

## 4

# AUTOMATI- SERINGENS TIDSALDER — EN TRUSSEL ELLER EN MULIGHED?

Af

**JOHANNA SEIBT**

PROFESSOR MSO, PH.D.  
INSTITUT FOR KULTUR  
OG SAMFUND,  
AARHUS UNIVERSITET

MODTAGET STØTTE TIL  
SEMPER ARDENS-PROJEKTET  
*WHAT SOCIAL ROBOTS  
CAN AND SHOULD DO  
— TOWARDS INTEGRATED  
SOCIAL ROBOTICS*

I februar 2018 offentliggjorde internationale eksperter en detaljeret rapport om de sikkerhedstrusler som kunstig intelligens formodes at udgøre. Ud over trusler mod vores digitale og fysiske sikkerhed – såsom muligheden for at hacke selvkørende biler – er der mulige trusler mod vores politiske sikkerhed. Professor mso Johanna Seibt står i spidsen for Carlsbergfondets Semper Ardens-projekt "What Social Robots Can and Should Do - Towards Integrative Social Robotics", der udforsker sociale aspekter af den hastigt voksende udbredelse af robotter i samfundet. Projektet giver samtidig bud på hvordan nye sociale robot-applikationer kan udvikles ansvarligt.



**H**vis cyberangreb kan fabrikere valgre-sultater eller ændre indholdet af live-streamede tv-debatter, er de grund-læggende demokratiske institutioner i fare. Eksperternes advarsler kommer på det rette tidspunkt – hvor vi allerede har set nok til at forstå og dele bekymringerne, men dog stadig er i stand til at sætte beskyttelsesforanstaltninger op. Der er dog andre aspekter af automatiseringens effekter der, i det lange løb, kan være lige så skadelige som sikkerhedstruslen. Her har vi lidt mere tid til at imødegå problematiske udviklinger, men vi har endnu ikke tilstrækkelig klarhed over hvad problemerne vil være.

Vi bør være opmærksomme på, at visse måder at implementere robotapplikationer på har potentiale til at forstyrre vores sociale og kulturelle praksis så fundamentalt at det ændrer selve kernen i det menneskelige liv. Ganske vist har menneskets vel-færd været bredt defineret over tid og sted, og 'menneskets natur' er i høj grad i stand til at tilpasse sig forandringer. Men der er gode grunde til at tro at "automatiseringens tidsalder" udgør en særlig tærskel. Det nye mål for industrien i det 21. århundrede er at placere robotter iblandt os, at skabe nye arbejdsmiljøer med menneske-robot-samarbejde og

nye interaktioner med robotter i vores sociale liv, både i det offentlige rum og i hjemmet. Dette kræver at robotter udvikler "sociale færdigheder".

Det kan faktisk siges at designere af såkaldte "sociale robotter" ikke længere programmerer maskiner, men i stedet skaber kunstige sociale agenter. Uanset om disse kunstige sociale agenter er menneskelige i deres fremtoning eller ej, er vi tilbøjelige til at opfatte "sociale" robotter som reelle interaktionspartnere og reagere på dem med ægte følelser. Det betyder at robot-ingeniører i realiteten kan påvirke den menneskelige kultur ved at producere kunstige agenter som er designet til at engagere mennesket i deres interaktive mønstre.

Hiroshi Ishiguro – samarbejdspartner på vores projekt – hvis robotiske tvillinger er verdenskendte. Foto: ATR Hiroshi Ishiguro Laboratory, Japan.

“  
**Automatiseringens tidsalder rejser et væld af komplekse etiske, sociale, kulturelle, endog metafysiske spørgsmål.**  
”

Robotter kan fremkalde ægte følelser i mennesker – robotsølen *Paro* findes på flere danske plejehjem.

### Humanister søges

Denne udvikling fremtvinger et spørgsmål om hvorvidt robot-ingeniører er uddannet til at være kultur-ingeniører? Robot-ingeniører og dataloger er i dag bevidste om at de skal forholde sig til konsekvenser der rækker langt ud over regulering af sikkerhed. Altså konsekvenser der ligger uden for deres ekspertise. Verdens største faglige sammenslutning inden for ingeniørvirksomhed, The Institute of Electrical and Electronic Engineers (IEEE), arbejder på en ny global standard i forhold til at introducere etik i ingeniøruddannelser på verdensplan.

Men hvem har den nødvendige ekspertise til at rådgive inden for denne nye form for kulturelt design? Til at besvare spørgsmål som de følgende: Skal plejerobotter være menneskelignende? Hvornår skal en menneskelignende robot have køn, og hvornår ikke? Hvordan vil brugen af pædagogiske robotter i skolerne påvirke børns kreativitet og sociale følsomhed? Skal vi tillade sexrobotter? Hvordan påvirker formerne for menneskelig robot-samarbejde vores velfærd? Humanioras og samfundsvidenskaberne ekspertise ligger præcis i denne type spørgsmål, men er på nuværende tidspunkt ikke tilstrækkelig involveret i diskussionen om robot-udvikling.

