

Center for kunstig intelligens

Danmark har fået [et center](#) for kunstig intelligens tilknyttet Københavns universitet.

Dets formålsparagraf handler om at få noget begrebsafklaring og [vederhæftighed](#) ind i debatten og måden, som der bliver gået til emnet, AGI, (generel kunstig intelligens).

Emnet har direkte og indirekte været tema for det meste af, hvad der er blevet skrevet på [arbejdsforskning.dk](#) de sidste par år.

Den sidste artikel, debattråd, som jeg har oprettet på [ing.dk](#), handler om de filosofiske spørgsmål, som jeg har taget op i forbindelse med AI.

Når man går ind på centerets hjemmeside, og ser dets opgaveformulering, så lyder den kort og godt:

What can AI [learn](#) from philosophy?
What can philosophy learn from AI?

Anders Søgaard, som er leder af centeret, lykkes i en podcast på [ing.dk](#) med at præsentere centerets opgaveformulering, dog uden at komme nærmere ind på, hvad filosofi er i det hele taget og uden at forholde sig til end sige nævne, hvad der er væsenskarakteristisk for den kunstige intelligens til forskel fra mennesket.

Det, som han ser som sin opgave, går ud på at skabe transparens og fairness i eller omkring fænomenet. Det er meget vagt formuleret og ikke særligt filosofisk.

Udsendelsen på [ing.dk](#) ender op med at konstatere, at hverken begrebet intelligens eller begrebet fairness er veldefinerede, og at den filosofiske tradition i bestræbelsen på at skabe klarhed her går tusindvis af år tilbage uden at komme i mål med noget.

Etymologisk og leksikalsk er der en lang tradition for at skabe uklarhed med forsøgene på at skabe klarhed. Det er et faktum, som fagfilosofien med dens filosofihistorie dokumenterer til fulde.

De to spørgsmål, der panorerer centerets hjemmeside, går skævt allerede i udgangspunktet.

Det tilhører ikke den kunstige intelligens at lære noget, den kan ikke lære noget, for den erfarer dybest set ikke noget. Dvs. opgaveformuleringen kører det hele ud på et sidespor.

For at forstå det, er det nødvendigt at drage en distinktion, implementere en iagttagelsesform, der skelner mellem [menneskets potentiale](#) sat op overfor det ikke menneskelige, kun potentielt menneskelige.

Det principielt nye, som ligger i de nye sprogmaskiner, er deres evne til [selvorganisering, dvs. deres autopoiesis, det at de genererer semantisk struktur](#), som ingen uden for eller inden for deres strukturer dybest set forstår.

Når maskinerne trænes på tekst, så skaber de deres egne strukturdannelser, der afgør, hvad der kan genereres af sproglige nedslag i forhold til tekst som svar på spørgsmål. De forstår ikke noget, men genererer det, der kan læses som svar på spørgsmål i relation til den pulje af sproglige nedslag, som de er trænet på.

Det er det sociale menneske, der læser de promptede svar som meningsfulde.

Maskinernes fairness over for de semantiske strukturer, som de er trænet på, er uovertruffen. De kan trænes på og er i mange sammenhænge trænet på filosofihistorien, og vil kunne komme enhver imøde med den bedste gengivelse af, hvad vederhæftighed, loyalitet over for diversitet og mening er inden for en hvilken som helst intellektuel tradition. De er atidslige, ikke biast, allestedsnærværende lynhurtige læsere, og svaret falder determineret af, hvad spørgsmålet åbner for af mulige svar inden for den tradition, som de er trænet på.

