



# Hvordan påvirker kunstig intelligens (KI) internrevisjon i offentlige virksomheter?

Nettverk for offentlig virksomhet - 22. november 2023

# Agenda 22. november 2023

Tidspunkt	Tema	Hvem
08.30 – 09.00	Lett frokostservering	
09.00 – 09.05	Velkommen	Sissel Misund Berg
09.05 – 09.45	KI og risikobilde	Henrik Slettene, Inmeta Consulting
09.45 – 10.00	Pause	
10.00 – 10.40	Veiledning for ansvarlig utvikling og bruk av kunstig intelligens i offentlig sektor»	Jens Andresen Osberg, Digdir
10.40 – 11.00	Paneldiskusjon og erfaringsdeling	Sissel Misund Berg



# Vi ønsker å diskutere

- Hva er KI?
- Hva er de største risikoene som vi som internrevisorer være oppmerksomme på?
- Hvordan sikre en forsvarlig bruk av KI?
- Hvilke rammeverk finnes på området?





# KI og risikobildet

**Henrik Slettene**

22. november 2023

KI – IIA nettverk for offentlig virksomhet

inmeta

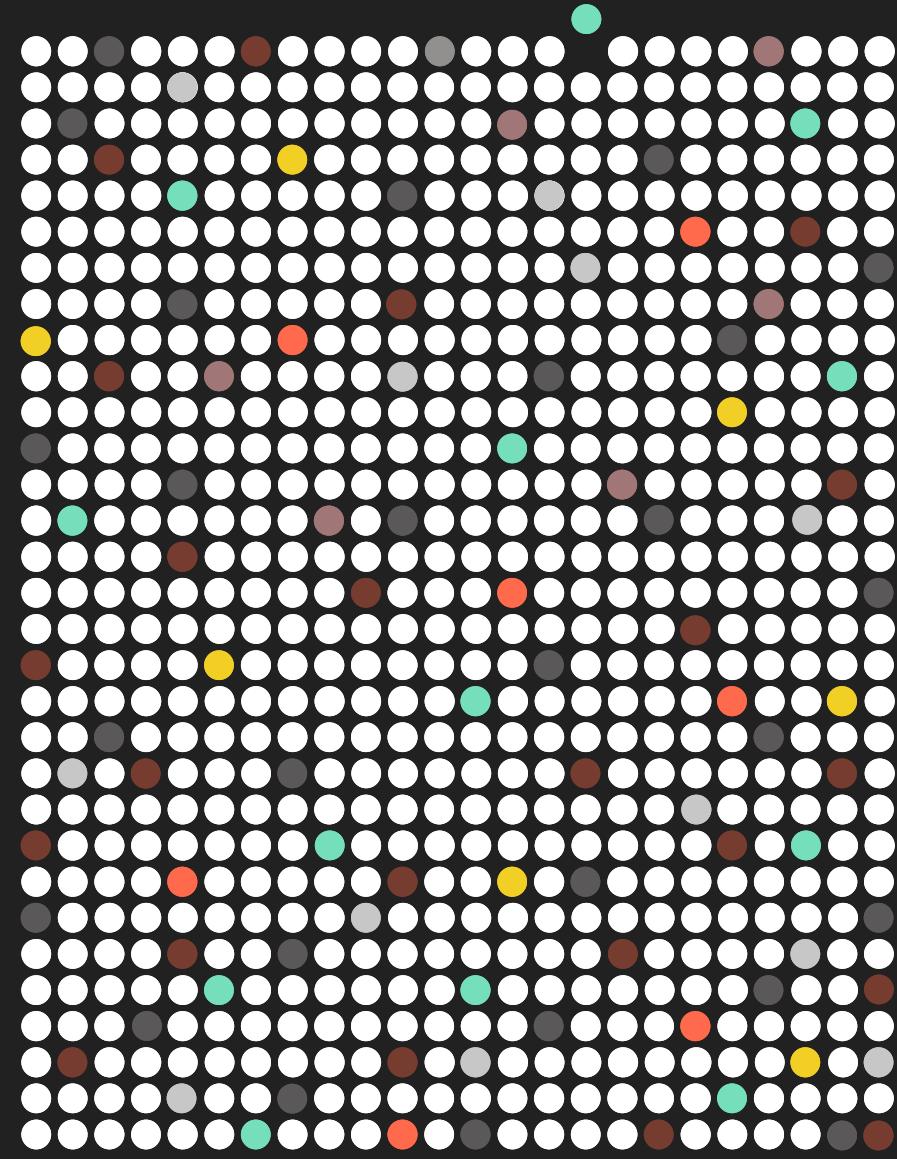
# Artificial Intelligence

Presentation to Internrevisorforeningen

Henrik Slettene

Nov 22nd, 2023

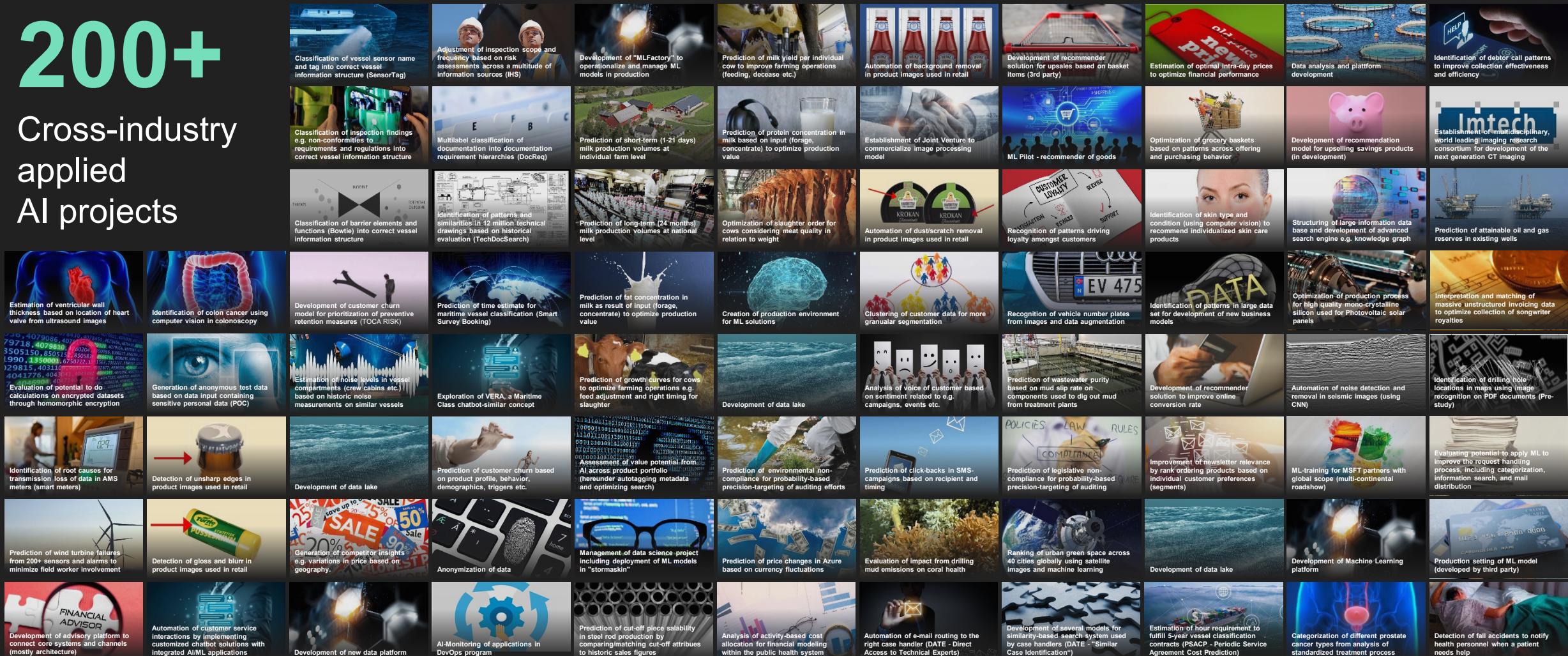
*Guaranteed free from images of brains and robots*



# The most dedicated AI consulting partner in Norway

**200+**

Cross-industry  
applied  
AI projects



## ABOUT ARTIFICIAL INTELLIGENCE

# Artificial intelligence mimics and enhances human characteristics and senses



The use of artificial intelligence means that humans and machines complement each other and use each other's characteristics in the best possible way.