I 2017 analyserede McKinsey Global Institute "automationspotentiale" af alle delaktiviteter i alle arbejdstyper. Du og dit arbejde har allerede et

"automationspotentiale" som økonomer regner med og som robotindustrien skal realisere. Bør vi ikke i det lys sætte spørgsmålstejn ved den sociale, psykologiske og eksistentielle betydning af arbejde? Samt diskutere hvordan vi kan erstatte denne betydning? Vil det øgede samspil med kunstige sociale agenter – i butikker og restauranter, i taxier og på kundeservicelinjer – forværre eller forbedre vores sociale verden? Når den oprindelige fascination er klinget af, hvordan vil vi så reagere på behovet for at stereotypisere vores sociale handlinger, at tilpasse dem til robotters kunnen? Vil vi miste evnen til at sondere mellem simuleret socialitet og "autentisk" socialitet? Vil vi komme til at give robotter rettighe-der? Kort sagt rejser "automatiseringens tidsalder" et væld af komplekse etiske, sociale, kulturelle og endda metafysiske spørgsmål.

### Socialt, men ikke levende

Ingen er i øjeblikket i stand til at besvare disse spørgsmål. Aktuell forskning i det nye område af "Human-Robot Interaction Studies" er i øjeblikket fragmenteret på grund af begrænsede erfaringer inden for tværfaglig forskning. Den udvikler sig langsomt på grund af de særlige empiriske, metodiske og teoretiske spørgsmål der melder sig når et helt nyt forskningsområde skal opdyrkes på tværs af meget forskelligartede traditioner. I modsætning





hertil udvikler robotindustrien sig hurtigt i en reaktion på stigende muligheder og efterspørgsel fra markedet. Denne forskel mellem forskningens og industriens tempo skaber en blokering der delvis kan sammenlignes med det velkendte Collingridge-dilemma, som mange nye teknologier plages af. Vi kan ikke vurdere den kulturelle virkning af en given teknologi før den er meget udbredt. Men vi kan heller ikke vente på resultaterne af langsigtede undersøgelser hvis mennesker potentielt kan lide skade, og teknologien ikke kontrolleres. Udviklingen af sociale robot-teknologier tilføjer endnu en dimension til dilemmaet. I øjeblikket kan vi ikke engang beskrive hvad vi ønsker at undersøge da vi ikke har begreber til at beskrive noget der er socialt, men ikke levende. At påstå at en robot kan "genkende" dig eller "reagere" på din kommando eller endda være din "ledsager", er at bruge meget vildledende metaforer. Vi kan ikke vurdere hvad vi ikke har konceptualiseret, og vi kan ikke regulere det vi ikke har vurderet.

Carlsbergfondets Semper Ardens-projekt "What Social Robots Can and Should Do - Towards

Integrative Social Robotics" udforsker en ny metode til at belyse og løse ovenstående problemstilling – og samtidig vise hvordan vi kan udvikle nye sociale robot-applikationer ansvarligt. Projektet har deltagelse af 25 forskere fra en bred vifte af discipliner: filosofi, robot-udvikling, antropologi, datalogi, pædagogisk videnskab, lingvistik, psykologi, neuropsykologi, medievidenskab, ledelse, sociologi og konfliktstudier. Metoden "Integrative Social Robotics" som vi er i gang med at udvikle, kan opsummeres i syv kerneprincipper (se Faktaboks).

Integrative Social Robotics har som overordnet mål at sikre at vi bevarer vores etiske og moralske værdier, såsom værdighed og retfærdighed, men også sociale og kulturelle værdier som autonomi, fleksibilitet, autenticitet i mødet mellem mennesker og robotter. Som mulig sidegevinst kan vi måske igennem social robotics også forbedre måden hvorpå disse værdier realiseres i samfundet. I samarbejde med ATR Hiroshi Ishiguro Robotics Laboratory i Japan og AU Social Robotics Lab undersøger vi for øjeblikket flere applikationer der skal øge vo-

Sociale robotter spreder de kulturelle og etiske forestillinger af deres skabere; Robot Sophia og Dr. Ben Goertzel.

### Principper for Integrative Social Robotics

(1) Princippet om at bruge alt relevant ekspertise: Vores sociale verden er den mest komplekse virkelighed vi kender. Da social robot-udvikling griber ind i vores sociale verden, skal hele udviklingsprocessen (forskning og design) frem mod en robotapplikation derfor involvere relevante eksperter fra alle discipliner. Især skal eksperter fra humaniora involveres i at vurdere robotters indvirkning på menneskelige værdier og trivsel.

(2) Princippet om at holde sig det rigtige mål for øje – det handler ikke om robotter: Erkend at de primære produkter af social robot-udvikling ikke er robotter, ikke objekter, men i stedet sociale processer – menneske-robot interaktioner. Beskriv altid sociale interaktioner ud fra tre perspektiver: (I) hvordan de opleves for den person der interagerer med robotten, (II) hvordan de er designet og programmeret fra teamet der udvikler applikationen, og (III) hvordan de opleves af samfundet i forhold til vores sociale og kulturelle normer.

