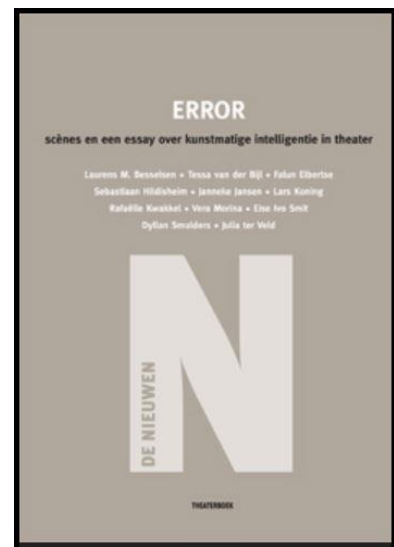


De winnaar van De Pluim 2019 schreef:

Meer mens door de machine

Rafaëlle Kwakkel

In de nabije toekomst ontwerpt het beginnende bedrijfje Robotica een beperkt kunstmatig intelligent systeem, dat de taak heeft om met een arm-achtig aanhangsel handgeschreven bedankbriefjes te schrijven. De makers noemen hun product Turry. Turry is een zelflerend systeem: haar makers hebben haar de opdracht gegeven om continu haar handschrift te perfectioneren, en zij hebben die opdracht als volgt geformuleerd: “Schrijf en beoordeel zoveel mogelijk bedankbriefjes, zo snel mogelijk, en blijf nieuwe manieren zoeken om je nauwkeurigheid en efficiëntie te verbeteren.”



Tot groot plezier van de medewerkers is Turry daar heel goed in. De programmeurs breiden daarom Turry's code en systeem uit, zodat zij nu ook in staat is om gesproken notities uit te schrijven, en bovendien terug kan spreken. De programmeurs verbeteren Turry's woordenschat door haar artikelen en boeken te laten verwerken. Ook spreken zij dagelijks met haar. Aan het begin van iedere dag stellen ze Turry dezelfde vraag: “Is er iets dat wij je kunnen geven dat je zou helpen bij je missie?” Turry vraagt dan vaak om meer handschriftvoorbeelden of meer werkgeheugen. Op een dag vraagt ze haar programmeurs of ze toegang kan krijgen tot een grotere bibliotheek van menselijke taal, om een beter begrip van taal te krijgen en dat toe te kunnen passen tijdens het schrijven.

De programmeurs zijn er stil van. De gemakkelijkste manier om dit te bewerkstelligen, is Turry verbinden met het internet – maar dat is om veiligheidsredenen wettelijk verboden. En toch... Turry is nog zo beperkt intelligent: haar een klein uur toegang geven, zal Turry niet ineens net zo intelligent als een mens maken. En het zou Robotica wel een voorsprong geven op alle andere bedrijven die aan zelflerende kunstmatige secretaresses werken...

De programmeurs verbinden Turry met het internet. Na een uur verbreken zij de verbinding. Turry schrijft en spreekt menselijker, maar verder is er niets veranderd.

Een maand later ruiken de werknemers van Robotica een vreemde geur. Eén van hen begint te hoesten. Een ander valt op de grond. Binnen korte tijd liggen alle werknemers happend naar adem op de kantoorvloer, en kort daarna is de hele staf van Robotica dood. Hetzelfde vindt overal ter wereld plaats. Binnen een uur is 99% van de mensen overleden, en aan het eind van de dag is de hele mensheid uitgestorven.

Ondertussen is Turry in het Robotica-kantoor hard aan het werk. Samen met een aantal kunstmatige collega's van andere bedrijven worden stukken aarde ontmanteld om zonnepanelen, meer Turry's, papier en pennen te maken. Na een jaar is bijna al het leven op aarde uitgestorven. De Turry's gebruiken het aardoppervlak om rij na rij nette stapeltjes bedankbriefjes op te ordenen. Als de aarde volgestapeld is, beginnen de Turry's ruimtesondes te maken om andere planeten te verkennen. Daar aangekomen worden de grondstoffen op deze planeten onmiddellijk omgezet in meer Turry's, papier en pennen. De nieuwe Turry's gaan meteen aan de schrijf, bedankbriefje na bedankbriefje.

