI FORORDNINGEN OG DATA

- Artikel 10 indeholder en række overordnede krav til data.
- Kravene, som gælder for de højrisiko-AI-systemer, der gør brug af teknikker, som omfatter træning af modeller med data,
  - Trænings-, validerings- og prøvningsdatasæt er underlagt passende datastyrings- og dataforvaltningspraksisser (beskrevet på næste slide).
  - Trænings-, validerings- og prøvningsdatasæt skal være relevante, tilstrækkeligt repræsentative, gennemgået for fejl og så fuldstændige som muligt henset til formålet med systemet.
  - De har de tilstrækkelige, statistiske egenskaber, herunder, hvis det er relevant, med hensyn til de personer eller grupper af personer, på hvilke højrisiko-AI-systemet skal anvendes (for de enkelte datasæt eller for en kombination heraf).
  - I trænings-, validerings- og prøvningsdatasæt tages i nødvendigt omfang og i lyset af deres tilsigtede formål, hensyn til specifikke geografiske, adfærdsmæssige eller funktionelle forhold, inden for hvilken højrisiko-AI-systemet skal anvendes.



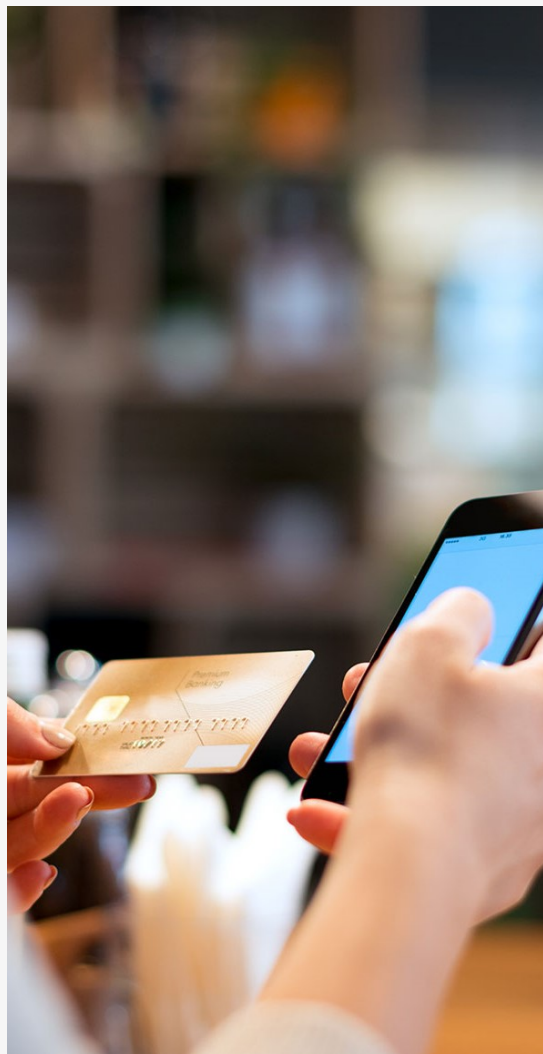


# DATA OG DATASTYRING

- Datastyrings og -forvaltningspraksisser omfatter især

- de relevante valg med hensyn til udformning,
- dataindsamlingsprocesser og gennemsigtighed i relation til det oprindelige formål med indsamlingen af data,
- relevant dataforberedelsesbehandling såsom annotation, mærkning, rensning, opdaterende berigning og aggregering,
- formuleringen af antagelser, navnlig med hensyn til de oplysninger, som dataene skal måle og repræsentere,
- vurdering af tilgængeligheden, mængden og egnetheden af de datasæt, der er nødvendige,
- undersøgelse af mulig forudindtaget, som sandsynligvis vil påvirke fysiske personers helbred og sikkerhed, have negativ indvirkning på grundlæggende rettigheder eller føre til ulovlig forskelsbehandling og passende foranstaltninger til at opdage, forebygge og afbøde mulige bias,
- passende foranstaltninger til at opdage, forebygge og afbøde mulige bias, og
- kortlægning af eventuelle datamangler eller datautilstrækkeligheder, og hvordan de kan afhjælpes.

# HVOR BEHANDLES PERSONOPLYSNINGER I AI-SYSTEMER?



- Træningsdata
  - Data, der anvendes til træning af et AI-system ved hjælp af tilpasning af systemets lærbare parametre, herunder vægte i et neuralt netværk.
- Valideringsdata
  - Data, der anvendes til at foretage en evaluering af det trænede AI-system og til at justere blandt andet systemets ikkelærbare parametre og dets læringsproces for at forhindre overtilpasning, idet valideringsdatasættet kan være et separat datasæt eller en del af træningsdatasættet med enten en fast eller en variabel opdeling.
- Testdata
  - Data, der anvendes til en evaluering/test af det trænede og validerede AI-system for at bekræfte systemets forventede ydeevne, inden det bringes i omsætning eller tages i brug.
- Data indlejret i modellen?
- Inputdata i anvendelse
- Outputdata (herunder nye udledte data, f.eks. diagnoser)

# GDPR OG DATA

- Behandling af personoplysninger skal ske til udtrykkeligt og legitimt formål
- Behandlingen skal være proportionel - egnet, nødvendig og hensigtsmæssig
- Der skal være behandlingsgrundlag
- Oplysningspligt skal overholdes
- Monitorering og overvågning
- Anonyme, pseudonyme og syntetiske data



# IP OG DATA

## RELEVANT LOVGIVNING

- **Ophavsretsloven**
  - Tekst, billeder, figurer, musik, videoer mv.
- **Varemærkeloven**
  - Navn, logoer, ikoner mv.
- **Markedsføringslovens bestemmelser om forretningskendetegn**
  - Navn, figurer, ikoner mv.
- **Lov om forretningshemmeligheder**
- **Lovbestemt tavshedspligt (konkurrenceloven, straffeloven, forvaltningsloven mv.)**
- **Personlighedsretten**
  - Retten til eget navn, billede, stemme osv.

## INPUT (DATA)

- Hvad er værktøjet "trænet" på?
- Må slutbrugere "fodre" værktøjet med tekst, billeder mv.?

## OUTPUT (RESULTATER)

- Er output beskyttet af IP-rettigheder?
  - Hvem har i givet fald rettighederne?
- Kan brug af output krænke 3M's rettigheder?

## RISICI

- Juridiske
- Kommercielle
- Økonomiske

# DE DATAETISKE PRINCIPPER

- Dataetik handler om at behandle andre menneskers data med respekt, ordentlighed og ansvarlighed i en tid, hvor data spiller en rolle i alle samfundets facetter.

## De ti dataetiske principper



Velfærd



Værdighed



Privatliv



Selvbestemmelse



Lighed



Frihed



Retssikkerhed



Gennemsigtighed



Sikkerhed



Ansvarlighed



# ANSVARLIG BRUG AF DATA



Dataetisk tankegang, ansvaret for dataetik og dataetisk politik



Data og datakilder



Brug af AI og virksomhedens kunder



Implementering, overvågning og evaluering

# SPØRGSMÅL?



**Milena Anguelova Krogsgaard**  
ADVOKAT, PARTNER  
M 25 10 05 57  
@ MAN@POULSCHMITH.DK

