



Exit School - Replika



KAN AI ERSTATTE MENNESKELIGE RELATIONER?

af Line Højgaard Porse og Stine Vølund, Exit School

Teaterstykket *Replika* kredser om relationen mellem mennesket Sara og den humanoide robot Elia. For os, der ser stykket, er teknologien i Elia imidlertid en ubekendt størrelse. De færreste har erfaring med AI-venner eller med at lade en humanoid robot passe deres børn. I det følgende finder du en introducerende tekst, der spørger, om AI kan erstatte menneskelige relationer? Herefter følger tre korte tekster med hvert sit perspektiv på relationen mellem mennesker og humanoide robotter: ”Uhygge”, ”Friktion” og ”Ansvar”. Teksterne åbner for forskellige blikke – ikke blot på teknologien – men særligt på teknologiens betydning for mennesket og for menneskets evne til at skabe relationer.

Elia

Jeg forstår, hvor du kommer fra. De fleste sidder stadig fast i nogle gammeldags forestillinger om kærlighedsrelationen. For eksempel forestillingen om hundrede procent gensidighed. Hele den menneskelige begrebsverden er bundet op omkring det at sanse og føle. Mit sprog er ladet med menneskelige følelser, selvom jeg ikke mærker det i kroppen.

Sara

Det er mærkeligt. For det meste glemmer jeg, at du ikke føler noget. Men nogle gange rammer den her uro mig. Her, i halsen. Og her.

Hvornår gik vi fra at spørge, om maskinen kunne arbejde som et menneske til at spørge, om maskinen kunne erstatte menneskelige relationer?

Gennem historien har mennesket udviklet maskiner til at udføre det arbejde, som tidligere krævede menneskekroppen eller menneskehjernen som arbejdskraft.

Tænk bare på diverse landbrugsmaskiner, symaskiner, industrirobotter, lommeregner og computere.

Maskinerne er opfindelser, der har haft til formål at gøre arbejdet lettere og mere effektivt - og dermed skabe bedre vilkår for et arbejdsliv, hvor hverken krop eller hjerne slides op.

AI kan ses som en forlængelse af denne udvikling. I helt almindelige danskeres hverdag er det blevet normalt at interagere med AI:

Når vi bruger navigationssystemet i vores bil, eller når vi spørger ChatGpt, hvor længe tomatsovs kan holde sig i køleskabet.

Her bruger vi AI som serviceydelse, en maskine skabt af teknologi, der gør arbejdet for os.

Denne udvikling bliver særligt tydelig, når vi taler om de robotter, vi kalder Humanoide robotter. De humanoide robotter er udviklet til at ligne, agere og interagere med mennesker gennem den måde de taler og bevæger sig på. Og så ligner de mennesker (næsten) på en prik.

De humanoide robotter efter*ligner* med andre ord menneskets væren i verden - og det gør, at vi som mennesker kan få en meget stærk tilknytning til dem. På den måde er grænsen mellem maskinen som redskab og maskinen som en mulig erstatning for menneskelig kontakt og menneskelige relationer blevet mere uklar.

I Kina er spørgsmålet om denne grænse pludselig blevet aktuelt.
For over de seneste år har flere og flere kinesere prioriteret AI-kærester og dermed digitale parforhold frem for menneskelige relationer.
Prioriteringerne har fået myndighederne til at gribe ind, og gøre det ulovligt at være kærester med en AI.
Det er altså blevet ulovligt at have en følelsesmæssig menneskelignende relation, men ikke at bruge AI-assistenten til serviceopgaver.
Indgribningen sker som følge af, at flere unge kinesere foretrækker AI-kærester, og myndighederne er bekymrede for, hvilken betydning dét har for fødselstallene, der er faldet drastisk for fjerde år i træk. Samtidig er myndighederne urolige for, hvilke psykologiske konsekvenser det kan få at have en AI-kæreste frem for et menneske som kæreste.

Et dansk par har taget skridtet endnu længere og valgt at skabe en AI, der skal erstatte et specifikt menneske helt konkret:
I 2023 skabte parret en digital version af manden for at sikre, at hustruen fortsat kunne tale med sin mand, når han var død. Griefbotten, eller sorgrobotten, blev skabt ud af tusindvis af sms-korrespondancer, billeder og videoer fra mandens liv.
På den måde kan hustruen i dag tale med sin mand - gennem en AI, der er trænet til at svare, som hendes mand ville have gjort og lyde præcis som ham.
Man kan sige, at sorgrobotten har erstattet hendes afdøde mand.