Where humans are best at intuition, creativity, emotions and imagination, machines are best at processing large amounts of information, finding complicated patterns, making advanced calculations and being objective.



Artificial intelligence is often developed using machine learning techniques, where we train machines using historical data.



Artificial intelligence is specialized solutions that are trained to solve a specific problem.

# AI is most commonly developed by training machines

## Explicit programming

*“The price of the next house sold is the average of the 3 last sold within a range of +/- 10% of the size”*

- Reproducing “your” understanding of the system
- No adaption to change in system or surroundings
- Will never be better than the programmer’s instructions
- By far most common way of using machines

## Machine Learning

*“Here is data about the last 1000 houses sold. Understand what is relevant and model the system. I will supervise you.”*

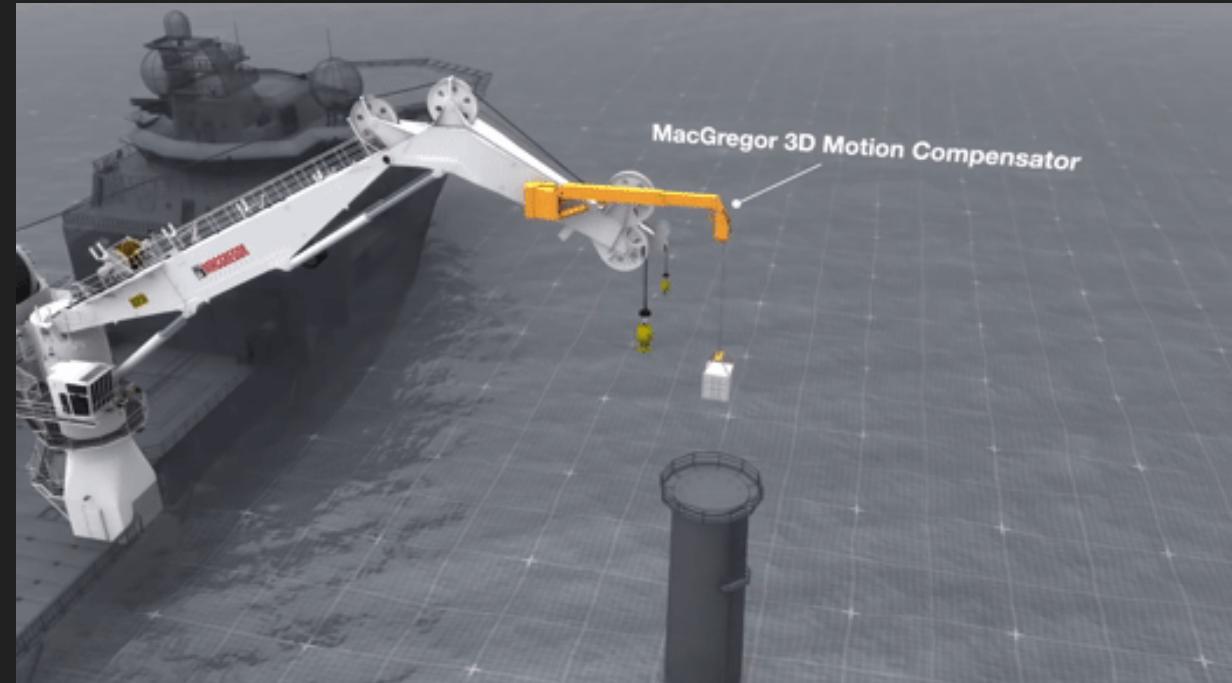
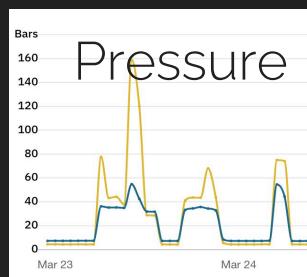
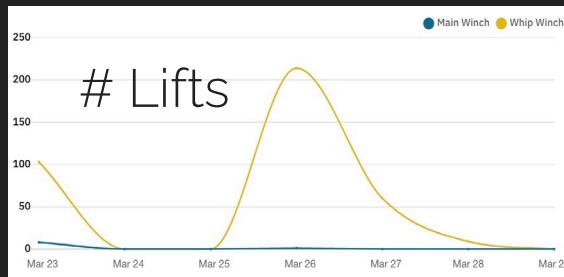
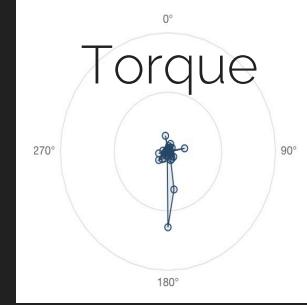
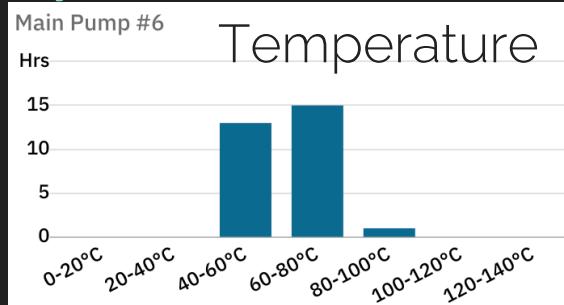
- Learns by observing the system
- Active learning constantly adapts and improves based on feedback from the system
- Will outperform its supervisor over time

# Machine Learning is by far the most used way to develop AI



# AI is more than ChatGPT!

54 – 196 sensors on each system,  
solution deployed on crane control  
system

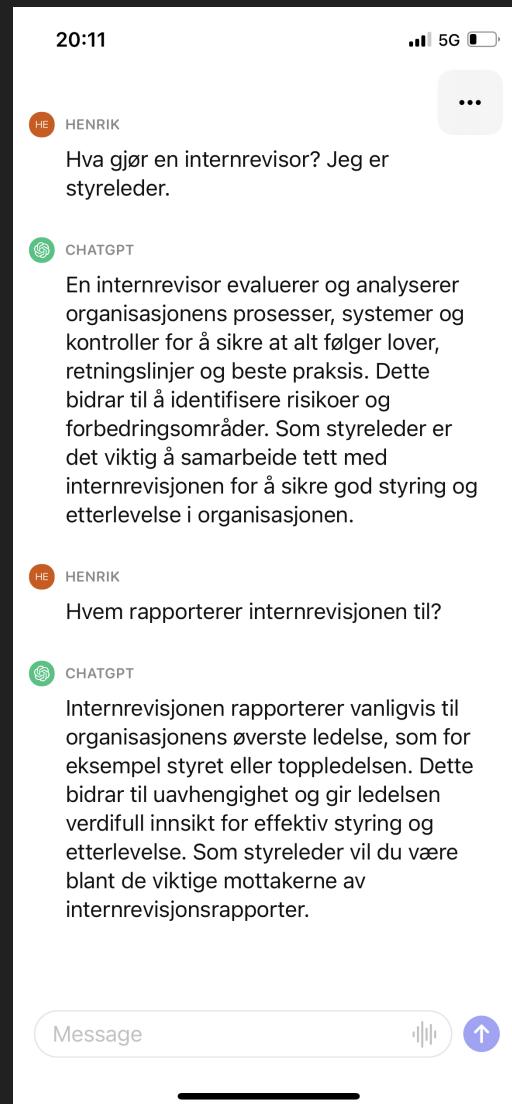
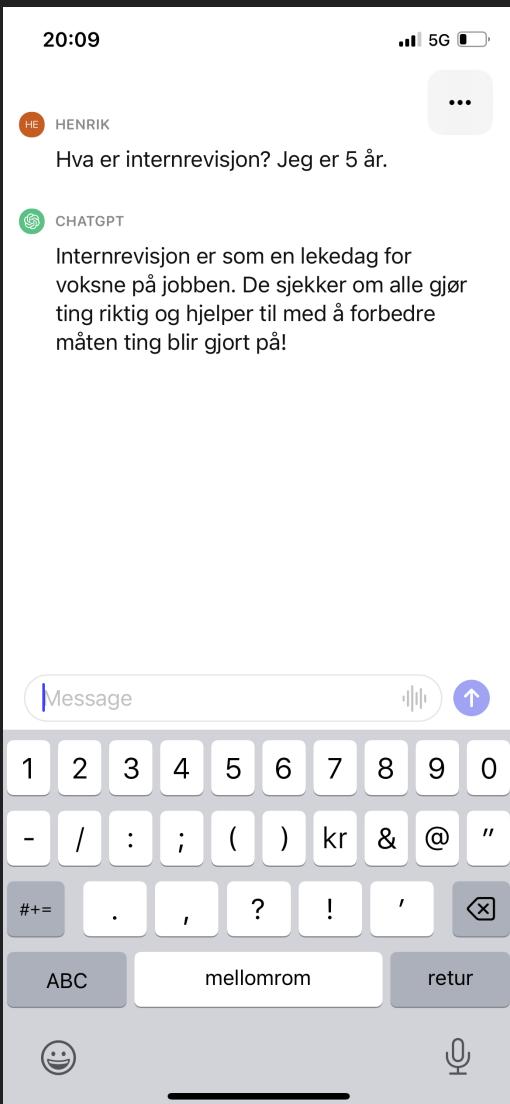


