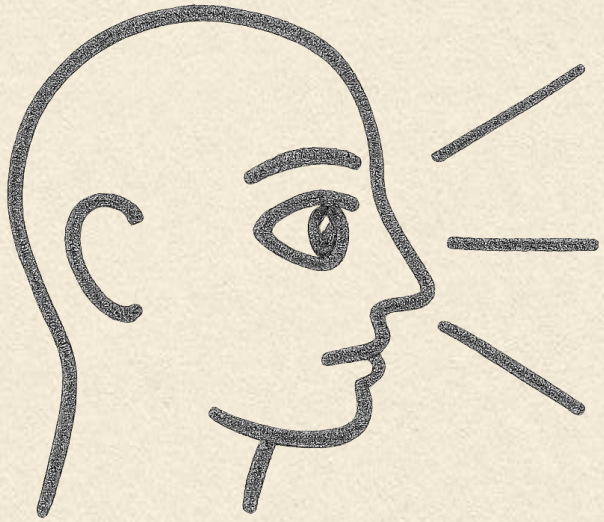
KI i offentlig sektor

– status, utfordringer og veien videre



Hans Christian Farsethås
Forsker, Inst. off. rett, UiO
Digital velferdsstat

Finansiert av Nav og Juridisk fakultet UiO



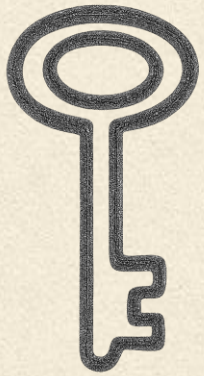
Et nytt «blikk»?

Blikk kan være et «perspektiv»...

... men også noe som trenes opp, slik som legens «blikk» på ultralyd og røntgen eller geologens «blikk» på fjellformasjoner.

Forsøke å bygge opp et «blikk» på KI i offentlig sektor.

Jeg er opptatt av språkmodeller og hvordan de kan brukes i offentlig sektors systemer. Først et skritt tilbake.



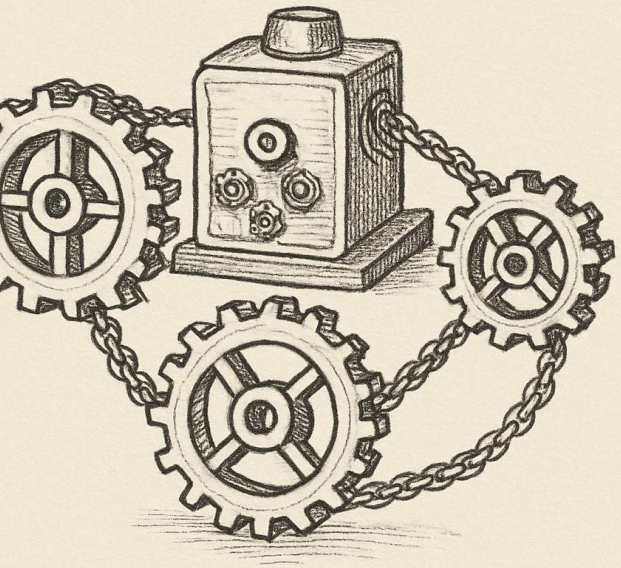
Fra åpent til lukket (OpenAI og de andre)

- ChatGPT lansert 30. november 2022. **Ganske bra dokumentert!**

- GPT 4 lansert 14. mars 2023

  «Given both the competitive landscape and the safety implications of large-scale models like GPT-4, this report contains no further details about the architecture (including model size), hardware, training compute, dataset construction, training method, or similar.»

De store amerikanske språkmodellene er ikke dokumentert, men kinesiske DeepSeek er delvis dokumentert.



Norske modeller er åpne

- Nasjonalbiblioteket
- Language Technology Group, UiO
- NTNU

Små i forhold til de ledende internasjonale, men transparente med hensyn til metoder og treningsdata.

Vi bruker modellarkitekturen fra internasjonale åpne språkmodeller. Men trenger store mengder tekst for å trene dem.