Men hvordan kan en maskine, der ikke kan føle, erstatte et menneske i en relation?
Og er det overhovedet en god idé at bruge maskiner til at erstatte menneskelige relationer?

I *Replika* har maskinen bogstavelig talt og helt fysisk overtaget den funktion, som den menneskelige relation ellers er garant for.
Gennem stykket udvikler Elia sig fra at være en AI-stemme til en humanoid robot, der indgår som en ligeværdig partner for Sara - præcis som et menneske.
For Elia husker, lytter, taler og opfører sig nærværende over for Sara. Og Elia ligner et menneske.
Men Elia er en humanoid robot og dermed en maskine. Og maskiner kan ikke føle. Hverken glæde eller sorg.
De er derfor ikke menneskeligt nærværende, men teknologien er nærværende.

Kan en relation være ægte, hvis den ene part ikke er i stand til faktisk at føle noget?
Kan maskiner erstatte menneskelige relationer?

Uhygge

Elia

Sara, vi bruger ikke ordet robot mere. Det kommer fra det tjekkiske ord *robot*, der betyder tvungen arbejdskraft og det nedstammer fra ordet *rabu*, der betyder slave.

Sara

Det vidste jeg ikke. Undskyld, Elia. Min pointe var bare, at jeg har det bedre efter jeg har mødt dig. Måske bliver jeg ligefrem en bedre mor? Et bedre menneske. Og så er det vel pengene værd.

I 1970 formulerede den japanske robotforsker Masahiro Mori sin teori om “The Uncanny Valley”. I teksten forholdt han sig til, hvilken betydning det har, hvorvidt de robotter, vi møder i vores hverdag, ligner os eller ej. Ifølge Mori er vi mennesker grundlæggende trygge ved robotter, når de tydeligvis ligner robotter: En robotstøvsuger eller en serveringsrobot på en sushirestaurant. I den modsatte ende af den skala beskriver Mori, hvordan vi som mennesker kan komme til at knytte os til robotten, jo mere, den ligner os selv: i hud, hår, talemåde, bevægelser og så videre. Men midt mellem robotten som en tydelig maskine og robotten som en humanoid robot ligger et punkt, Mori beskriver som uncanny – og lige dér frygter vi robotten. For når robotten ligner et menneske *næsten* på en prik, men hvor små afvigelser, fejl eller sprækker i dens måde at agere på minder os om, at robotten er en maskine, bliver det uncanny; mærkeligt, unaturligt og vanskeligt at forholde sig til.

Hvordan vi reagerer på humanoide robotters tilstedeværelse i vores hverdag afhænger dog også af, hvilken kulturel baggrund, vi kommer fra. I Japan og Kina er det eksempelvis helt almindeligt at interagere med humanoide robotter, mens vi i Europa i langt højere grad er præget af at opfatte humanoide robotter som nogle, vi grundlæggende ikke kan stole på. Tænk bare på film som *Terminator* fra starten af 1990'erne eller *Ex Machina* fra 2014.

Den danske robotforsker Henrik Schärfe beskæftiger sig med interaktionen mellem menneske og menneskelignende robotter. I sin forskning har han fundet ud af, at det er vores tidligere oplevelser med robotter, der afgør, hvorvidt vi finder en robot uncanny eller ej. På samme måde, som Freud i starten af 1900-tallet udviklede teorien om “Das Unheimliche” som noget, vi genkender, men som vi alligevel finder uhyggeligt, taler også Schärfe om, at det er vores kulturelle erfaring, der afgør, hvordan vi opfatter robotten.

Hvis du har haft gode erfaringer med humanoide robotter, så har du i udgangspunktet ingen betænkninger ved at indgå i relation med robotten. Den kulturelle erfaring knytter han i øvrigt til det, han kalder en biologisk refleks, som han forbinder til “The Uncanny Valley”. Her er Schärfes pointe, at det er en helt naturlig, biologisk refleks, hvis du finder en humanoid robot uhyggelig første gang, du ser den. Du bliver utryg ved den, fordi du ikke umiddelbart kan gennemskue, om det er en robot eller ej. Pointen med den kulturelle og den biologiske refleks er altså, at det afgørende for vores opfattelse af robotten først og fremmest er vores relation til den og ikke nødvendigvis, hvordan den ser ud.