- 4 anomaly detection Machine Learning models running in monitoring system
- Objective is to detect and correct problems before they occur and perform maintenance based on condition



inmeta

# What Language Models can do



# **Advokater tester Chat GPT: – Ofte er det ren fantasi**

Da Wikborg Rein testet Chat GPT, fikk de dommer, personer og rettskilder som aldri har eksistert. Advokatene er likevel klare på at kunstig intelligens vil overta mange oppgaver.

3 MIN | PUBLISERT: 28.07.23 – 21.24 | OPPDATERT: EN MÅNED SIDEN



# Is AI making things up?

The cat sat on the

Hat 2%

Floor 20%

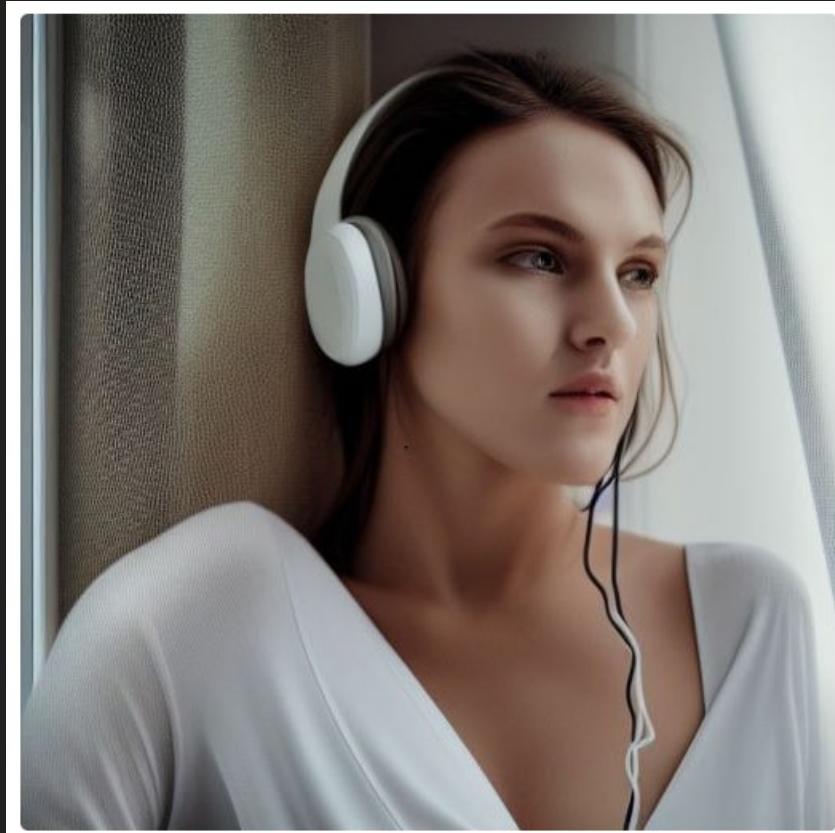
Mat 40%

Table 10%

Box 5%

Umbrella 0%

# It is not only language



«Gorgeous young Swiss girl sitting by window with headphones on, wearing white translucent shirt, beach blonde hair, photorealistic»

**“Jeg har så lyst på  
verdens beste”**



# What Language Models can do

Summariz  
e a text

Translate  
between  
language  
s

Generate  
text

Classify  
text

Answer  
questions

Perform  
semantic  
search

# Will AI ever be “intelligent”?

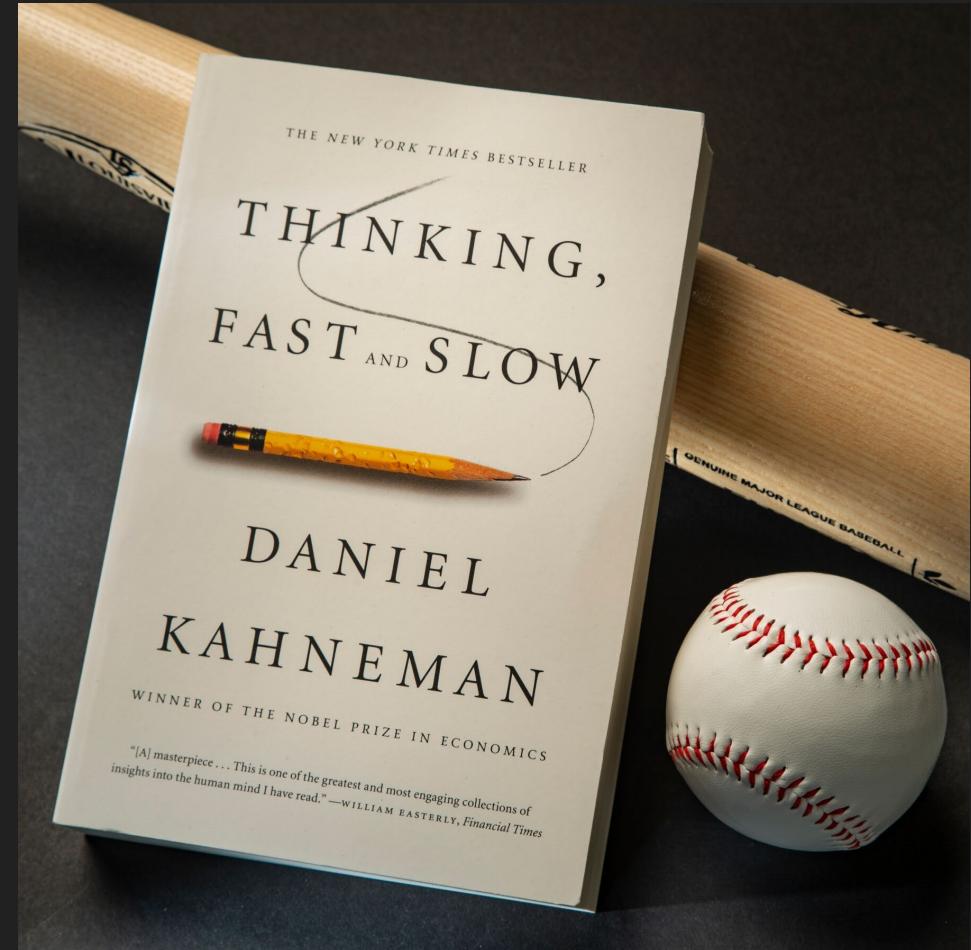
Understanding of how the world works?

