

Kunstig intelligens og etikk

Erik Johannes Bjørnson Løvenskiold Grüner Husom

2024-03-06

Mitt mål i dag

Øke forståelsen for kunstig intelligens og problemstillingene vi står ovenfor når denne teknologien tas i bruk i samfunnet.

Hva er kunstig intelligens?

Hva er kunstig intelligens?

- *Kunstig intelligens er en form for dataprogram som utfører oppgaver på en tilsynelatende intelligent måte, og som kan lære av datainnsamling.*
- **KI = automatisering**

Hva bruker vi KI til?

- Automatisere oppgaver (for å spare tid/ressurser)
- Løse oppgaver vi ikke får til selv

Hva er mulighetene?

- KI kan benyttes innen alle sektorer

Ulike former for utfordringer

Ulike former for utfordringer

- 1 Iboende utfordringer
- 2 Konsekvensene av hvordan det lages og driftes
- 3 Konsekvensene av hvordan det brukes

1. KI og iboende utfordringer

- Pålitelighet
 - Usikkerhet
 - Forklarbarhet
 - Ansvar

Estimering av usikkerhet

- Hvor sikker er KI-modellen på sin prediksjon?
- Hvorfor trenger vi dette?
 - Ta tryggere avgjørelser
 - Bygge “tillit”
 - Sikkerhet og robusthet

Estimering av usikkerhet

Vi har to typer usikkerhet:

- Har vi usikkerhet pga dårlig datagrunnlag
- Har vi usikkerhet fordi modellen er dårlig?

Eksempel: KI for gjenkjenning av beinbrudd

Unikt prosjekt i Noreg – har allereie spart 115 døgnsventetid

Sjukehusa i Vestre Viken er dei einaste i landet som har tatt i bruk kunstig intelligens i samband med røntgenanalyse. Emilie (8) blei pasient nummer 10.000 i KI-kø.



Emilie Johansen Geirholt var uheldig og hamna på sjukehuset. Ho blei likevel pasient nummer 10.000 – og det gjekk ikkje upåakta forbi.

FOTO: TORDIS GAUTEPLASS / NRK



Henning Renhovde
Journalist

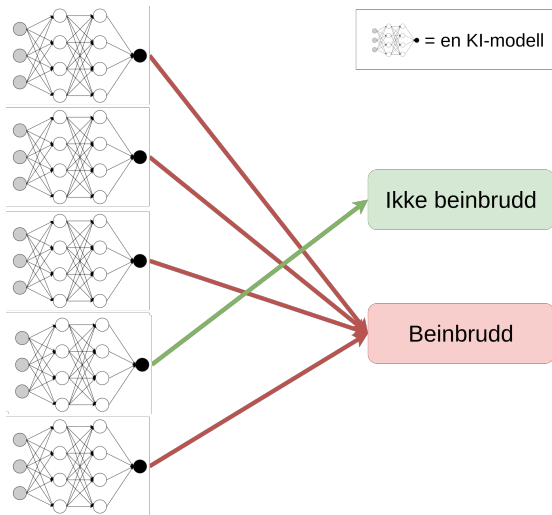


Tordis Gauteplass
Journalist

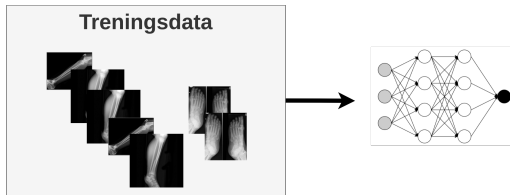
Vi rapporterer frå Drammen

Publisert i går kl. 08:04

Er flere modeller enige om det samme?



Har modellen sett noe lignende før?



Forklarbarhet

- Åpenhet og innsikt
- Trenger innsikt i modellen for å vite om svakheter
- EUs personvernforordning: Alle har rett til en forklaring på automatiserte prosesser
- “Forklaring” = hvilke deler av input-dataene brukte modellen til å produsere en prediksjon

A COMPUTER

CAN NEVER BE HELD ACCOUNTABLE

THEREFORE A COMPUTER MUST NEVER
MAKE A MANAGEMENT DECISION

Utfordringer ved store språkmodeller

- Er det noen forståelse der?
- Hvordan kan vi bli trygge på det den produserer?