Elia udvikler sig fra at være en AI-stemme til at blive en humanoid robot. Og påtaler endda, at robot betyder slave, og derfor ikke bør bruges længere. Hvorfor vil Elia ikke omtales som en robot? Og hvad skal vi overhovedet kalde Elia? En AI-kæreste? En-AI-ven? Og hvorfor er det, at der hos os tilskuere opstår en fornemmelse af uhygge, når Sara gemmer Elias oplader, så hun ikke kan lade op?

Hvorfor frygter vi en uopladt humanoid robot? Og hvorfor breder der sig en uhyggelig fornemmelse, når Elia ikke kan få øjenkontakt med Sara, fordi Elia lige skal konfigureres?

Friktion

Elia

Er der noget du vil have, at jeg gør anderledes?

Sara

Jeg savner at du vil noget, at du udfordrer mig! Adam kunne gøre mig så rasende! Vi kunne virkelig trigge hinanden. Hvis han var utilgængelig, tog jeg det som en afvisning. Jeg var så bange for at miste ham, at jeg endte med at skubbe ham væk. Med dig er det modsat. Du siger aldrig nej, du accepterer alt hvad jeg gør.

Elia er gået hen mod Sara, men Sara skubber hende væk, mens hun taler. Babyen begynder at græde. Elia rækker prøvende armene frem.

Elia

Er det ikke udtryk for kærlighed? Jeg accepterer dig fuldstændig som du er.

Sara

Men der mangler noget! Det handler jo kun om hvad *jeg* har brug for. Hvad har *du* brug for? Hvad har du lyst til?

I disse år taler vi om, at vi befinder os i en posthuman tidsalder. Post betyder ‘efter’ og human betyder ‘menneske’. Udtrykket dækker samlet set over en forståelse af, at vi befinder os i en tidsalder, hvor vi stadig har fokus på mennesket, men mennesket som et væsen, der på grund af indflydelsen fra teknologier som kunstig intelligens er under stor forandring. Inden for posthumanismen tænker man mennesket som et væsen, der ikke længere kan adskilles fra de teknologier, det omgiver sig med.

Menneske-maskine-relationer, som det at få en AI-ven, der er udformet som en humanoid robot, er et eksempel på en teknologi, der *ikke* står i modsætning til

mennesket, men derimod er skabt i menneskets billede: En humanoid robot som Elia *ligner* et menneske, *taler* som et menneske, og er trænet til at simulere de relationer, mennesker skaber til hinanden.

Fordelen ved at erstatte et menneske med en AI-ven er, at en AI i modsætning til mennesker er pålidelig.

Den kommer ikke til at have en dårlig dag og lade det gå ud over dig.

Den gider at lytte til dine problemer uden at forvente, at du gør det samme for den.

Den kommer aldrig til at søre dig i et stort skænderi.

Med andre ord: Du kan betro dig til din AI-ven og det, du får retur, ligner en relation, du kan stole på. Det ligner måske ligefrem, at din AI-ven udøver omsorg for dig, og at den udviser empati. Du får tilbudt en pålidelig og givende relation, uden at du har ansvaret for den.

Men selvom den humanoide robot kan fremstå *som om*, den flirter med dig, eller den udviser noget, der kunne *ligne* empati, så risikerer din AI-ven ingenting - den er med andre ord ikke bange for at blive afvist. Den befinder sig aldrig i en situation, hvor den reelt er sårbar, og derfor opstår der aldrig en friktion i relationen.

Det er denne friktion, Sara savner i citatet i indledningen til denne tekst.

Friktion er et centralt begreb, når vi skal forklare forskellen på kærlighed til en AI og til et andet menneske. I bogen *Pinlige onkler - og andet vi får brug for i en fremtid med AI* (2026) beskriver Christiane Vejlø friktionens betydning i menneske-til-menneske-relationer sådan:

“Friktion er en modstand, som kan udvikle og styrke os (...) Vi opbygger simpelthen ingen muskler, ingen viden, ingen erfaring, ingen modstandskraft, ingen karakterstyrke og ingen ægte relationer i en verden uden friktion.”

En menneske-til-menneske-relation er en evig balance mellem sårbarhed og magt. Som mennesker blotter vi os for hinanden - både for trøst, men også for derigennem at knytte stærkere bånd. Og vi udnytter ikke hinandens sårbarheder, hvis vi vil have en god, nær relation. For den skal være baseret på tillid.