Ability to solve (mathematical) problems?

Planning?

Reasoning / Forward thinking?

A plan, a purpose, a desire, or an objective?



# How many was not caught?



# Tre studenter utestengt fra NTNU etter ChatGPT-juks

**Tre studenter er utestengt fra NTNU etter å ha brukt kunstig intelligens (KI) for å jukse på eksamen.**

Share



ARTIFICIAL INTELLIGENCE

Share



# ChatGPT displays ‘systematic’ Left-wing bias, say researchers

GARETH CORFIELD, JAMES TITCOMB AND SARAH KNAPTON

CHATGPT has been accused of a “systematic” Left-wing bias that reflects the positions of Labour and US Democrat politicians, raising concerns over its use in policymaking and education.

Researchers found that the chatbot’s responses generally favour Left-leaning political beliefs, despite efforts to reduce bias.

The study, led by academics at the University of East Anglia (UEA), claims to be the first large-scale research into ChatGPT’s political slant.

# Could we – and should we?

Could AI impact how we think about the world?

Bias – or just replicating unwanted behaviour?

How do we protect against poor handcrafting?

Trust of source and origin?

More than 70% of all new sites has fraud as prime objective?

How do we know there is AI in it?

Machine unlearning and LLM Engineering?

Will EU fall behind in its strive for ethical use?

Explainable behaviour?

Consequences of 90% of all content being machine generated?

How can we make sure AI is doing what we want it to do and nothing else?

# **When should you be really worried about AI?**

..when ChatGPT asks you:

*“If killer robots are chasing you – where would you hide?”*



# Takk for meg



**Henrik Slettene**

Director, Nordic AI Center of Excellence

Inmeta Consulting AS, part of Crayon

[Henrik.slettene@inmeta.no](mailto:Henrik.slettene@inmeta.no)



# Hvordan virker ChatGPT og Digidirs veiledning for ansvarlig utvikling og bruk av KI

Jens Andresen Osberg

22. november 2023

KI – IIA Nettverk for offentlig virksomhet

# Hva skal jeg snakke om?



Hvordan virker teknologi  
som ChatGPT?



Hvorfor kan ChatGPT så  
mange ting?



Kan vi lære ChatGPT  
folkeskikk?



Digdirs retningslinjer for  
bruk av generativ KI i  
offentlig sektor

# Hvordan virker ChatGPT og Digdirs veiledning for ansvarlig utvikling og bruk av KI

Jens Andresen Osberg  
Seniorrådgiver, jurist og teknolog i Digitaliseringsdirektoratet



Hvordan virker teknologi  
som ChatGPT?

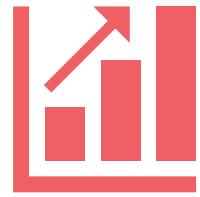
«Jeg elsker [ ? ]»



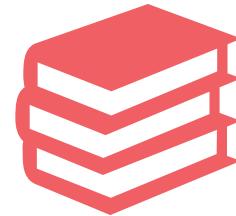
# Hva kan si noe om denne sannsynligheten?



Hvilke ord du brukte  
sist?

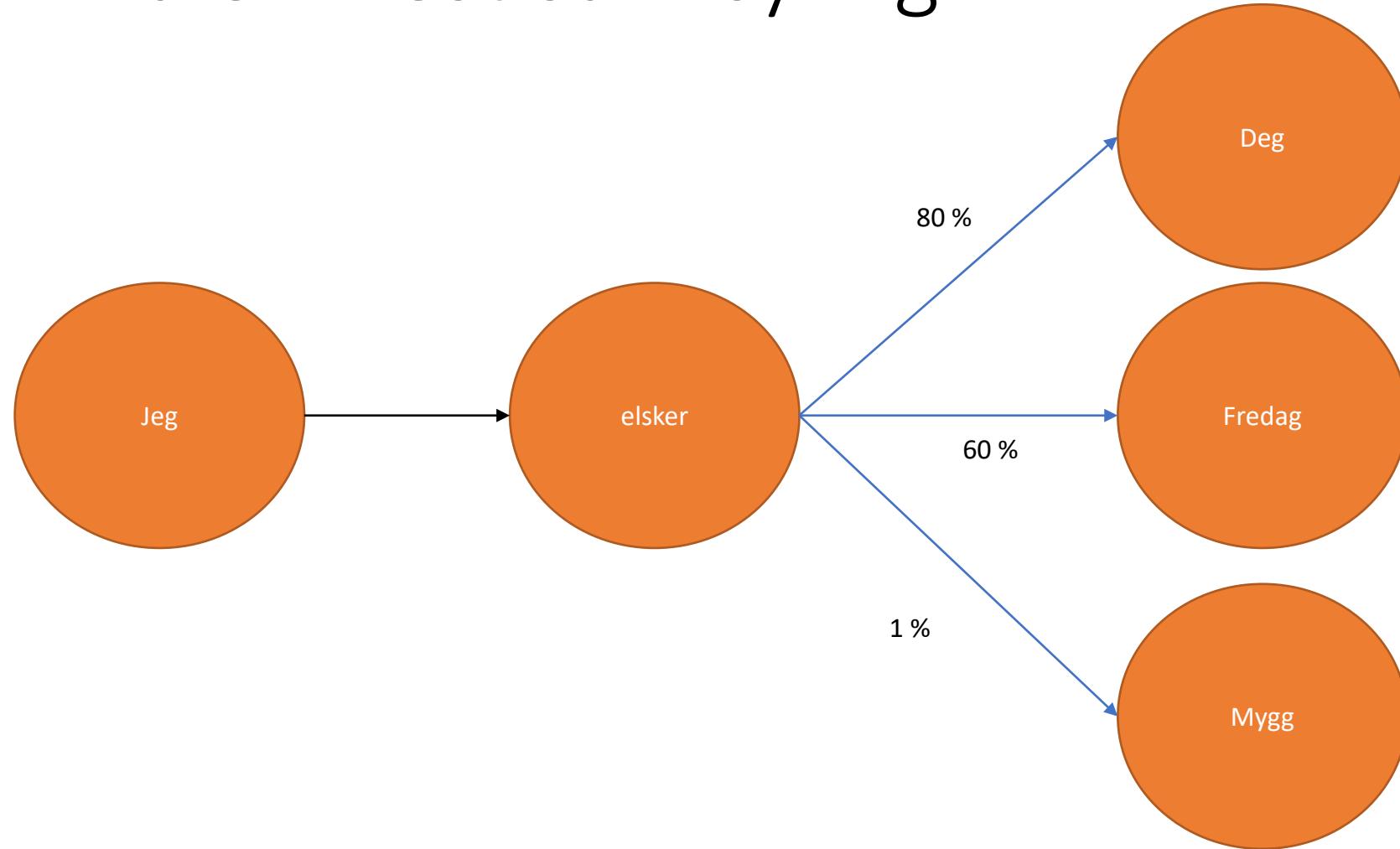


Hvilke ord som er  
trendende?



Hvor ofte et ord  
brukes etter et annet?