(3) Princippet om værdi-baseret udvikling: Alle forsknings-, design- og udviklingsprocesser bør reguleres af en "non-replacement norm": *Robotter må kun gøre hvad mennesker ikke kan, men skal gøre.* I starten af hver proces identificeres en værdi hvis realisering i den givne sammenhæng er ukontroversiel. I udviklingsprocessen skal det sikres at denne værdi realiseres med kontinuerlig tilbagemelding i hvert trin i forsknings-, design- og udviklingsprocessen.

(For de resterende fire principper se Seibt, Damholdt og Vestergård 2018.<sup>8</sup>)



***For at forme den måde hvorpå robotrevolutionen vender, er humaniora nødvendig som aldrig før. At bidrage til udviklingen af robotrevolutionen på en kompetent og ikke mindst proaktiv måde er ikke alene en spændende ny udfordring for humaniora, det er også en krævende forpligtelse. Det kræver at humaniora genopfinder sig selv og tilføjer en ny proaktiv og konstruktiv rolle til den traditionelle reflekterende fortolkning af samfundet.***



res viden om opfattet retfærdighed samt følelsesmæssig og videnskabsmæssig autonomi.

#### Verdens største konference i humanistisk robotforskning

For at gøre offentligheden opmærksom på de omfattende sociale og kulturelle konsekvenser af udviklingen af sociale robotter, har vi etableret en konferencereserie i "robophilosophy" ([www.robophilosophy.org](http://www.robophilosophy.org)). Siden 2014 har vi hvert andet år organiseret events, som har udviklet sig til verdens største møder inden for tværfaglig humanistisk forskning

i Human Robot Interaction. Det er hovedsageligt forskningskonferencer, men vi har også inviteret politiske beslutningstagere på europæisk og nationalt plan for at overbevise dem om at fremtidige robot-applikationer bør udvikles af tværfaglige forskningsteams med deltagelse af alle interessenter.

"Automatiseringens tidsalder" fordrer at nationale og internationale politiske beslutningstagere, fonde og forskere inden for humaniora påtager sig en ny form for social ansvarlighed. Ansvar for at forme robotrevolutionen på måder der passer til vores samfundsmæssige værdier.

Psykologen Sherry Turkle har meget passende beskrevet begyndelsen af det 21. århundrede som "the robotic moment" – et afgørende vendepunkt i den menneskelige historie: "the way we contemplate [robots] on the horizon says much about who we are and who we are willing to become."

At bidrage til denne udvikling på en kompetent og ikke mindst proaktiv måde er ikke alene en spændende ny udfordring for humaniora, det er også en krævende forpligtelse. Det kræver at humaniora genopfinder sig selv og føjer en ny proaktiv og konstruktiv rolle til den traditionelle reflekterende fortolkning. Denne nyorientering kræver nye tværfaglige uddannelser og især nye grænseflader mellem uddannelser som fx ingeniørvidenskab og humaniora. Vi håber at regeringer verden over reagerer på IEEE's globale initiativ og lignende opdrag som "responsible robotics"<sup>7</sup> og erkender omfanget og den historiske betydning af deres nuværende forpligtelser i uddannelsespolitikken.

Noter

1 Brundage et al 2018, <https://arxiv.org/abs/1802.07228>

2 [https://standards.ieee.org/develop/indconn/ec/autonomous\\_systems.html](https://standards.ieee.org/develop/indconn/ec/autonomous_systems.html) 3 <https://www.mckinsey.com/global-themes/europe/a-future-that-works-the-impact-of-automation-in-denmark> 4 Se artikel 59 af European Parliament Resolution 20170210IPR61808, februar 2017, afvist fra European Commission i april 2017. 5 Se J. Seibt (2016), "Integrative Social Robotics: A New Method Paradigm to Solve the Description and the Regulation Problem?" in:

Seibt, J., Nørskov, M., Schack Andersen, S., (eds.) What Social Robots Can and Should Do: Proceedings of Robophilosophy 2016/TRANSOR 2016 (Amsterdam: IOS Press) pp. 104-115.

6 S. Turkle (2011), Alone Together (New York: Basic Books), p. 26. 7 Se f.e. Foundation Responsible Robotics, [www.responsiblerobotics.org](http://www.responsiblerobotics.org) 8 Se Seibt J., Damholdt, M., Vestergård, C. (2018), "Integrative Social Robotics: Interdisciplinary or Transdisciplinary?", forthcoming in: Coeckelbergh et al (eds.) Envisioning Robots in Society - Power, Politics, and Public Space. Proceedings of Robophilosophy 2018 (Amsterdam: IOS Press).

Artiklen er oversat fra engelsk af postdoc Christina Vestergård.



Den kontroversielle filosofiske position "funktionalisme" påstår, at alt hvad vi oplever, selv vores følelser, kan oversættes i en fuldstændig funktionel beskrivelse. Hvis funktionalismen har ret, så burde alt vi oplever kunne simuleres i en maskine.

Bygger vi menneskelignende robotter for at forstå os selv?

Vi har defineret et nyt felt for interdisciplinær anvendt forskning i filosofi: "Robophilosophy: philosophy of, for, and by social robotics".

