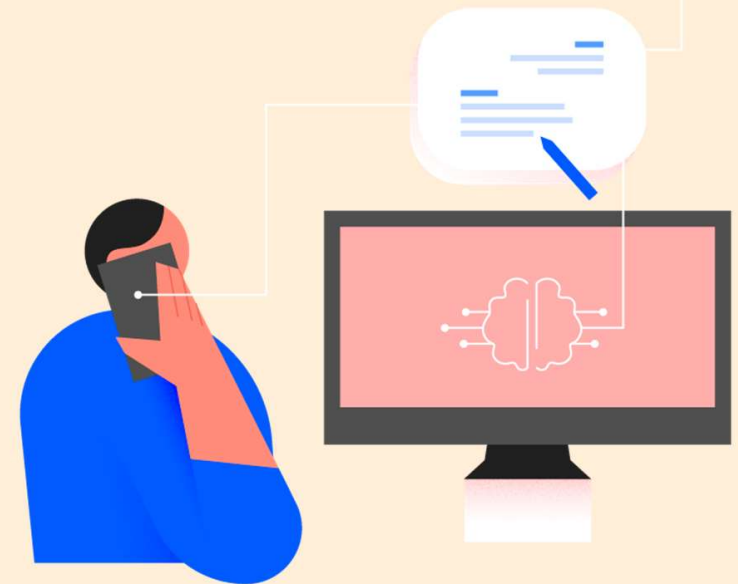


# Generativ AI og offentlige myndigheder

– Pålidelighed og dataetik

5. februar 2025



# Dataetisk Råds projekt

-  Brugsscenarier
-  Dataetiske udfordringer
-  Eksisterende retningslinjer
-  Risiko for fejl
-  Dataetisk Råds anbefalinger



# Offentlige myndigheder bruger GenAI til...

- Idegeneration
- Tekstredigering
- Udkast til brugerdefineret tekst
- Informationssøgning
- Opsummering af materiale
- Vurdering af sagsforhold