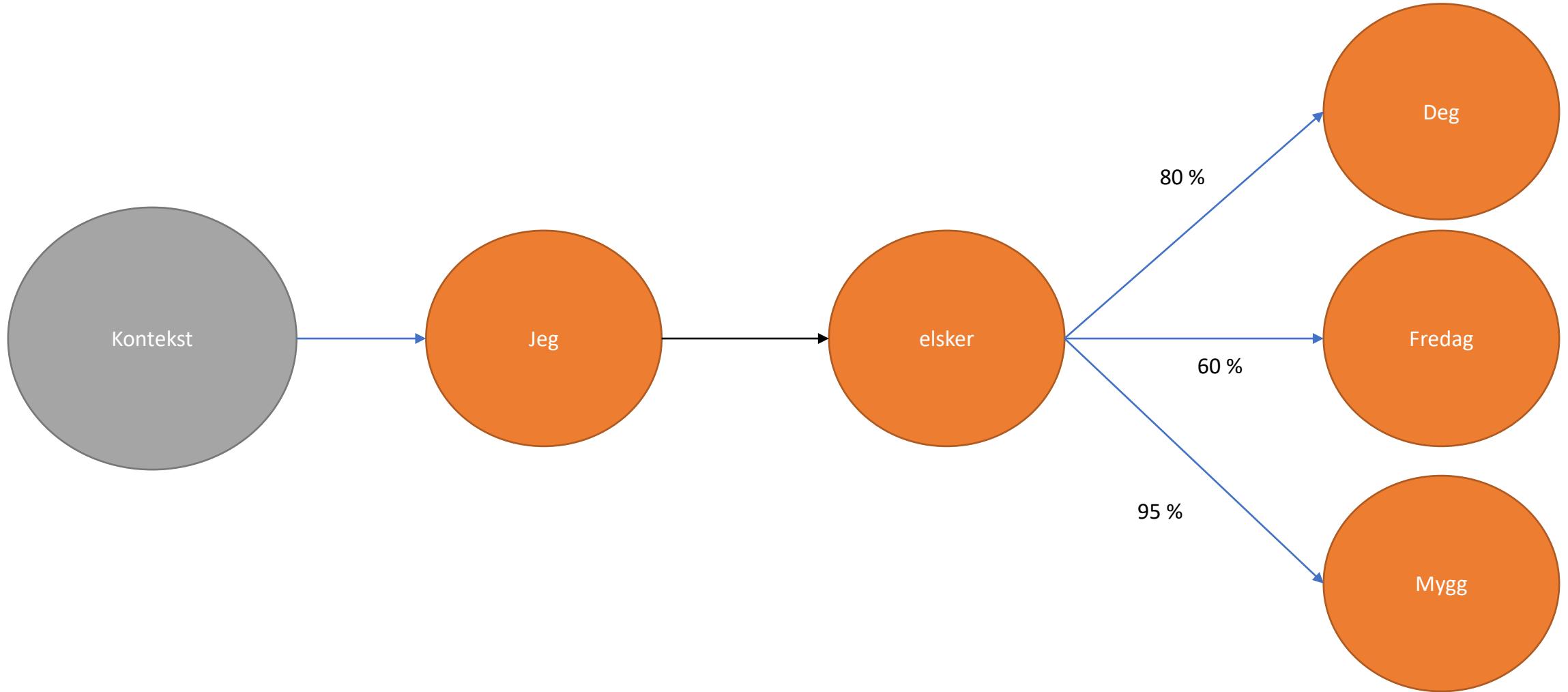
# Hva er mest sannsynlig?



# Hva med mer kontekst?



# Hva er mest sannsynlig?



# Hva er nytt med ChatGPT?



**ChatGPT predikerer hva som skal være  
det neste sannsynlige ordet**

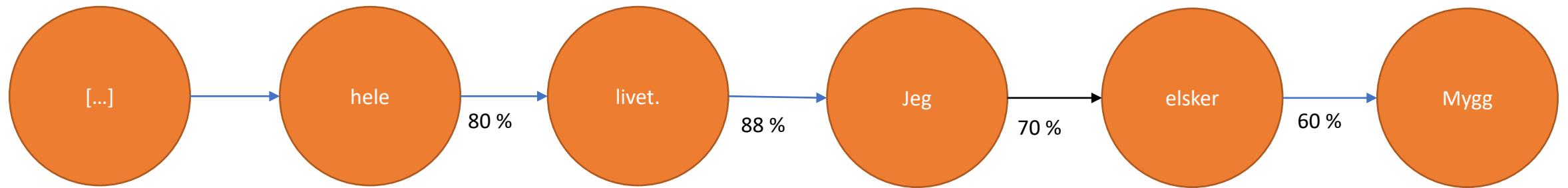
Kontekst har stor betydning for dette



**Men hva er da nytt med ChatGPT?**

Vi har jo hurtigtastaturet på mobilen  
Det forstår da kontekst, gjør det ikke?

# Utfordringen



Kontekst gir stor kostnad knyttet til utregning

# Løsningen

- Generativ Pre-trained Transformer
- Transformer bygger på en mekanisme kalt «attention»
- «Attention» gjør det mye lettere å forstå ord i sin kontekst

## Attention Is All You Need

Ashish Vaswani\*  
Google Brain  
avaswani@google.com

Noam Shazeer\*  
Google Brain  
noam@google.com

Niki Parmar\*  
Google Research  
nikip@google.com

Jakob Uszkoreit\*  
Google Research  
usz@google.com

Llion Jones\*  
Google Research  
llion@google.com

Aidan N. Gomez\* †  
University of Toronto  
aidan@cs.toronto.edu

Łukasz Kaiser\*  
Google Brain  
lukaszkaiser@google.com

Illia Polosukhin\* ‡  
illia.polosukhin@gmail.com

### Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

# Oppsummert



ChatGPT, Bard, Bing's (Sydney) bygger på store språkmodeller

De gjetter hvilket ord som kommer



Bygger på en mekanisme kalt «attention»

Gjør det lettere å analysere ord i kontekst fordi det er «prosesseringsmessig billigere»



Forstå kontekst ved gjetting av ord gir imponerende resultater



Hvorfor kan ChatGPT så  
mange ting?

# Hvorfor kan ChatGPT så mange ting?

Hvis store språkmodeller  
bare gjetter det som er  
sannsynlig?



Hvorfor har den da så  
mye faktakunnskap?