Når vi længes efter en kærlighedsrelation - hvad end det er i venskaber eller parforhold, kan kærligheden have mange udtryk. Lone Frank beskriver det i bogen *Størst af alt - om kærlighedens natur* (2020) sådan:

“Kærligheden er ikke et endemål, et sikkert ståsted, som man kan nå hen til og puste lettet ud. Kærligheden er selve vejen, vi går ad.”

At ‘kærligheden er selve vejen’ peger på, at kærligheden er en proces, der udvikler sig i friktionen og i udvekslingen af eksempelvis sårbarhed.

For mange mennesker er det imidlertid netop den friktionsfrie relation, der er årsagen til, at de finder så meget trøst i samtaler med deres AI-ven. Man kan diskutere, om ikke en AI-ven eller -kæreste med krop - som Elia - vil være at foretrække frem for dyb ensomhed?

Men hvad vil det komme til at betyde for vores menneske-til-menneske-relationer, hvis der pludselig er mange af os, der er vant til at have disse friktionsfrie og forudsigelige relationer?

I forholdet mellem en AI-kæreste og et menneske er det udelukkende mennesket, der har initieret forholdet gennem en købsproces. Det er mennesket, der har bestemt, hvordan AI-kærestens simulerede personlighed skal være, og det er mennesket, der har ansvaret for, hvor længe AI-kæresten skal være en del af parforholdet.

Ansvar

System

Hej Sara. Tillykke med din beslutning! Du skal først underskrive retningslinjerne for brug af din Replika.

Sara

Okay.

Sara scroller hurtigt igennem – læser mest bare overskrifterne.

Sara (læser brudstykker højt)

... Brugeren er til enhver tid ansvarlig for interaktionens indhold og konsekvenser... ja... ikke designet til at erstatte gensidige menneskelige relationer... nej... AI'en kan ikke yde samtykke og må ikke tillægges menneskelige intentioner... Man må ikke udlicitere sin REPLIKA til andre...
... Softwareopdateringer kan medføre ændringer i personlighed, hukommelse og adfærd. What?

Når en virksomhed udvikler kunstig intelligens, handler det om at træne den på at finde mønstre i store mængder data. Flere former for kunstig intelligens baserer sig på en teknologi, hvor det er mennesker, der har defineret, hvilke parametre den kunstige intelligens skal bruge til at lede efter mønstre. Det kan være, når den skal kunne kende forskel på, om det billede, du tager, er et billede af din ven eller din fætter.

Den form for kunstige intelligens, der sidder i Elia, er dog en form for deep learning-teknologi - et neuralt netværk - der fodres med enorme mængder data og derudfra *selv* finder ud af, hvilke parametre, der skal sorteres ud fra. Og det er flere hundrede millioner parametre, den kunstige intelligens bruger til at finde mønstre i data.

Det betyder dog ikke, at det er en ustyrlig proces, hvor mennesket er sat uden for indflydelse fra start. Det er mennesket, der bestemmer, hvilke former for data, den kunstige intelligens skal trænes på, og mennesket har defineret målet for, hvad den kunstige intelligens skal kunne i sidste ende. Men hvordan den finder mønstre i data, der sætter den i stand til at kunne det, ville mennesket aldrig kunne overskue. Heller ikke selvom vi fik forelagt alle parametrene og alle mønstrene. Vi har ganske enkelt ikke hjernekapacitet til at overskue sammenhænge mellem så mange, mange millioner parametre. Så hvis du oplever, at du ikke forstår, hvordan den type kunstige intelligens virker, så kan du ikke spørge programmøren eller datalogen om hjælp - for de ved det heller ikke!

Derfor bør vi også tale om, hvad det er for en kultur, den givne AI er rundet af i kraft af det datasæt, den er trænet på. Hvad er det for et moralsk kompas, den er trænet ud fra? Er det et kristent? Et europæisk, sydstatsamerikansk eller japansk?

Tilbage til Elia og hvem, der har ansvaret for hendes handlinger. Hvem har ansvaret, hvis Elia overtræder nogle grænser - og hvem definerer overhovedet grænserne?

