

Sprudler av overskudd



Trond Kjærstad
Promethevs AS

Kunstig intelligent psykologi viser oss hva kunstig intelligens er, og hva det ikke er.

Anmelderen har fylt ut interessekonflikt skjema og oppgir ingen interessekonflikter.



BOK *Kunstig intelligent psykologi*

FORFATTER Jan Ketil Arnulf

ÅR 2025

FORLAG Fagbokforlaget

SIDER 208

Språkteknologien – de snakkende maskinene – har på kort tid utfordret vår selvforståelse som den eneste skapningen med intelligens. Jan Ketil Arnulf viser i sin nyeste bok, *Kunstig intelligent psykologi*, at våre nye lekekamerater, språkmodellene, heller gir oss et nytt blikk på hva det vil si å være menneske – hva som skiller menneskelig intelligens fra den kunstige. Det gir oss samtidig en unik mulighet til å forstå hva språk egentlig er: et bindemiddel mellom oss mennesker, et forsøk på å gjøre hverandre forutsigbare. Språket er en kulturell programvare som gjør oss i stand til å samhandle. For maskinene oppfører seg ikke – tenker ikke – som oss, selv om de nå snakker, henvendt til oss.

Kunstig intelligent psykologi er på mange måter en oppsummering av temaer Arnulf har vært opptatt av gjennom mye av sin karriere. Han er professor i organisasjonspsykologi ved Handelshøyskolen BI, og har særlig vært opptatt av språk og dets betydning for temaer innen organisasjon og ledelse. Boka tar først for seg hvordan det står til med oss mennesker. Den går steg for steg gjennom hvordan vi bør forstå begrepet intelligens, at språket er utviklet som et verktøy for å koordinere individer med sammenfallende interesser, og hvordan språket derfor er et forenklet kart over en mangfoldig virkelighet. Deretter følger en kort innføring i informasjonsteori og prinsippene bak hvordan de kunstige språkmodellene er bygget opp, og hvordan de skiller seg fra oss. Kjernen i boka er at vi mennesker *vet* hva vi vet, vi er de eneste som evner å skille mellom kart og terreng. Dette

kan vi fordi vi til forskjell fra alle andre arter sammenligner kartene våre med andres kart, altså andres forestillinger om verden. Den sammenligningen gjør vi ved hjelp av språket. Dessverre er språket et så forførende verktøy at vi fort glemmer at virkeligheten slett ikke ligger i språket, men «der ute», i andre mennesker og vår delte forståelse av virkeligheten.



Arnulf er selv overrasket over hvor lett det er å lage snakkende maskiner, bare treningsmengden er stor nok. Maskinene snakker og snakker, basert på alt vi mennesker allerede har snakket om. Når språket blir automatisert, er det kanskje lettere å se hvordan «sannheter» oppstår. Maskinene «hallusinerer», de skrøner forbeholdsløst i vei, basert på svimlende kalkulasjoner av hva som bør være neste ord, neste bokstav. Men de er og blir kalkulatorer. Det er dessverre ikke så mye bedre stilt med oss. Vår evne til å gjøre hverandre forutsigbare gjennom et felles språk har samtidig redusert vårt behov for utforskning av virkeligheten.

På sett og vis kan man si at boka er «vanskelig», men ikke fordi den er innfløkt skrevet. Tvert imot sprudler den av overskudd og konkrete, morsomme eksempler. Men det *er* vanskelig å se seg selv utenfra. Vi lever og tenker gjennom språket, og derfor er det krevende å ta inn over oss at virkeligheten ikke finnes der. Dette er altså det sentrale poenget i boka: Språket er kun et tynt og skjørt kart over en uforutsigbar verden. Samtidig er dette kartet veldig effektivt for å navigere i verden. Det har gjort oss i stand til å dele ideer og løsninger og dermed gjort teknologiske nyvinninger mulige, fra de enkleste steinalderredskaper til kunstig intelligens. Slik har språket forført oss, til vi til slutt glemmer at kart og terreng er svært forskjellige.

Det er her språkmodellene faktisk hjelper oss. Ikke gjennom deres beherskelse av språk, men ved å vise oss vårt eget slumsete forhold til dette språket. Maskinene viser oss hvor upresist kartet egentlig er, for vi hallusinerer, vi også. Abstrakte begreper kan være nyttige for kommunikasjonen, men det er lett å glemme at de ikke er virkelige. Vi sier «motivasjon» og later som ordet presist beskriver noe virkelig. Men kartet sier kanskje mer om den som tegner kartet, enn om terrenget. Tidligere har fenomener som motivasjon eller mangel på sådan blitt forklart med for eksempel nærvær eller fravær av demoner. Slike kart har vi byttet ut etter som vi har ervervet oss ny kunnskap om verden. Men er det noen grunn til å tro at dagens kart er så mye bedre? Det er først når vi skal prøve å formidle disse tilstandene til hverandre, i hovedsak gjennom språket, at vi må forenkle virkeligheten såpass at den gir mening for andre. Så blir vi enige om en merkelapp som fungerer godt nok til at vi kan handle på det, sammen. Denne merkelappen springer alltid ut av en bestemt relasjon i en bestemt kontekst, men verden blir mye enklere hvis vi antar at vi nå har «forklart» noe en gang for alle. Det er her det blir ekstra brennbart for oss psykologer. Vår yrkesgruppe er spesielt flinke til å lage språklige konstruksjoner som vi så tror gjenspeiler en fysisk virkelighet. Som professor i organisasjonsfag på Handelshøyskolen BI er det naturlig at Arnulf retter skytset spesielt mot egen fagkrets, men det er ingen grunn til å tro at andre psykologifelt bør slippe billigere unna. «Alle lederteorier er egentlig de samme, ettersom de stort sett bare er ulike semantiske formuleringer av samme tema», skriver han. De faktiske handlingene som finner sted mellom mennesker, er selvsagt langt mer mangfoldige enn et enkelt begrep kan favne. Likevel benytter og utforsker vi det som om det er en entydig beskrivelse av virkeligheten. Men da har vi altså blandet sammen kartet og terrenget. På samme måte lar han begreper som «motivasjon» og «velferd» få gjennomgå.

Aller mest nådeløs er Arnulf når han viser hvordan psykologers måleinstrumenter, stort sett spørreskjemaer, ofte måler sammenhenger i språket, ikke i virkeligheten. Når det man forsker på, henger litt sammen i språket, er ofte svarmønstrene forutsigbare *før* noen har svart. Forskningen gir inntrykk av sterkere sammenhenger enn det som faktisk gjøres i praksis.

Det vi psykologer og andre samfunnsforskere driver med når vi forsøker å sette ord på fenomener, er å ta fenomenet ut av sin opprinnelige kontekst og tro at det gir like mye mening overalt. Med språkalgoritmenes inntog blir dette mer akutt, og til syvende og sist blir hovedagendaen med boka derfor å skjerpe blikket for hva kunstig intelligens er og ikke er. KI mestrer språket vårt relativt lett, men vi mennesker er ikke laget for å leve i språket. Vi er heldigvis ufattelig mye mer komplisert konstruert enn maskinene for å kunne håndtere den virkeligheten som språket simpelthen bare kan rapportere om. Les denne boka, og bli (enda) litt smartere enn maskinene.