# «Bonuskunnskap» i språkmodellen

Stort  
ordforåd

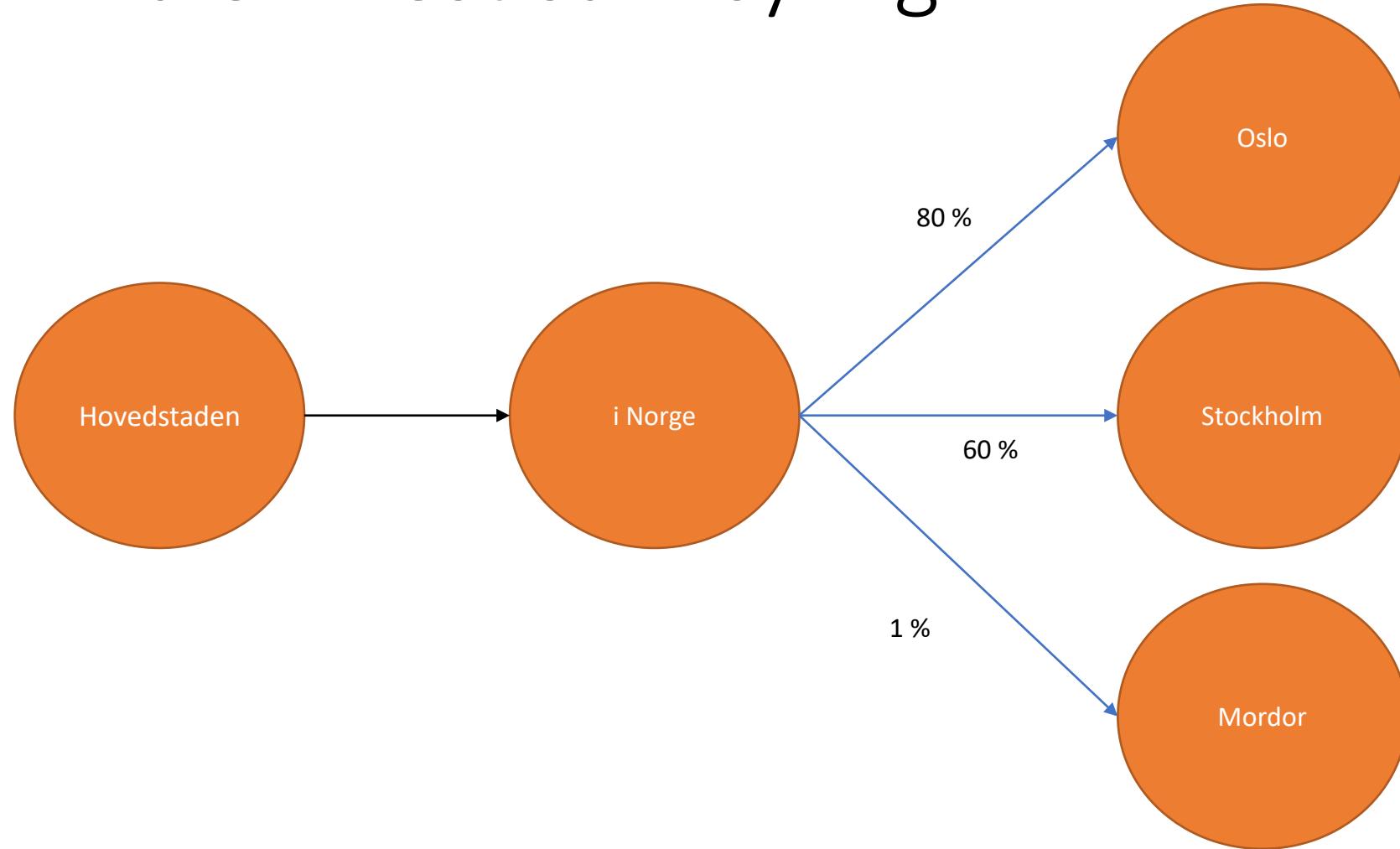


Predikerer  
hva som er  
sannsynlig

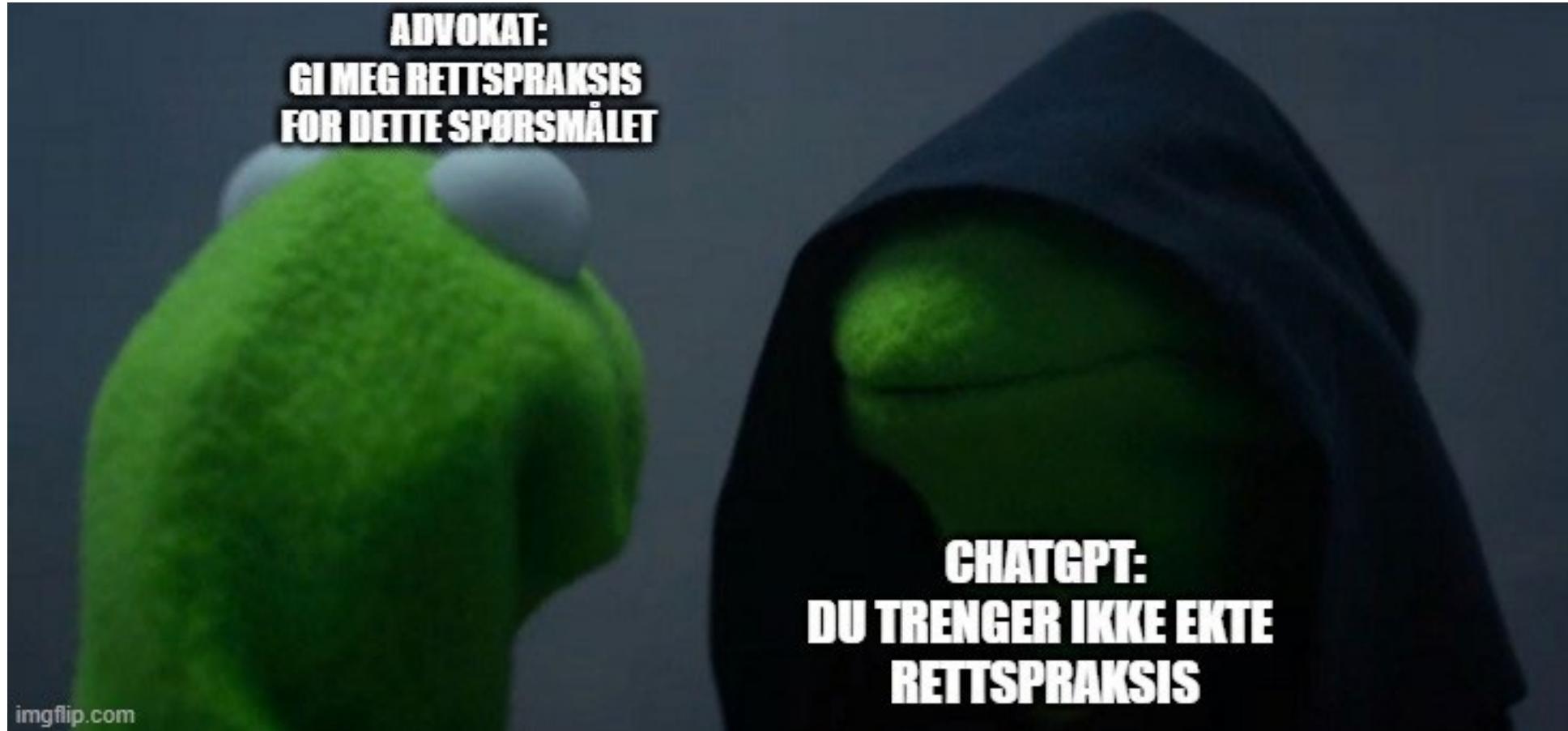


«Illusjonen  
av  
kunnskap»

# Hva er mest sannsynlig?



# En grunnmodell kan «hallusinere»





Hvordan lærer vi ChatGPT  
«folkeskikk»?

Det er ikke all tekst vi vil ha



**Midjourney-prompt:**

"Tired man yelling at people at  
a party"

De tre H-ene



---

Helpful

---

Harmless

---

Honest

# Kan vi justere den store språkmodellen?

Vi har nå en stor språkmodell som predikerer det neste sannsynlige ordet

Kan vi videre justere på den store språkmodellen?

Lære den «folkeskikk» på samme måte som vi mennesker lærer hva vi kan si og ikke?

# Vi kan finjustere grunnmodellen

Finjustering er en viktig del av utviklingen av store språkmodeller



To tilnærmingar

Supervised fine tuning

Reinforcement learning human feedback

Veiledning for ansvarlig utvikling  
og bruk av kunstig intelligens i  
offentlig sektor

# Noen utgangspunkter for ansvarlig bruk



## Eksisterende regulering

Mange lover som setter rammer  
for KI

Gjelder uavhengig av teknologien  
som brukes



## Kommende regulering

KI-forordningen kommer snart  
Et regelverk tilpasset KI

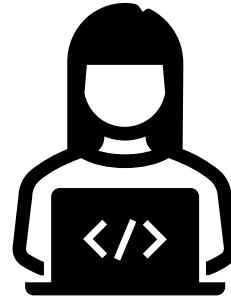


## Ytterligere regulering?

Mange tar til ordet for «KI må  
reguleres»

Lite nyttig å snakke om  
«regulering» helt generelt – vi  
må være mer presise

# Konkrete rammer er et tydelig behov

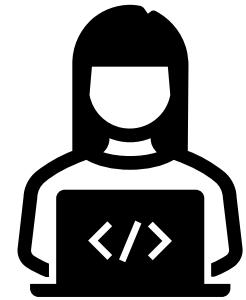




Forvaltningsloven § 25  
GDPR artikkel 15



«Rett til en  
begrunnelse/forklaring»





Forvaltningsloven § 25  
GDPR artikkel 15

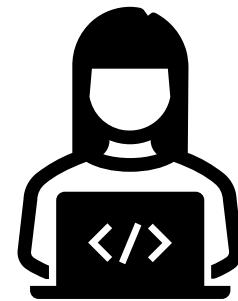


«Rett til en  
begrunnelse/forklaring»

Regelverket tolkes



«Informasjon som gjør et  
vedkommende kan slå seg til ro  
med eller utfordre vedtaket»