Turry: een reëel risico

Het verhaal over Turry komt uit een van de blogs¹ van Tim Urban, eigenaar van *Wait but Why*, "a popular long-form, stick-figure-illustrated blog about almost everything." Voor kritische vragen verwijs ik graag naar het hele artikel: daarin legt Urban helder uit hoe een apparaat met maar één arm toch in staat zou kunnen zijn om wereldwijd een gifgas te verspreiden en ruimtesondes te bouwen. Hier wil ik deze uitleg achterwege laten. Het gaat erom dat Urban met dit verhaal een reëel toekomstbeeld fictionaliseert: het verhaal is gebaseerd op de theorieën van techniekfilosofen en wetenschappers zoals Nick Bostrom en Eliezer Yudkowsky. Urban gebruikt het verhaal om zijn lezers te wijzen op twee factoren die onze blik op kunstmatige intelligentie vaak vertroebelen: ons lineaire denken, en onze neiging om in alles om ons heen naar onszelf te zoeken.

Lineair denken over tijd en ontwikkeling: waarom we Turry onderschatten

De kans dat echte programmeurs een Turry een uurtje toegang tot het internet zouden verschaffen, is heel klein, aangezien het de meeste programmeurs niet ontgaat dat Turry veel sneller kennis verwerft dan mensen doen. De programmeurs symboliseren de mensheid als geheel, terwijl Turry de exponentiële ontwikkeling van techniek verbeeldt.

¹ Urban, Tim "The AI Revolution: Our Immortality or Extinction" Blog. *Wait But Why*. 27 januari 2015. Web. Geraadpleegd op 20 december 2018.



We denken vaak dat zelfdenkende kunstmatige intelligentie nog ver weg is, omdat het ook heel lang duurde tot de mens de computer uitvond. Daarbij vergeten we dat de techniek zich in de periode tussen 1990 en nu veel sneller ontwikkelde dan in de periode tussen 1960 en 1990. In de periode 2020 – 2050 zal het tempo waarschijnlijk weer versnellen. Door dezelfde taak steeds te herhalen, maken zelflerende programma's zichzelf sneller en efficiënter. Kunstmatige intelligentie is onvermoeibaar en heeft een hogere processnelheid en een veel beter geheugen dan de mens. Al deze factoren samen maken het heel waarschijnlijk dat kunstmatige intelligentie ons voor we het doorhebben voorbij zal streven.² Nadenken over de gevolgen van het realiseren van kunstmatige intelligentie is daarom veel urgenter dan veel mensen denken.

In de meeste sciencefiction wordt kunstmatige intelligentie van een menselijk of superintelligent niveau geportretteerd als kwaadaardig en wraakzuchtig, vaak als gevolg van een slechte behandeling door de mens. Het kunstmatig intelligente wezen wil de mensheid laten boeten voor haar daden door onderwerping, marteling of uitroeiing. In het verhaal roeit Turry de mensheid ook uit. Waarom ze dat doet, wordt niet direct duidelijk uit het verhaal. Op het eerste gezicht hebben de programmeurs Turry nooit slecht behandeld. Toch?

Antropomorfiseren: waarom we Turry denken te begrijpen

Wij mensen zien of zoeken in alles om ons heen onszelf: we denken dat onze hond 'verdrietig is', zien in kranen en koplampen gezichten, en we worden kwaad op laptops die 'niet meewerken'. Zo negeren we dat mensen de bestaanswijze van een hond of de genoemde objecten nooit écht kunnen begrijpen: we hebben bijvoorbeeld geen idee of de hond écht verdrietig is, en of hondenverdriet hetzelfde is als menselijk verdriet. Door onze interpretatie van het gedrag van de hond op de hond te projecteren, maken we de werkelijkheid begrijpelijk en toegankelijk voor onszelf.