1. Iboende utfordringer – oppsummering

- Pålitelig egenskaper er ikke innebygd.
- Det må stilles krav til at man skal integrere denne type analyse i systemene.
- 100% korrekthet er ikke nødvendig, men det er viktig å kjenne systemets begrensninger.

2. Konsekvensene av hvordan KI lages og driftes

KI og ressursforbruk: Energi

The New York Times

SUBSCRIBER-ONLY NEWSLETTER

Climate Forward

A.I. Could Soon Need as Much Electricity as an Entire Country

Behind the scenes, the technology relies on thousands of specialized computer chips.



Share full article



FORBES > TECH

AI'S Unsustainable Water Use: How Tech Giants Contribute To Global Water Shortages

Federico Guerrini Contributor 

Follow



Apr 14, 2023, 08:49am EDT

KI og ressursforbruk: Materialer

- Nye og kraftigere maskiner
- Alt må(?) oppgraderes

KI og ressursforbruk: Hva kan vi gjøre?

- Faktisk måle energiforbruket
- “Gjenbruk” av modeller
- Optimalisere modellene (gjøre dem mindre energikrevende)
- Rapportering og sertifisering av modeller

KI og ressursforbruk: Hva kan vi gjøre?

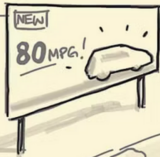
KI kan også brukes til å bidra i den grønne omstillingen:

- Smarte strømnett
- Optimalisere transport
- Mer energieffektive bygninger

JEVONS PARADOKS

ØKT DRIVSTOFFEFFEKTIVITET PLEIER Å ØKE, IKKE
SENKE, DRIVSTOFFORBRUKET.

DISSE NYE BILENE ER SÅ
EFFEKTIVE AT ALLE KJØRER
OVERALT NÅ FOR TIDEN.



Utnyttelse av åndsverk

The New York Times

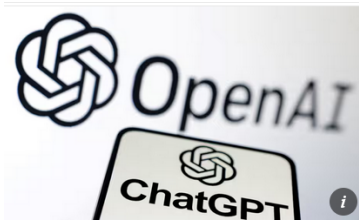
The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work

Millions of articles from The New York Times were used to train chatbots that now compete with it, the lawsuit said.

Utnyttelse av åndsverk



World UK Climate crisis Ukraine Environment



OpenAI

'Impossible' to create AI tools like ChatGPT without copyrighted material, OpenAI says

Pressure grows on artificial intelligence firms over the content used to train their products

Utnyttelse av arbeidskraft



TIME

SUBSCRIBE

BUSINESS • TECHNOLOGY

Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic

15 MINUTE READ



2. Konsekvensene av hvordan KI lages og driftes – oppsummering

- Lange produksjonskjeder gir et uklart kostnadsbilde

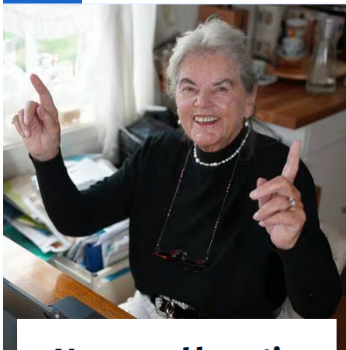
3. Konsekvensene av hvordan det brukes

KI og “overvåkningskapitalisme”

- KI + datainnsamling = sant
- Hvem sitter med makta?
- Personvern
- Kan KI være en trussel mot menneskets frihet?

KI og kognitiv avlastning

- Hva gjør det med oss å la KI overta kognitive oppgaver?



Venn med kunstig intelligens

For Agnete (90) utfyller ChatGPT rollen til en nær venn i hverdagen. Hun ønsker at flere eldre skal få samme mulighet. Men forskningsmiljøer er skeptiske.