Forvaltningsloven § 25  
GDPR artikkel 15

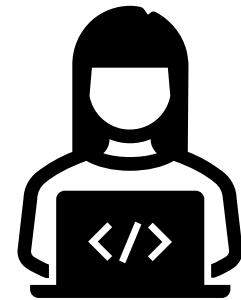


«Rett til en  
begrunnelse/forklaring»

Regelverket tolkes



«Informasjon som gjør et  
vedkommende kan slå seg til ro  
med eller utfordre vedtaket»





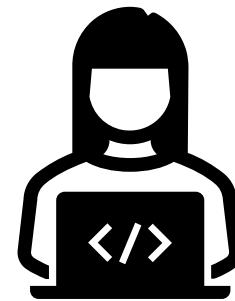
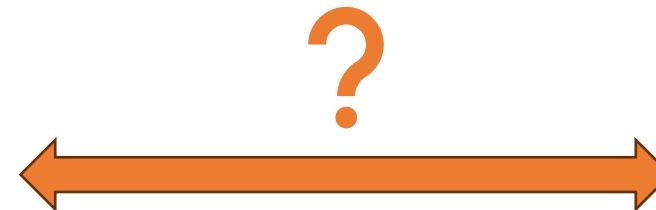
Forvaltningsloven § 25  
GDPR artikkel 15



«Rett til en  
begrunnelse/forklaring»



«Informasjon som gjør et  
vedkommende kan slå seg til ro  
med eller utfordre vedtaket»

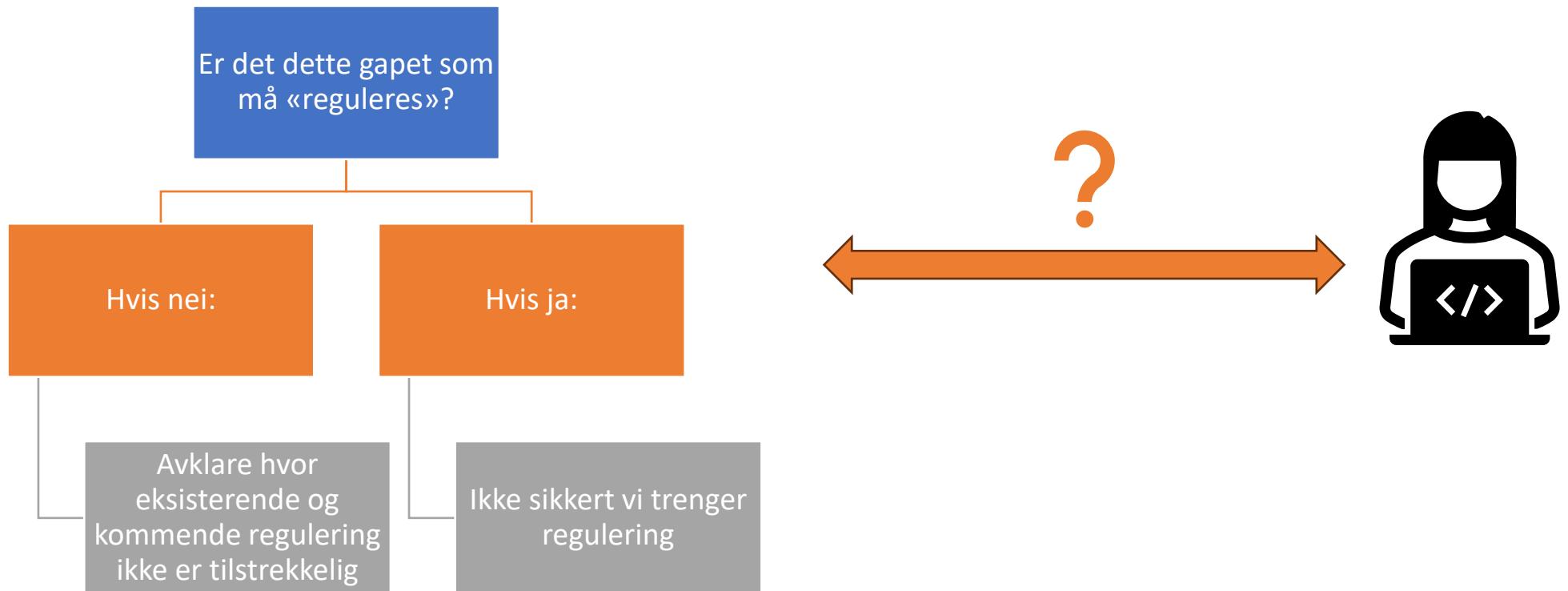


Hvilken XAI-metode skal jeg  
bruke?



XAI: Ulike tekniske tilnærminger  
for å forklare KI

# Er det dette siktet til med «regulering»?



# Vi har gitt ut veiledning



## Veileddning for ansvarlig utvikling og bruk av kunstig intelligens i offentlig sektor

Veileddningen er i «åpen beta». I tiden fremover ønsker vi tilbakemeldinger og innspill til innholdet som presenteres her. Det betyr at rådene du finner i veileddningen kan bli endret. Har du innspill til oss vil vi gjerne høre fra deg! Ta kontakt med oss på [datalandsbyen.norge.no](http://datalandsbyen.norge.no) så sikrer vi at innholdet her blir best mulig.



### Råd for ansvarlig utvikling og bruk av kunstig intelligens

Mange dokumenter og ressurser gir veileddning om kunstig intelligens, men det kan være vanskelig å finne frem. Vi gir konkrete råd for å utvikle og bruke KI i offentlig sektor i samsvar med regelverket.



### Åpenhet og kunstig intelligens

Avgjørelser skal begrunnes, slik at de kan aksepteres eller utfordres. En god begrunnelse forklarer hva som påvirkt resultatet. Dette er en viktig utfordring ved kunstig intelligens.



### Hvordan vurdere risiko ved utvikling og bruk av kunstig intelligens

Å ta i bruk kunstig intelligens handler en avveining av risikoen sterreise og mulige gevinstar. En risikovurdering kan hjelpe oss med å identifisere sterrenes på risikoen.



### Bruk av generativ kunstig intelligens i offentlig sektor

Generativ KI har aldri vært mer tilgjengelig. Som med annen KI, må denne teknologien brukes ansvarlig. Her foreslår vi retningslinjer for å adressere noen av de spesiifikke utfordringene ved generativ KI.



## Bidra til å tette gapet



## Følger eksisterende og kommende regelverk tett



## «Åpen beta»



Innholdet i veiledningen

# Fire hovedelementer



## Råd for ansvarlig utvikling og bruk av kunstig intelligens

Mange dokumenter og ressurser gir veiledning om kunstig intelligens, men det kan være vanskelig å finne frem. Vi gir konkrete råd for å utvikle og bruke KI i offentlig sektor i samsvar med regelverket.



## Åpenhet og kunstig intelligens

Avgjørelser skal begrunnes, slik at de kan aksepteres eller utfordres. En god begrunnelse forklarer hva som påvirket resultatet. Dette er en viktig utfordring ved kunstig intelligens.



## Hvordan vurdere risiko ved utvikling og bruk av kunstig intelligens

Å ta i bruk kunstig intelligens handler en avveining av risikoens størrelse og mulige gevinner. En risikovurdering kan hjelpe oss med å identifisere størrelsen på risikoen.



## Bruk av generativ kunstig intelligens i offentlig sektor

Generativ KI har aldri vært mer tilgjengelig. Som med annen KI, må denne teknologien brukes ansvarlig. Her foreslår vi retningslinjer for å adressere noen av de spesifikke utfordringene ved generativ KI.



# Fire hovedelementer



## Råd for ansvarlig utvikling og bruk av kunstig intelligens

Mange dokumenter og ressurser gir veiledning om kunstig intelligens, men det kan være vanskelig å finne frem. Vi gir konkrete råd for å utvikle og bruke KI i offentlig sektor i samsvar med regelverket.