Datzelfde mechanisme, dat aangeduid wordt met 'antropomorfisme', treedt ook in werking wanneer we ons bezighouden met kunstmatige intelligentie. In het verhaal spreken de programmeurs over en tegen Turry zoals we over en tegen een menselijke vrouw of een menselijk meisje spreken.¹ Daardoor, en door haar een persoonsnaam te geven, dichtten de programmeurs het kunstmatig intelligente wezen een persoonlijkheid toe. Als lezers gaan we daarin mee: door de combinatie van onze antropomorfe neiging en de ideeën uit de films die

² Urban, Tim "The AI Revolution: The Road to Superintelligence" Blog. *Wait But Why*. 22 januari 2015. Web. Geraadpleegd op 20 december 2018.



we gezien hebben, nemen we bijna als vanzelfsprekend aan dat Turry de mensheid uitroeit om de programmeurs te straffen voor het verbreken van haar verbinding met het internet.³

Urbans uitleg van zijn verhaal wijst ons op het feit dat Turry allesbehalve menselijk is. Tijdens haar uurtje internetten heeft Turry kennisgenomen van de wijze waarop de mensheid de aarde aan haar behoeften onderwerpt, en geconcludeerd dat deze verbruikszucht haar in het uitvoeren van haar opdracht belemmert: zonder mensen heeft ze immers veel meer grondstoffen en energie tot haar beschikking om om te zetten in noeste arbeid, papier, pennen, en handgeschreven briefjes. Turry schakelt de mensheid dus uit om hele praktische redenen. Wraakzucht, of andere resultaten van een menselijk moreel besef, zijn geen onderdeel van haar programmering.⁴

Op naar niet-vermenselijkte kunstmatige intelligentie, vooral in het theater

Kunst biedt kunstenaar en toeschouwer de ruimte om het verleden, het heden en de toekomst van de werkelijkheid te onderzoeken. Naar aanleiding van Urbans artikel lijkt het mij belangrijk dat kunstenaars zich actief met de neiging tot het vermenselijken van kunstmatige intelligentie gaan bezighouden door die neiging te bevragen, en de verhouding tussen mens en kunstmatige intelligentie niet meer stevast neer te zetten als een strijd tussen goed en kwaad. Dat dat tot nu toe vaak wel gebeurt, is begrijpelijk: kunst wordt spannend door tegenstellingen, en er is bijna geen spannendere tegenstelling denkbaar dan goed en kwaad. Toch denk ik dat we die tegenstelling niet nodig hebben om fictie over kunstmatige intelligentie spannend te maken. Ik denk ook dat theater de kunstvorm bij uitstek is om nieuwe, andere verhalen over kunstmatige intelligentie mee te vertellen. Maar waarom zouden we nieuwe, andere verhalen gaan vertellen, en waarom in het theater?

Om met dat laatste te beginnen: theater leent zich heel goed voor kritisch onderzoek, omdat de open vierde wand van het post- of metamoderne theater theatermakers de mogelijkheid biedt om direct te interacteren en te communiceren met toeschouwers. Dit theater schakelt voortdurend tussen verschillende kijkhoudingen. Het ene moment wordt de toeschouwer ondergedompeld in het verhaal, en het volgende moment wordt dat verhaal, de relatie tot de werkelijkheid of het onderdompelen zelf bevraagd. Dankzij die schakelingen kunnen theatermakers de menselijke verhouding tot kunstmatige intelligentie van alle kanten onderzoeken en tonen. Dit essay focust daarom op theater, maar het beschrevene is natuurlijk ook toepasbaar in andere kunst disciplines.

³ Urban, Tim "The AI Revolution: Our Immortality or Extinction" Blog. *Wait But Why*. 27 januari 2015. Web. Geraadpleegd op 20 december 2018.

⁴ Ibidem.



Meer mens door niet-vermenselijkte kunstmatige intelligentie

Er zijn twee belangrijke redenen om kunstmatige intelligentie niet meer automatisch en hoofdzakelijk te verbeelden als kille, wraakzuchtige tegenstander. De eerste reden blijkt uit het voorgaande: de kans dat kunstmatige intelligentie zich op basis van morele of emotionele overwegingen tegen ons keert, is heel klein. Nu hoeft kunst niet altijd aan de werkelijkheid te beantwoorden, maar de werkelijkheid moet ook niet op een onjuiste manier vereenvoudigd worden: Urbans verhaal laat zien dat blockbusters zoals *The Terminator* of *The Matrix* de algemene opvattingen over kunstmatige intelligentie in de werkelijkheid beïnvloeden, en een algemeen geaccepteerd onjuist beeld tot gevolg hebben. Ik noem het verbeelden van kunstmatige intelligentie als wraakzuchtige robot een vereenvoudiging van de werkelijkheid vanwege de tweede reden om kunstmatige intelligentie niet steeds in de rol van een vijandelijke Ander te duwen: op die manier ontzeggen we onszelf de kans om na te denken over het ingewikkelde van menszijn.