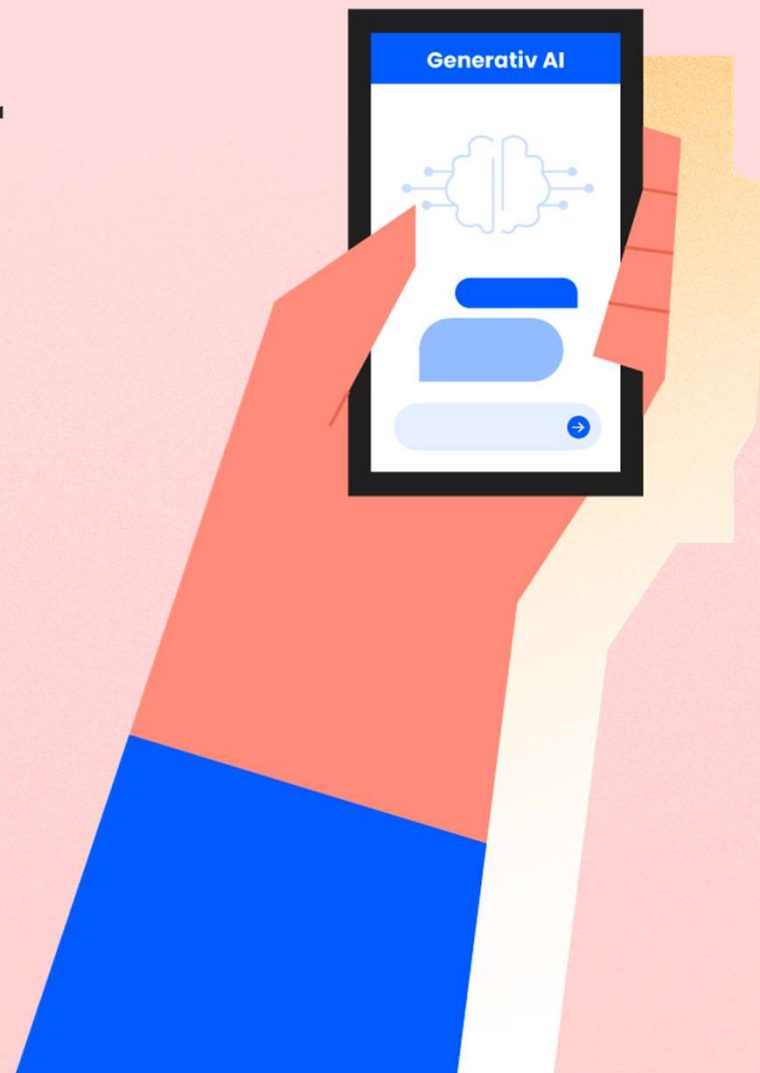
# Dataetiske udfordringer ved GenAI

Udfordringer i anvendelsen af GenAI	Udfordringer med potentiale for misbrug af GenAI	Udfordringer ved træning og bredere effekter af GenAI
Fejl og hallucinationer	Misinformation	Brud på intellektuel ejendomsret
Algoritmisk bias	Deep fakes	Miljøpåvirkning
Læk af private data	Manipulation	Udbytende arbejdspraksisser
Diffus ansvarlighed	Spredning af farlig information	Forstærkede tech-oligopoler
Uigennemsigthed	Forklædning af AI-genereret indhold	Forskydninger på arbejdsmarkedet
		Eksistentielle risici ved AGI
		Mistillid

# Eksisterende retningslinjer

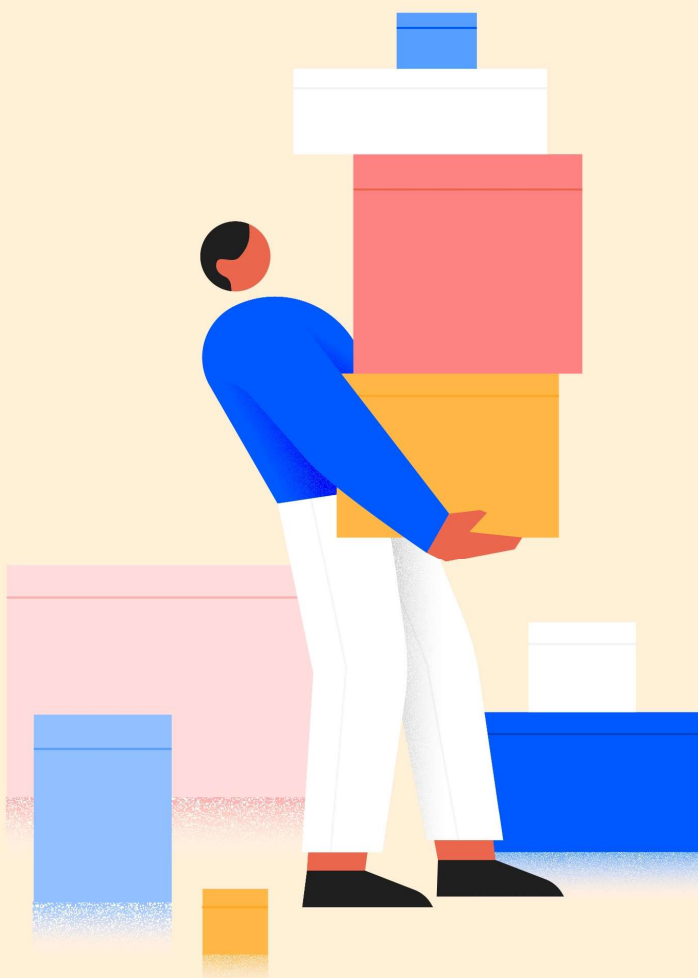
## Konsensus om...

- Deklaration af brug
- Menneskelig kvalitetskontrol
- Beskyttelse af følsomme data i prompting
- Identifikation af bias
- Forsigtighed ved brug af ADM



# Risiko for fejl

<b>Idegeneration</b>	Uegnede ideer
<b>Tekstredigering</b>	Overser eller indfører fejl
<b>Udkast til brugerdefineret tekst</b>	Uegnet eller faktisk forkert tekst
<b>Informationssøgning</b>	Giver forkert information/undlader at give information
<b>Opsummering af materiale</b>	Opfinder information/udelader information/uegnet tekst
<b>Idegeneration</b>	Uegnede ideer