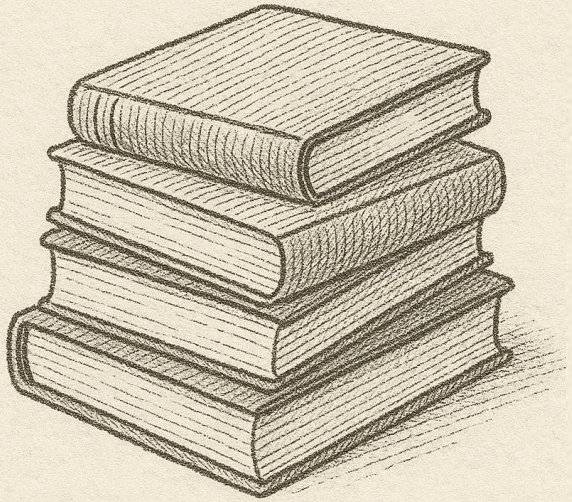
Trenes på enorme mengder tekst

2018: GPT-1 8-900 millioner ord. (8000 romaner)
2019: GPT-2 10 milliarder ord (Webcrawl <- reddit)
2020: GPT-3 225 milliarder ord (Web+Wikipedia+bøker)
2024: GPT-4 10 billioner ord ?
2025: GPT-5 85 billioner ord ??



Til sammenligning

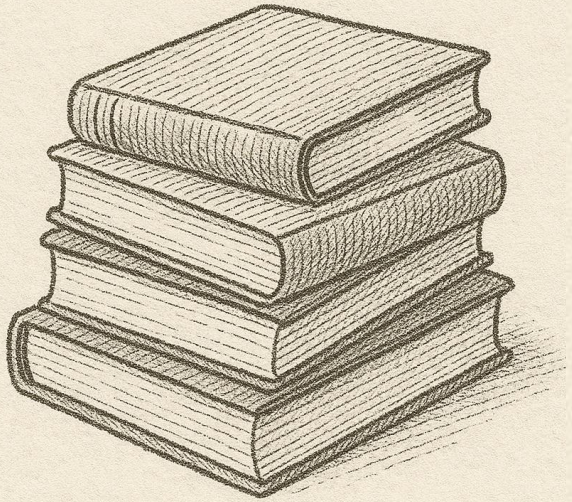
Stortingsforhandlinger 1815-2000:	1,5 milliarder ord
Alle norske kilder i Lovdata:	2,0 milliarder ord
Norske språkmodeller:	24/40 milliarder ord
Jf. de største modellene	85 billioner ord



Treningsdata i norske modeller

Målfrid (regjeringen.no 1996–2023)

- NOU-er
- Stortingsmeldinger
- Stortingsproposisjoner
- Rundskriv og regelverk



Treningsdata i norske modeller

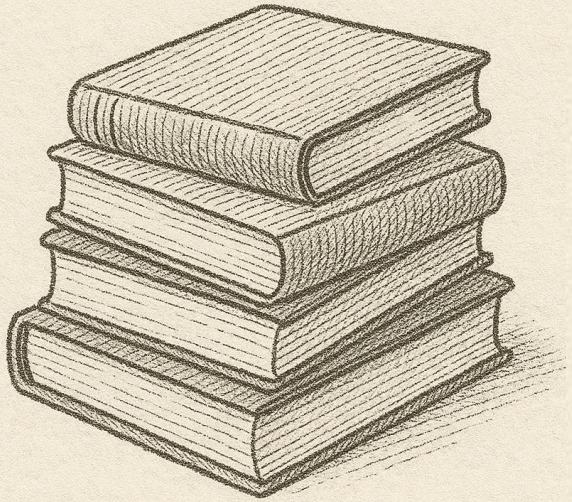
Målfrid (483 domener fra statsforvaltningen)

- v Arbeidsretten
- v Arbeidstilsynet
- v BUF

...