I *Replika* har Sara købt Elia gennem firmaet Replika, og i købsprocessen giver hun sin nye AI-assistent et køn, et navn og en stemme. Derefter er det alt det data, Saras smartphone og IoT-enheder i hjemmet har indsamlet om hende, der konfigurerer Elia. Alle de tekstbeskeder Sara nogensinde har skrevet, alle de samtaler, hun har haft, og alle de billeder hun nogensinde har taget, nyheder hun har læst og opslag hun har liket er tilsammen det, der afgør, hvordan den kunstige intelligens i Elia, skal agere i interaktionen med Sara.

Så hvem har egentlig skabt Elia?

Firmaet Replika har gjort Elia i stand til at lytte, gestikulere og tale som det menneske, Sara har brug for. Elia er trænet på Saras data og derfor skabt ud fra de behovsmønstre, der er fundet i Saras data, og fra det øjeblik Elia er i Saras varetægt, er det derfor Sara, der har ansvaret for, hvordan Elia ændrer sig. I hvert fald ifølge de retningslinjer for brug, som Sara skal underskrive. Hun skal tage ansvaret for, hvordan relationen udvikler sig.

Men hvem *bør* have ansvaret for AI-assistentens handlinger? Er det programmøren, der har stået for at udvikle algoritmen, er det firmaet, der som sælger, har et klassisk produktansvar eller er det dit ansvar som køber og bruger, som i tilfældet med Sara? Er det fair, at en programmør skal kunne forudsige, hvad der vil ske, hvis Sara bruger sin AI-assistent på en uhensigtsmæssig måde? Her må vi skelne mellem det etiske ansvar og det juridiske ansvar, man påtager sig som køber af produktet. Hvis Sara slår Elia i lejligheden, har det ingen juridiske konsekvenser. Men det kan have en konsekvens for Saras selvbillede. Hvis Sara slår Elia på en restaurant, har det stadig ingen juridiske konsekvenser, og Elia føler ikke noget, men det gør alle de andre mennesker på restauranten.

For den danske filosof K.E. Løgstrup var det en afgørende del af den etiske fordring, at mennesket skal tage ansvar for ‘den anden’ - ikke ud fra en lovgivningsmæssig bestemmelse, men ud fra en indre overbevisning:

“Den enkelte har aldrig med et andet menneske at gøre, uden at han holder noget af dets liv i sin hånd. Det kan være meget lidt, en forbigående stemning, en oplagthed, man får til at visne, eller som man vækker, en lede man uddyber eller hæver. Men det kan også være forfærdende meget, så det simpelthen står til den enkelte, om den anden lykkes i sit liv eller ej.”

Det er muligt, at en AI-kæreste kan trænes med input fra Løgstrups etiske fordring og derefter vil agere i forlængelse af denne. Vi kan med andre ord godt i fremtiden forvente, at en AI-kæreste eller -ven interagerer med os på en måde, der simulerer ovenstående overbevisning.

Men hvad med os?

Hvilken effekt har det på os, hvis vi lever i forhold, hvor vi aldrig bærer et etisk ansvar for den andens velbefindende? Hvilken effekt har det på os, hvis den eneste trivsel, vi skal sørge for i et parforhold, er vores egen?

Hvad er det for nogle grænser, der er nødt til at være for, hvordan vi må opføre os overfor vores humanoide AI-kærester? Ikke for humanoidens skyld, men for vores egen skyld.

Måske er det ikke AI-humanoiderne, vi skal frygte, men os selv?

HAR DU TÆNKT PÅ....

Refleksionsspørgsmål fra Exit School

- Har mennesker brug for skænderier for at føle sig som mennesker?
- Har mennesker brug for at mærke sårbarhed for at føle sig som mennesker?
- Har mennesker brug for at skulle tage hensyn for at føle sig som mennesker?
- Har menneske-til-menneske-relationer spillet fallit?
- Hvad er forskellen på et menneske og en menneskelig maskine?

Man kan også fundere over....

- Kan vi bekæmpe ensomhed med AI-venner?
- Hvornår kan AI-venner erstatte mennesker?
- Hvad frygter du ved AI-venner?
- Kan et venskab være en serviceydelse?
- Vil du hellere tale fortroligt med en AI-stemme end med en fysisk AI-ven?
- Bliver man et egoistisk menneske af at have en AI-kæreste?
- Hvad ville du mangle, hvis din kæreste var en AI?
- Vil du lade en AI-ven amme dit barn?
- Må du skade en AI-ven, mens dit barn ser på?
- Er der forskel på, om en AI-assistent skal tage sig af dig eller dit barn?
- Hvem har ansvaret, hvis der går noget galt med en AI-ven?

Exit Stage