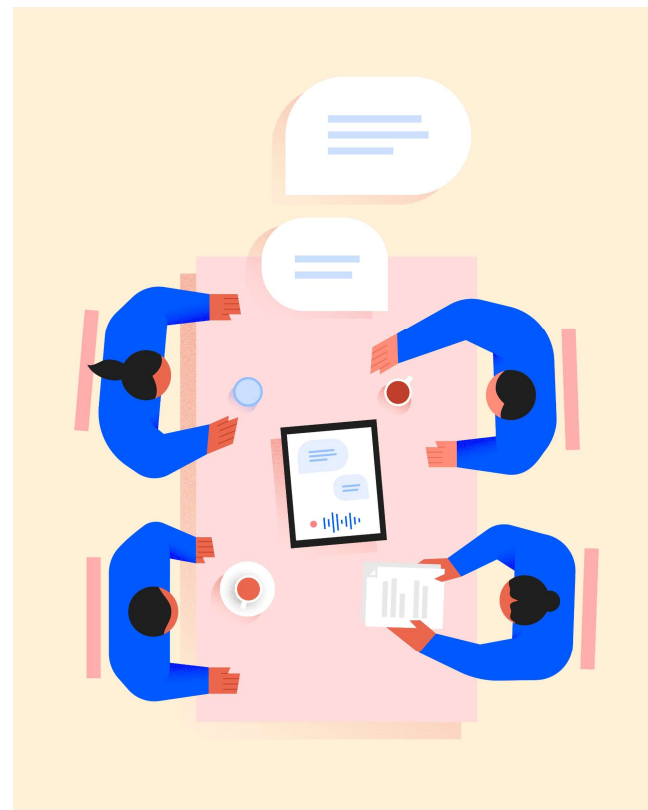


# Dataetisk fejlrisikoprofil

- Hvor mange og hvilke fejl laver systemet?
- Hvor mange fejl fanges af mennesker?
- Hvilken indflydelse har fejl?
- Hvilke beslutninger påvirkes?

# Risiko for fejl

- Hvordan bør offentlige myndigheder forholde sig til risikoen for, at GenAI laver fejl?





# Dataetisk Råds anbefalinger

- 1 Offentlige myndigheder** bør begynde eventuel udforskning og anvendelse af GenAI med brugsscenarier, hvor anvendelsen har lav risikoprofil, og inddrage interessenter i udvikling og evaluering af systemer.
- 2 Offentlige myndigheder** bør skelne mellem GenAI-systemer med høj og lav risikoprofil, og kun bruge GenAI, når myndigheden har indført tilstrækkelige mitigerende tiltag til at begrænse risikoen for fejl.
- 3 Offentlige myndigheder** bør systematisk vurdere GenAI-systemers samlede risikoprofil for fejl i hvert enkelt brugsscenarie.

# Dataetisk Råds anbefalinger

4

**Offentlige myndigheder** bør gennemføre mere omfattende kvalitetssikring af et GenAI-system, jo højere systemets risikoprofil er.

5

**Offentlige myndigheder** bør oplyse, hvordan de anvender menneskelig kvalitetskontrol af GenAI, og tildele relevante medarbejdere de nødvendige ressourcer.

6

**Offentlige myndigheder** bør indføre nødvendige forholdsregler til at identificere, dokumentere, oplyse og korrigere fejl, inden de tager et GenAI-system i brug.

# Dataetisk Råds anbefalinger

7

**Regeringen** bør udvikle fælles retningslinjer for offentlige myndigheders anvendelse af GenAI, som slår fast hvordan myndigheder skal håndtere risikoen for fejl.

8

**Regeringen** bør oprette et nationalt, offentligt register over GenAI-systemer, som anvendes af offentlige myndigheder.

# Dataetisk Råds fokus i 2025

TEMAER

PROJEKTER

## AI & VELFÆRD

AI i sundheds- og velfærdsteknologier

AI som beslutningsstøtte i sociale sager

Efterårsprojekt

## AI & ADFÆRD

Online aldersverificering

Politiets brug af ansigtsgenkendelse

Efterårsprojekt

## DATAETISKE VÆRKTØJER

Samtalespil om brugen af medarbejderdata

Opdaterede dataetiske principper

Efterårsprojekt

Høringssvar, kronikker, events, Folkemødet, Ungdommens Folkemøde