- v Nav (kun informasjonssider)
- v Skattedirektoratet (mange rettskilder)

- ÷ Det meste som er lagt ut på Lovdata
- ÷ Avgjørelser fra domstolene og en del nemnder

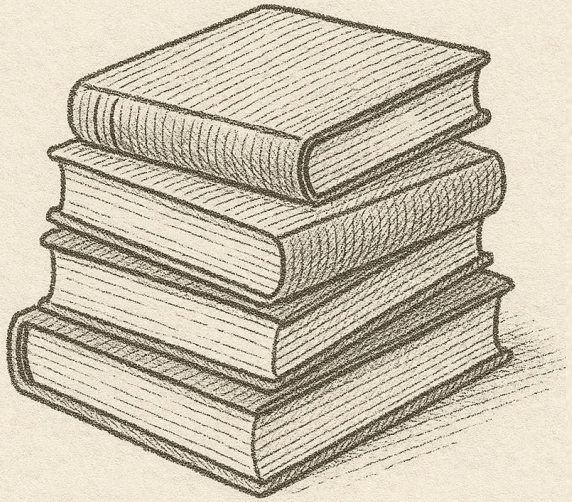


Treningsdata

Webcrawl (CulturaX og HPLT)

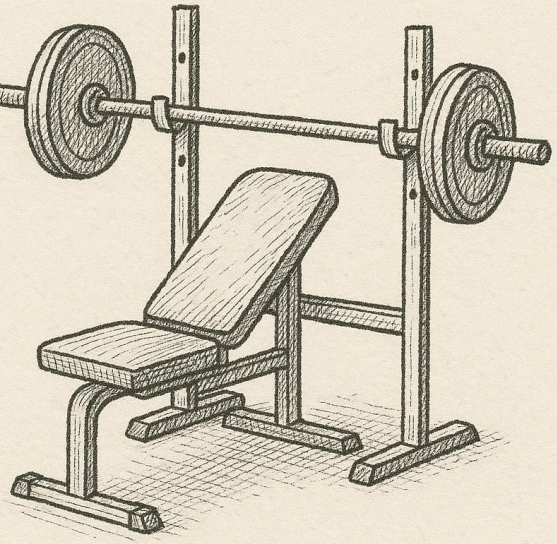
Henter inn store mengder norsk tekst fra internett. Får blant mye annet med seg:

- Stortinget.no – i hovedsak spørretime og skriftlige spørsmål og svar
- Europalov.no – i hovedsak politikkdokumenter, ikke rettskilder



Vi mangler

- Retts-, nemnds- og forvaltningsavgjørelser
- EU/EØS, EFTA, EMK
- Rundskriv og instrukser
- Til nå komplette og oppdaterte samlinger av lover og forskrifter



Trening av modellene

Finjustering

Etter å ha trent på 40 milliarder ord har vi en grunnmodell.
Den må finjusteres med trening på spørsmål og svar.

Stor mangel på

a) Datasett med spørsmål og svar
b) Datasett som ligger tett på mulige bruksområder i offentlig sektor.
c) Rundskriv, veiledninger, håndbøker osv.

Avgjørende med mer tekst for bedre ytelse!



Evaluering av modellene

v Vi har NorEval

÷ Generelle evalueringer for norsk jus
÷ Bruksspesifikke evalueringer for norsk jus.

Hvordan skal vi vite om modellene kan brukes i offentlig sektor uten grundig evalueringsmetodikk?
→ Like relevant for norske som internasjonale modeller
→ Også relevant for revisjon.

Viktigå utvikle evalueringer som ligger tett på potensielle bruksmåter



Status: Hva brukes de til?

- Rettsinformasjonssystemer
- Chatboter med spørsmål og svar
- Sortering av henvendelser
- Ulike former for støtte til saksbehandlere

Inntrykket er at det er noe i produksjon, stor vilje og mange forsøk.

Dette er bruksområder der norske modeller bør kunne matche utenlandske etter hvert.



Ønskeliste

1. Mer data, mer tekst.
   → Krever støtte fra beslutningstakere
2. Samarbeid om evalueringsdata og metodikk
   → Nasjonal infrastruktur for evaluering av språkmodeller til bruk i offentlig sektor
3. Tettere samarbeid mellom modellbyggere, tekstsamlere og offentlig sektor. Modellene er dyre og tidkrevende å lage. Bør ha som mål å lære noe nytt om bruk i offentlig sektor for hver ny modell som lages (eks. samisk)

Ta gjerne kontakt!

h.c.farsethas@jus.uio.no

Hans Christian Farsethås, «Rettskilder i de norske språkmodellenes treningsdata. En analyse av Mímir-base», Lov og Rett, nr. 8 2025

Hans Christian Farsethås og Anders Løvlie, «Lovdatas nye KI-søk», Lov og Rett, nr. 9 2025.