

# Når etikk betyr alt

Skal vi glede eller grue oss til fremtidens intelligente teknologi? Det er ikke bare filmskapere som er opptatt av dette.

**Forskning viser at ...**  
**Jim Tørresen**



**H**vordan blir fremtiden vår når roboter og teknologier vi omgir oss med kan ta til seg kunnskap på egen hånd og lære seg å gjøre oppgaver og tilpasse seg til nye situasjoner? Dette kalles intelligente og lærende systemer, eller kunstig intelligens (KI). Svarene varierer med hvem vi spør.

Filmskapere er ganske entydige på at fremtiden blir lite hyggelig. Siden filmmediets barndom har det strømmet ut dystopiske filmer på rekke og rad. Disse gir nok ikke de mest sannferdige bildene. I det siste har det kommet advarende utsagn fra mer tillitvekkende personer som vitenskapsmannen Stephen Hawking og Microsoft-gründer Bill Gates.

Er det noe i dette? Og hvordan skal vi i så fall tilpasse videre utvikling for å komme faresignallene og problemene i forkjøpet?

Mange forskere er opptatt av dette, som det fremgår av en nylig publisert artikkel i *Frontiers in Robotics and AI*.

To faktorer kan bidra til at teknologien oppfører seg slik vi ønsker.

For det første må de som utvikler intelligente systemer ha fokus på etiske problemstillinger.

For det andre, og mer utfordrende, må systemene på egen hånd ha evnen til å gjøre



**Filmskapere er ganske entydige på at fremtiden blir lite hyggelig. Siden filmmediets barndom har det strømmet ut dystopiske filmer på rekke og rad. Disse gir nok ikke de mest sannferdige bildene. Her roboten «Wall-E» fra amerikansk dataanimert science fiction-film. Foto: Pixar/Walt Disney Studios**



**Jim Tørresen,** gruppeleder Robotikk og intelligente systemer (ROBIN), Universitetet i Oslo.

## Forskningen

**Hvem:** Jim Tørresen

**Hva:** «A Review of Future and Ethical Perspectives of Robotics and AI»

**Hvor:** Artikkel i *Frontiers in Robotics and AI* (januar 2018) <https://www.frontiersin.org/journals/robotics-and-ai> <https://doi.org/10.3389/frobt.2017.00075>

etiske vurderinger for situasjoner som ikke er kjent når systemene blir utviklet.

Den største utfordringen er å beskytte datasystemer mot at uvedkommende får tilgang og tar kontroll - eller misbruker informasjon. Elektronisk sikkerhet (eng. security) et sentralt begrep.

Men vi vil ikke bare forholde oss til data som kan feile og misbrukes, men også fysiske

roboter og andre teknologier som beveger seg i vår nærhet. Innføringen av selvkjørende biler er et godt eksempel på vi også må være opptatt av fysisk sikkerhet (eng. safety). Her trengs det god beskyttelse mot misbruk fra andre, samtidig som systemene må ha gode sanseegenskaper når de befinner seg fysisk nærmere oss enn hva de gjør i dag.

Det jobbes mye med å lage systemer som er generelt intelligente, fremfor spesialisert på avgrensede oppgaver som dagens systemer stort sett er. Da blir det samtidig stadig vanskeligere å teste hvordan systemet vil oppføre seg under ulike situasjoner.

Det er dermed behov for at systemet selv kan resonnerer rundt etiske problemstillinger. Dette kan være krevende å gjøre på generelt grunnlag, men noen enkle mekanismer kan være viktige. Det angår i hvilken grad et system på egen hånd skal få lov til å ta avgjørelser.

I noen tilfeller er dette helt nødvendig som for eksempel i selvkjørende biler der det ikke vil være tid til menneskelig

vurdering. I andre sammenhenger der mange ulike samvirkende systemer potensielt kan være koblet sammen, som innen energiproduksjon, finans, produksjon, med mer, kan mulige skadevirkninger bli ganske store hvis det tas ukloke valg. Her vil det være fornuftig med en begrenset selvstendighet og beholde mennesket i beslutningsløkken til samvirkende systemer. Personer kan følge med i avgjørelser som tas og gripe inn hvis det tas uønskede eller feilaktige valg.

Det er samtidig opp til politikere å lage lover og reguleringer for å begrense uønskede endringer i samfunnet, enten om det er som følge av selvkjørende biler eller roboter til bruk innen eldreomsorg.

Vi jobber på Universitet i Oslo med forskningsprosjekter finansiert av Norges forskningsråd der etiske vurderinger er viktig. I *Engineering Predictability with Embodied Cognition (EPEC)* utvikles systemer med interne modeller for sin egen tilstand og sine omgivelser. Målet er forbedrede løsninger

innen så ulike anvendelser som en robots oppførsel og en brukers preferanser ved automatisk generering av musikk. Dette er også relevant i prosjektet *Multimodal Elderly Care Systems (MECS)* der vi ser på bruk av mobile roboter som hjelpemiddel i oppfølgingen av eldre som bor alene hjemme.

Vi opplever stor internasjonal interesse for disse problemstillingene, blant annet gjennom prosjektet *Collaboration on Intelligent Machines (COINMAC)* der vi samarbeider med partnere i Brasil, Japan og USA om teknologiske og tilhørende etiske problemstillinger. Verdens største og mest betydningsfulle ingeniørorganisasjon innen data og elektronikk (IEEE) har engasjert forskere til å utvikle standarder for etisk tilpasset design med fokus på «A Vision for Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems».

Det er også andre engasjementer i EU og Japan for å diskutere problemstillinger og utvikle retningslinjer for fremtidens intelligente teknologi. Så forskere synes å være likeså opptatt av fremtiden som filmskapere, men får noe mindre oppmerksomhet.

**Jim Tørresen,** gruppeleder Robotikk og intelligente systemer (ROBIN), Universitet i Oslo og medlem av Den nasjonale forskningsetiske komité for naturvitenskap og teknologi (NENNET)

**Forskning viser at ...**  
Oppsiktsvekkende funn?  
Ny innsikt?  
Skriv til spalten «Forskning viser at ...». Tekstlengde inntil 3500 tegn (inkludert mellomrom).  
**debatt@dn.no**