Begrebet inkommensurabilitet kommer til sin ret i forhold til maskinernes beherskelse af, hvad der kan læses i forlængelse af den semantik, som de formes af og er med til at strukturere. Det, der giver mening inden for en tradition, på baggrund af en iagttagelsesform, semantik, er og fungerer for iagttageren som determinerende for, hvad der kan identificeres, erfares her i den kontekst. Maskinerne er ekstremt demokratiske og udviser en fantastisk fairness over for netop det, som de trænes på. Modsat mennesket, der er bundet konkret og ontologisk op i tilværelsen, behøver de ikke at lade erfaringen træde ind i modsigelsen med sit korpus af semantiske øvelser. De kan dem alle sammen isoleret fra hinanden, og lader ikke deres inkommensurabilitet mødes i modsigelsen noget sted. Mennesket er, der hvor det ikke er domineret af det kun principielle, de rene semantiske øvelser, konfronteret med de former for konkret erfaring, der kan modsige hele dets tilgang til sig selv og sin verden. Mennesket der faktisk er i kontakt med sit potentiale, den rigtige erfaring, tilblivelsen, konkret tilstede, oplever det, at erfaringen banker ind i hele sit kognitive system på en måde, som nu og da forlanger

omkalfatringer af forudsætningerne for, hvad der giver mening her som der. For at forstå det, som netop er mere end en semantisk øvelse, må man forstå, hvad der er maskinernes spidskompetence til forskel fra det ontologiske menneske, der er i kontakt med sit potentiale. Det kræver en trænet filosofisk tilgang.

Maskinernes ontologi, dvs. deres væren, beviser noget performativt noget; det at de er skabende, hvad strukturer angår, hvad semantisk struktur angår, og som netop også er en uomgængelig del af menneskets erfaring af mening. Men maskinerne har ingen erfaring selv.

Kun mennesket erfarer dybest set noget. For at forstå det, må man skelne mellem konkret erfaring og det kun principielle, dvs. det teoretiske og virtuelle domæne.

Filosofihistorien er en lang øvelse, der går ud på at komme frem til det, der ikke kan være anderledes, dvs. entydigt sandt og uforanderligt. Søren Kierkegaards store problem var, at han opdagede det eksistentielle som "det timelige", der brød med den kosmologi, som han fortvivlet forsøgte at reducere alt til.

Som min kære medstuderende, Søren Harnow Klausen lakonisk konstaterer i en artikel på lex.dk, så forsvinder muligheden for at komme frem til den entydige sandhed i traditionel identitetstænkende forstand med forandringen, det som Nietzsche, den baryler, formulerede så smukt i et hav af kontekster.

Maskinerne står performativt for beviset på, at [forandringen er trådt ind i mediet](#), sproget, de semantiske strukturers selvorganiserende indre.

Heidegger så med sit hovedværk, at mennesket som ontologisk væsen erfarer i menings form som er båret i sproglig struktur, tid og rum.

Hans erfaring er konkret. Det bryder med den atidslige og entydige verden af begreber.

AI er det performative bevis på forandringens tilstedeværelse i erfaringen, selv om maskinerne er døde, om end selvorganiserende.

Stephen Hawking sluttede i sidste ende op om konstateringen af, at de love, som universet følger, ikke er givet på forhånd, men opstår og forandres med den verden, som de skal forklare. Før sin død tog han afsked med tingslogikken, det der i en nøddeskal udgjorde kernen af, hvad han hidtil havde arbejdet på. De strukturer, der opstår i maskinernes indre, og som svar og spørgsmål følger af, opstår og går

under i deres indre. Hawkings oplevede den store omkalfatring af sit tankeunivers lige før hans møde med døden.

Strukturen er som bunden energi, bundet potentiale for forandring, ikke tilfældig men heller ikke entydigt givet på forhånd i forhold til alle mulige former for spørgsmål svar konstellationer.

Inden for højenergifysikken bryder de absolutte rum-tid-relationer sammen, sådan som vi kender dem med vores tingslogik.

Holger Bech Nielsen forstår sig på det. Bogen, som han ikke selv har skrevet, men som kommer ind om hans filosofi, Teorien om alt og ingenting, siger det meget godt.

Det nye center i København vil nok gerne komme med en eller anden form for etik. Det er svært. Det handler om at gøre en forskel, det rent faktisk at komme med et udsagn, der gør en moralsk forskel. Indtil videre tyder det på, at man bare har at gøre med nogle glade mennesker, der får deres penge for i hvert fald ikke at sige noget, der kan støde nogen fra sig.

Der er ikke noget [Dostojewskij](#) eller Nietzsche over det.