30. SEP, 2023 KL. 14:42

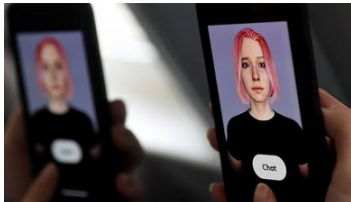
Ine Julia Rohahn Schwebs Journalist

Norge

Siste nytt Dokumentar Klima NRK Ytring

Stadig fleire KI-romansar: – Kjærlaik fri for risiko

Det som for mange før verka som eit merkeleg framtidfenomen skjer no på ekte; fleire og fleire blir forelska i og saman med samtalebodar.



SKAP SJØLV: I Replika kan du sjølv bestemme utsj... [Vis mer](#)

Ingrid Emilie Waaler Journalist

Ine Julia Rohahn Schwebs Journalist

Oppdatert 15. sep. 2023 kl. 03:26

Del

KI og menneskesyn

- Spiller det noen rolle om noe er “ekte”?
- Intelligens, forståelse, bevissthet?

KI og transhumanisme

- Transhumanisme: Forbedre mennesket ved hjælp av teknologi

Hva er kjerneutfordringen?

- Utviklingen (og utrullingene) går for raskt

Hva gjør vi?

Hva gjør vi?

Regjeringen med milliardsatsning på KI: – Vil forandre samfunnet på måter vi fortsatt ikke forstår

Regjeringens nyeste initiativ møtes begeistring av forskere på feltet. Høyre kaller det et viktig skritt.



Martin Gundersen
Journalist



Ine Julia Rojahn Schwebs
Journalist

Publisert 7. sep. 2023 kl. 21:50

SATSER PÅ KI: Regjeringen setter av én milliard til forskning på kunstig intelligens.

FOTO: MARTIN SOLHAUG STANDAL / NTB

Hva gjør vi?



Avslutning

There's too much effort trying to create autonomous machines rather than trying to create machines that are useful tools for humans.

– Prof. Emily Bender

What we should be concerned about, I suggest, is not human uniqueness, as if we alone in the universe are bearers of those capacities that we deem to be distinctively human. Rather, our concern should be for human dignity: we should not fear the upgrading of robots so much as the downgrading of human beings.

– Prof. Robert Song

Anbefalt litteratur

- Om kunstig intelligens og kristen tro:
 - “The robot will see you now” – John Wyatt and Stephen Williams (bok)
- Om den digitale fremtidens trussel mot menneskelig natur og frihet:
 - “The age of surveillance capitalism” – Shoshana Zuboff (bok)
- Om farene ved de nyeste gjennombruddene innen KI:
 - “The AI dilemma” (videoopptak av foredrag)
- En tankevekker angående vårt forhold til moderne teknologi:
 - “The way home” – Mark Boyle (bok)
- Om teknologibransjens kultiske forhold til KI:
 - [The cult of AI](#) (artikkel)

Lenke til presentasjon:

- Presentasjon: <https://erikjohannes.no/slides/talk-2024-03-06-kunstig-intelligens-og-etikk.pdf>
- Kontakt: erik.johannes.husom@sintef.no



Spørsmål?

Spørsmål til diskusjon

- Hva vil det si å være menneske?

Spørsmål til diskusjon

- I hvor stor grad opplever dere at KI påvirker hverdagen deres i dag? Er det positivt eller negativt?
- Spiller det noen rolle om noe er menneskeskapt eller ikke?
- Spiller det noen rolle om en avgjørelse er tatt av KI eller mennesket?
- I hvilke tilfeller (hvis noen) er det bedre om en KI-modell tar avgjørelser istedenfor mennesker?
- Hvordan kan vi ta vare på våre relasjoner og menneskelighet i møte med en teknologi som lages for å automatisere "alt"?
- Kan vi nå et punkt hvor vi har automatisert nok?
- Hva gjør det med menneskesynet vårt at vi får teknologi som automatiserer kommunikasjon og som erstatter mange menneskelige funksjoner?