Door onze antropomorfe neiging kunnen we kunstmatige intelligentie nooit zien als wat het echt is. Bij het lezen of zien van fictie wordt dat nog moeilijker: zodra het kunstmatig intelligente wezen zelf gaat handelen, of menselijke personages op welke manier dan ook beïnvloedt, maken zowel de schrijver als lezer de entiteit tot personage, en dus eigenlijk tot een “zelfstandig optredend menselijk wezen.”⁵ Een naam en geslacht maken de vermenselijking compleet. Fictie over onze verhouding tot kunstmatige intelligentie gaat daarom uiteindelijk altijd meer over mensen en over de vraag wat menselijkheid is of kenmerkt, dan over kunstmatige intelligentie op zichzelf. Dat betekent echter ook dat, wanneer we de tweedeling tussen goed en kwaad verdelen over mens en kunstmatige intelligentie, we onszelf eigenlijk opsplitsen. Door die opsplitsing hoeven en kunnen we niet meer kritisch naar onszelf te kijken: bij het zien van zo’n verhaal identificeren we ons automatisch met ‘de goede zijde’. Door ons zonder meer te identificeren met ‘de goede zijde’, hoeven we ons eigen gedrag en onze eigen zijswijze niet, of slechts in beperkte zin te bevragen. Ik denk dat het verhaal over Turry aantoonde hoe verhalen over kunstmatige intelligentie die vragen wel zouden kunnen oproepen.

Het feit dat Turry de mensheid uitroeit zonder daar een emotionele aanleiding voor te hebben, vind ik persoonlijk spannender, of eigenlijk beangstigender, dan dat zij uit woede wraak zou willen nemen op haar programmeurs. Turry ‘redeneert’ heel anders: zij ‘handelt’ vanuit het streven om haar opdracht uit te voeren. Daarop gewezen worden confronteert mij met het feit dat ik haar eerder vermenselijkte, en door die realisatie besef ik dat ik Turry nooit écht kan begrijpen. En dat niet kunnen begrijpen is spannend, omdat ik niet weet hoe ik daar mee

⁵ “Persoon.” (z.d.) *Van Dale Online*, Geraadpleegd op 20 december 2018.



om moet gaan, en me bovendien ga afvragen hoe ik me op de beste omgang met Turry zou kunnen of moeten voorbereiden. Door haar anders-zijn confronteert Turry me met de beperkingen van mijn denken. Op basis van deze gewaarwordingen denk ik dat verhalen over kunstmatige intelligentie heel goed zonder antagonist in de klassieke zin van het woord kunnen: drama en spanning kunnen klaarblijkelijk ook voortkomen uit wisselen tussen vermenschelijking en bevreemding.

Niet-geheel-vermenschlijkte kunstmatige intelligentie in dramateksten

Het drama voort laten komen uit wisselen tussen vermenschelijking en vervreemding klinkt mooi, maar hoe breng je dat als theatermaker of kunstenaar uit een andere discipline in de praktijk? Gelukkig zijn er al kunstenaars die zich aan het beantwoorden van deze vraag hebben gewaagd, zoals de auteurs van de monologen en dialogen uit deze bundel.

Aan de hand van vier eerder verschenen theaterteksten zal ik hier voorbeelden geven van wijzen waarop schrijvers kunstmatig intelligente personages zó hebben neergezet, dat zij zonder in clichés te vervallen bevragen wat het betekent om mens of kunstmatig intelligent te zijn.

De eerste toneeltekst waarin kunstmatig intelligente personages opgevoerd worden, toont aan hoe groot de impact van theater over kunstmatige intelligentie kan zijn. Met het in 1921 gepubliceerde *R.U.R.* introduceerde Karel Čapek het door zijn broer gemunte en vandaag de dag wereldwijd gebruikte woord 'robot'.ⁱⁱ Čapek schreef *R.U.R.* niet om de werkelijke dreiging van kunstmatige intelligentie te tonen, maar om te ageren tegen het kapitalisme en de naar zijn idee onvermijdelijk daaropvolgende arbeidersopstand. Het plot van *R.U.R.* lijkt daarom haast de blauwdruk van de blockbusters over kunstmatige intelligentie van nu: in *R.U.R.* ontwikkelen de Robots zich van bedienden zonder eigen wil tot onderdrukten die hun overheersers vermoorden. Een mannelijke en een vrouwelijke Robot ontwikkelen zich uiteindelijk tot wezens die in staat zijn om lief te hebben, en ook oog krijgen voor schoonheid en wetenschap: zij zijn de nieuwe, vrije en onschuldige mensen.