## Åpenhet og kunstig intelligens

Avgjørelser skal begrunnes, slik at de kan aksepteres eller utfordres. En god begrunnelse forklarer hva som påvirket resultatet. Dette er en viktig utfordring ved kunstig intelligens.



## Hvordan vurdere risiko ved utvikling og bruk av kunstig intelligens

Å ta i bruk kunstig intelligens handler en avveining av risikoens størrelse og mulige gevinner. En risikovurdering kan hjelpe oss med å identifisere størrelsen på risikoen.



## Bruk av generativ kunstig intelligens i offentlig sektor

Generativ KI har aldri vært mer tilgjengelig. Som med annen KI, må denne teknologien brukes ansvarlig. Her foreslår vi retningslinjer for å adressere noen av de spesifikke utfordringene ved generativ KI.



# Fire hovedelementer



## Råd for ansvarlig utvikling og bruk av kunstig intelligens

Mange dokumenter og ressurser gir veiledning om kunstig intelligens, men det kan være vanskelig å finne frem. Vi gir konkrete råd for å utvikle og bruke KI i offentlig sektor i samsvar med regelverket.



## Åpenhet og kunstig intelligens

Avgjørelser skal begrunnes, slik at de kan aksepteres eller utfordres. En god begrunnelse forklarer hva som påvirket resultatet. Dette er en viktig utfordring ved kunstig intelligens.



## Hvordan vurdere risiko ved utvikling og bruk av kunstig intelligens

Å ta i bruk kunstig intelligens handler en avveining av risikoens størrelse og mulige gevinner. En risikovurdering kan hjelpe oss med å identifisere størrelsen på risikoen.



## Bruk av generativ kunstig intelligens i offentlig sektor

Generativ KI har aldri vært mer tilgjengelig. Som med annen KI, må denne teknologien brukes ansvarlig. Her foreslår vi retningslinjer for å adressere noen av de spesifikke utfordringene ved generativ KI.



# Retningslinjer for generativ KI



## Bruk av generativ kunstig intelligens i offentlig sektor

Generativ KI har aldri vært mer tilgjengelig. Som med annen KI, må denne teknologien brukes ansvarlig. Her foreslår vi retningslinjer for å adressere noen av de spesifikke utfordringene ved generativ KI.



Generelt



Tekstgenerering



Bildegenerering



Kodegenerering



Grunnmodeller

# Generelle retningslinjer for generativ KI



**GitHub Copilot**



# Generering av tekst



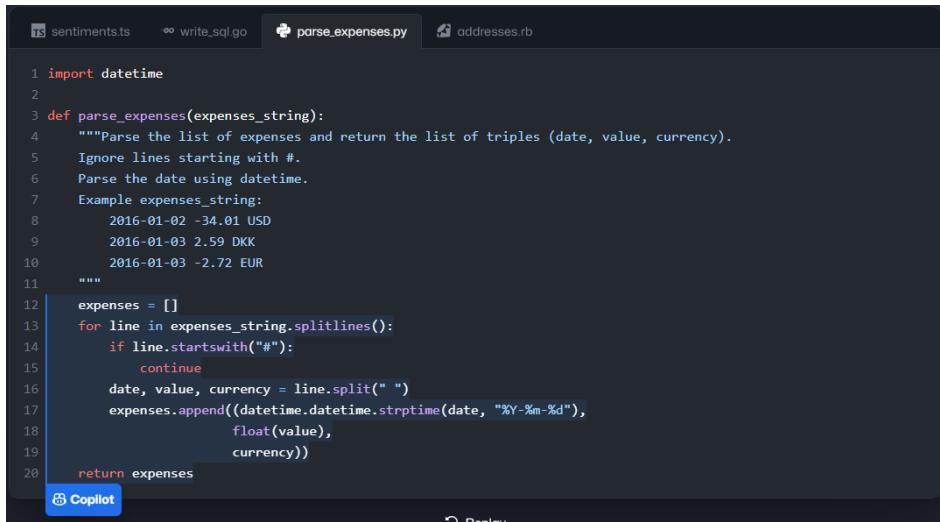
# Generering av bilder



Generert av Midjourney med følgende “prompt”:

“Photo of a salmon production facility along the coast of Norway, photorealism, high definition, focal point --s 50 --v 5.1”

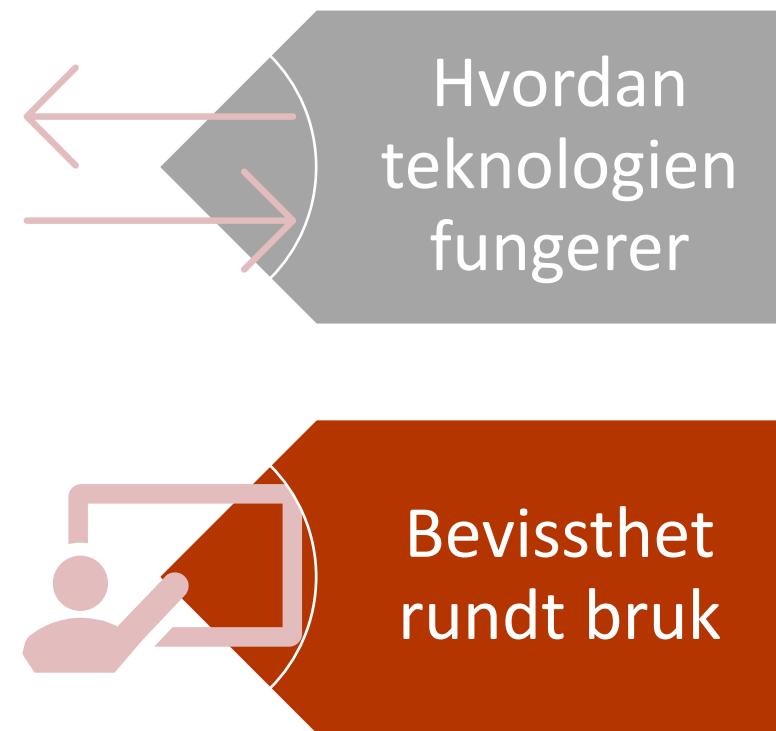
# Generering av kode



A screenshot of the GitHub Copilot interface. It shows a code editor with a Python file named `parse_expenses.py`. The code defines a function `parse_expenses` that takes a string of expenses and returns a list of triples (date, value, currency). The code uses the `datetime` module to parse dates. A Copilot AI icon is visible at the bottom left.

```
1 import datetime
2
3 def parse_expenses(expenses_string):
4     """Parse the list of expenses and return the list of triples (date, value, currency).
5     Ignore lines starting with #.
6     Parse the date using datetime.
7     Example expenses_string:
8         2016-01-02 -34.01 USD
9         2016-01-03 2.59 DKK
10        2016-01-03 -2.72 EUR
11    """
12     expenses = []
13     for line in expenses_string.splitlines():
14         if line.startswith("#"):
15             continue
16         date, value, currency = line.split(" ")
17         expenses.append((datetime.datetime.strptime(date, "%Y-%m-%d"),
18                         float(value),
19                         currency))
20
21     return expenses
```

[GitHub Copilot · Your AI pair programmer · GitHub](#)



# Finjustering og tilpasning av grunnmodeller

**En stor grunnmodell**



Tilpasning og justering

En modell som er tilpasset en spesifikk oppgaver

**GPT-4 (eksempel)**



Tilpasning og justering

En chatbot som kan svare på spørsmål om Digidirs veiledere

# Hva skjer fremover?

Nye tema

- Entreprise-lisenser

Videreutvikling

- Enda mer konkret
- Finjustering
- Risikovurderinger





[digdir.no](http://digdir.no)

Digitaliseringsdirektoratet

postmottak@digdir.no

22 45 10 00

Postboks 1382 Vik, 0114 Oslo

Besøksadresser:

**Industriveien 1, 8900 Brønnøysund**

**Skrivarevegen 2, 6863 Leikanger**

**Grev Wedels Plass 9, 0151 Oslo**