

Anders Søgaard: Hvorfra ved vi, hvad en sprogmodel ved?

Vi taler tit om, hvad en sprogmodel ved. Ved ChatGPT, hvem der vandt det sidste Europaparlamentsvalg - eller hvem Steen Steensen Blicher var? Men mange filosoffer - og en del forskere i kunstig intelligens - rynker på næsen over den måde at snakke om sprogmodeller. Sprogmodeller er papegøjer, der blot kombinerer forskellige tekstbidder på mekanisk vis. Så hvad vil det sige - for en filosof - at vide noget. Hvad skal der til, før vi kan sige, at et menneske - eller en sprogmodel - ved, at et sagsforhold S er tilfældet? Typisk involverer det noget med at have en teori om verden og en begrundelse for S. Og det er også derfor filosofferne rynker på næsen. Sprogmodellerne har jo ikke teorier om verden og kan ikke begrunde deres udsagn, vel? Jeg vil her forsøge at komplicere det hele lidt. Det er nemlig ikke åbenlyst rigtigt, at filosofferne har ret. Sprogmodeller - af en vis størrelse og kvalitet - opbygger noget, der minder temmelig meget om teorier om verden, og man kan tit også udlede begrundelser for deres udsagn - med ligeså stor eller større validitet end med mennesker.