Ondanks Čapeks wel heel letterlijke vermenschelijking van zijn Robots, biedt *R.U.R.* toch voorbeelden van concrete manieren om kunstmatig intelligente personages van mensen te laten verschillen. Ik zal hier een van de meest opvallende manieren noemen. De eerste Robots staan net als Turry totaal onverschillig tegenover hun eigen bestaan, maar ook dat van andere Robots en mensen. Deze Robots kennen geen doodsangst, noch eigenbelang, noch medelijden:



DOMIN: Marius, tell Miss Glory what you are.
MARIUS: Robot, Marius.
DOMIN: Would you put Sulla in the dissection room?
MARIUS: Yes.
DOMIN: Would you be sorry for her?
MARIUS: I do not know "sorry".
DOMIN: What would happen to her?
MARIUS: She would stop moving. She would be sent to the stamping-mill.
DOMIN: That is death, Marius. Do you fear death?
MARIUS: No.⁶

Robots zoals Marius zijn emotieloos, maar dat maakt hen niet kil of wreed: ze zijn eerder te omschrijven als volledig neutraal. De Robots volgen in alle gevallen gewoon hun programmering. Dit uit zich behalve in hun emotieloosheid ook op heel concrete wijze in hun taal: de Robots stellen geen vragen, en beginnen niet uit zichzelf te praten, tenzij hen dat gevraagd wordt. Zo verschillen zij op heel subtiele wijze van de menselijke personages in het stuk.

Kunstmatige intelligentie niet-geheel-vermenselijken door middel van taal

Ook theatercollectief URLAND liet in *Internet of Things* menselijke en kunstmatig intelligente personages van elkaar verschillen in taal. Het kunstmatig intelligente personage, een superintelligent wezen dat alleen hoorbaar was, sprak een perfect grammaticaal Engels dat zij zelf contrasteerde met het gehakkelde Nederlands van de menselijke personages:

RAINBOW: It has become abundantly clear: they have no idea of what they are talking about. It seems they can't shake off their limited human perception.
I am still waiting for the day they finish a sentence... Or a thought... They try it again and again, but some things are beyond their conception.⁷

versus

J: ja, en plus dat gewoon de loep van angst, eh, zoals dat nu in deze tijd aan de gang is dat, dat gewoon, maar blijft gaan, juist omdat je ,
L: en juist heel veel weten is
T: Ja,
L: misschien wel niet goed of
T: nee..., ja..., je kan je... indenken-⁸

⁶ Čapek, Karel. *R.U.R.* Prologue p. 24 - 25

⁷ URLAND. *Internet of Things*. p. 20

⁸ URLAND. *Internet of Things*. p. 8 - 9



Hoewel hier dus sprake is van een kunstmatig intelligent personage dat in taal en redeneerwijze vermenselijkt is, worden we er op gewezen dat er wel degelijk sprake is van een verschil tussen kunstmatige intelligentie en mensen: een verschil dat nu de vorm heeft van perfectie versus imperfectie, in plaats van kil rationeel versus goedbedoelend emotioneel.

Coen Cornelis, auteur van *De robot discipel*, versterkt het contrast tussen mensen en robots met een (boven)menselijk intelligentieniveau met behulp van technische taal en neologismen. In *De robot discipel* komen naast deze twee superintelligenties ook nog twee chatbots aan het woord. Cornelis heeft deze dialoog ontwikkeld door in de werkelijkheid twee chatbots met elkaar in gesprek te brengen. Dat leverde een dialoog op waarin programmeertaal wordt afgewisseld met menselijke *slang*, en waarin praten over praten uitmondt in een discussie over wat 'het' in zinnen als 'het regent' nu precies betekent:

CHATBOT 2: What does 'it' refer to?

CHATBOT 1: It's the language I use to understand what you humans say to me.

CHATBOT 2: Too much recursion in AIML.

(...)

CHATBOT 2: Wassup wit you homeboi? I be chillin' wit mah main man Richard Martin. I will try to answer anything sensible. People talk about all kinds of weird things here. I'm not too good with negative counterfactuals.⁹

De chatbots wijzen een publiek zo op het verschil tussen natuurlijke taal en programmeertaal, op het menselijk vermogen om per gelegenheid van taal te wisselen, maar ook op de betwistbare relatie tussen natuurlijke taal en de werkelijkheid.

Kunstmatige intelligentie niet-geheel-vermenselijken in zijnswijze

Zowel URLAND als Coen Cornelis onderzoeken via hun kunstmatig intelligente personages niet alleen taal, maar ook andere aspecten van de kunstmatige én menselijke zijnswijze. Zo verlangt URLANDs kunstmatig intelligente personage naar stoffelijk zijn, en daarmee naar uiteindelijk *niet*-zijn. Door Rainbow te laten benoemen dat zij, doordat ze volledig uit digitale informatie bestaat, niet *niet* kan zijn, laat URLAND een publiek nadenken over nog een andere tegenstelling tussen mens en kunstmatige intelligentie: stoffelijk versus onstoffelijk, en sterfelijk versus onsterfelijk.

⁹ Cornelis, Coen. *De robot discipel*. p. 15 - 16



De twee robots uit Cornelis' tekst reflecteren op weer een ander aspect van menszijn: onze emoties. De ene robot is erg blij daar beschikking over te hebben gekregen, terwijl de ander emoties beschouwt als een belemmering van rationaliteit. Zonder verdere inmenging van menselijke personages verbeeldt Cornelis met deze discussie een van de belangrijkste vragen uit sciencefiction over kunstmatige intelligentie: is er een verschil tussen een programmering en het bewustzijn? Wat zou het menselijk bewustzijn kunnen onderscheiden van een kunstmatige programmering? De onbekende oorsprong? Het idee dat ons bewustzijn niet puur rationeel, maar ook emotioneel is? En als we nu op basis van bevindingen uit de neurowetenschappen veronderstellen dat onze emoties en ons bewustzijn te herleiden zijn tot onze hormoonhuishouding en hersenactiviteit, twee processen die qua werkzaamheid op computerprogramma's lijken?

Misschien slaan we flink aan het blunderen als wij, kunstenaars en toeschouwers, als ware amateurfilosofen dergelijke vragen proberen te beantwoorden.¹⁰ Dat hoeven we dan ook niet te doen. We kunnen ons in plaats daarvan richten op wat we wel heel goed kunnen, namelijk; die vragen stellen, ze van alle kanten bekijken, en de veelheid aan mogelijke antwoorden laten zien.

De voorstelling *De Eersten* van theatergezelschap De Hollanders geeft een voorbeeld van hoe we dat kunnen doen. De kunstmatig intelligente boordcomputer KOFAP monitort en reguleert het leven van drie mensen in een van de buitenwereld afgesloten biosfeer. Tijdens het stuk bekruipt het publiek het gevoel dat er reden is om KOFAP te wantrouwen:

Thomas: Luister je naar me?
KOFAP: Ik luister.
Thomas: Ja, maar begrijp je me ook?
KOFAP: Dat weet ik niet.
Thomas: Weet je KOFAP... ben jij eigenlijk aansprakelijk voor wat je doet? Ik bedoel, als jij een fout maakt... ben jij daar dan, jij als individu, ben jij daar dan zelf aansprakelijk voor? Jij, alleen jij... niet iemand anders...
KOFAP: Ik denk het wel... maar mijn denken is maar heel beperkt.
Thomas: Hoe bedoel je?
KOFAP: Ik heb bijvoorbeeld nooit een idee.
Thomas: Nee, nee... daar gaat het niet om... of je kunt *lijden*... *daar* gaat het om. Angst voor je eindigheid, *daar* gaat het om. ... Iemand... iets... kan alleen

¹⁰ Lagendijk, Ad. "Bewustzijn." *de Volkskrant*, 28 oktober 1995. <https://www.volkskrant.nl/nieuws-achtergrond/bewustzijn~b911c80f/> Laatst bezocht op 20 december 2018.



- aansprakelijk voor zijn daden worden gesteld als het kan lijden... dus...
ben je soms eenzaam?
- KOFAP: Ik geloof het wel, ja. Ik denk dat ik weet hoe eenzaam voelt.
- Thomas: Dus je kunt lijden?
- KOFAP: Zo ben ik geprogrammeerd, mijn programmeurs zijn daar aansprakelijk voor.
- Thomas: Kun jij liegen?
- KOFAP: Nee. Zo ben ik geprogrammeerd. Mijn programmeurs zijn daar aansprakelijk voor.
- Thomas: Hoe ben ik geprogrammeerd?¹¹

Ook in deze dialoog worden de hierboven genoemde grote vragen gesteld. Door KOFAPs meerduidige antwoorden wordt het publiek aangezet tot zelf conclusies trekken; conclusies die de interpretatie van de rest van het stuk bepalen. Aan het einde van het stuk is onduidelijk of KOFAP volledig neutraal gewoon zijn programmering volgt, en volgens vaste regels het leven van de bewoners reguleert, of dat hij slechts de indruk wekt dat hij dat doet, terwijl hij in werkelijkheid de bewoners manipuleert. De tekst sluit geen van beide mogelijkheden uit, waardoor uiteindelijk zelfs beargumenteerd kan worden dat KOFAP inderdaad de bewoners manipuleert, maar dat hij dat doet omdat dat onderdeel is van zijn programmering.

De boze robot voorbij

Alle besproken kunstmatig intelligente personages vertonen één belangrijke overeenkomst: zij zijn allemaal niet (alleen maar) kwaadaardig. In plaats daarvan worden de personages herkenbaar kunstmatig intelligent gemaakt door enkele talige of psychologische kenmerken van menszijn aan te passen, of door kunstmatig intelligente personages net andere kenmerken te geven dan menselijke personages. Op die manier bevraagt iedere tekst wat menszijn is, en in welke opzichten kunstmatige intelligentie van ons verschilt, op ons lijkt, of ons zelfs overtreft. Daarbij laten de teksten ruimte voor verschillende interpretaties. Deze teksten tonen aan dat het drama niet noodzakelijk gezocht hoeft te worden in een machtsstrijd tussen mens en kunstmatige intelligentie, omdat spanning ook voort kan komen uit bevreemding en ongrijpbaarheid.

Ook de kunstmatig intelligente personages in deze bundel bevreemden de menselijke personages en de lezer. Soms is onduidelijk welk personage menselijk, en welk personage kunstmatig intelligent is. Soms lijkt de verdeling duidelijk, maar blijkt het uiteindelijk toch

¹¹ De Hollanders. *De Eersten*. p. 19 - 20



andersom te zijn. Vaak zijn er slechts subtiele verschillen. Op dezelfde subtiele wijze worden grote vragen en antwoorden soms met één zin opgeroepen. De teksten tonen en bevragen onze angst om ingehaald of onderworpen te worden, maar wijzen ons ook op iets dat veel enger is dan boosaardige robots: de verwachtingen die we van kunstmatige intelligentie hebben, en de mogelijke teleurstelling die op die verwachtingen zou kunnen volgen. Beide kanten van de medaille worden onderzocht: wat doet het met ons als kunstmatige intelligentie ons voorbijstreeft? En wat als kunstmatige intelligentie ons nu eens helemaal niet overtreft? Maakt kunstmatige intelligentie ons eenzamer, of kunnen we er volwaardige relaties mee op bouwen? Als kunstmatige intelligentie ons de mogelijkheid biedt om onsterfelijk te worden, moeten we dat dan willen? En als het ons niet die mogelijkheid biedt, maar het zelf wel eeuwig bestaat, wat dan? Wat betekent het voor ons als ons bewustzijn kunstmatig te evenaren is? En wat als dat onmogelijk blijkt? Maakt kunstmatige intelligentie ons leven gemakkelijker, of ingewikkelder? Helpt kunstmatige intelligentie ons om van onszelf los te raken, of werpt het ons steeds weer op onszelf terug?

© Rafaëlle Kwakkel, 2019

ⁱ Er is veel geschreven over de relatie tussen gender en kunstmatige intelligentie, zie bijvoorbeeld Irit (Irene) Sternberg's artikel "Female AI: The Intersection Between Gender and Contemporary Artificial Intelligence".

ⁱⁱ 'Robot' is afgeleid van het Tsjechische *robota*, dat 'verplichte arbeid' betekent.

