

Verslag van de Hollands Spoor-bijeenkomst over Artificiële Intelligentie op dinsdag 2 oktober 2018

Aanleiding

Het kabinet heeft de WRR in juni 2018 gevraagd om discipline-overstijgend onderzoek te doen en te adviseren over de impact van AI op publieke waarden. AI biedt volgens het kabinet veel kansen, bijvoorbeeld op het terrein van het voorspellen van strafbare feiten en gezondheidsrisico's, het opsporen van ongewenste inhoud op sociale mediaplatforms en voorkomen van schooluitval. Maar er zijn ook uitdagingen, die te maken hebben met discriminatie, onwenselijke sturing van het gedrag van burgers en de mate waarin AI-systemen zijn te controleren en corrigeren. Een veelbesproken onderwerp is bovendien de impact van AI op de arbeidsmarkt en productiviteit.

De WRR heeft positief gereageerd op de adviesaanvraag. Het onderwerp sluit nauw aan rapporten als *Staat zonder land* (1998), *iOverheid* (2011), *De robot de baas* (2015) en *Big Data in een vrije en veilige samenleving* (2016).

Als aftrap voor het adviestraject is op 2 oktober een de Hollands Spoor bijeenkomst over AI georganiseerd, in samenwerking met het Strategiebeeraad Rijksbreed. De ruim zeventig deelnemers waren afkomstig van onder meer de departementen, het bedrijfsleven, universiteiten en onderzoeksinstituten, adviesraden en planbureaus. Gedurende het eerste deel van de bijeenkomst was ook de premier aanwezig. De bijeenkomst bestond uit een korte introductie tot de thematiek en de bijeenkomst door WRR-voorzitter Corien Prins en drie lezingen gevolgd door discussie. De lezingen waren georganiseerd rond de vragen 'wat is en kan AI', 'welke impact heeft AI', en 'wat doen overheden wereldwijd om AI te bevorderen'.

AI: Wat het is. Wat het kan

Inleiding door Maarten de Rijke (Universiteitshoogleraar Artificial Intelligence en Information Retrieval, UvA)

Er bestaat veel onduidelijkheid over wat AI is en wat het kan, nu en vooral ook in de nabije toekomst, als de ontwikkelingen verder zijn. Als gevolg hiervan bediscussiëren we niet altijd de voor de samenleving meest relevante vragen. Maarten de Rijke gaf een eerste aanzet voor een gezamenlijk perspectief op wat AI is.

De Rijke definieert AI als een systeem dat (tot op zekere hoogte) autonoom kan waarnemen, handelen en zich aanpassen. In toenemende mate beschikken machines dankzij AI over menselijke kwaliteiten als waarnemen, leren, plannen, realiseren, zoeken en handelen. Een belangrijke oorzaak voor deze ontwikkeling ligt in de beschikbaarheid van meer rekenkracht, meer data en betere algoritmes. De voor AI gebruikte systemen zijn daarbij data- of modelgestuurd. Datagestuurde systemen kunnen op drie manieren leren: 'supervised' (de supervisor nadoen), 'unsupervised' (data verkennen en een structuur zoeken) en 'reinforced' (met beloningen). De Rijke illustreert 'supervised' leren aan de hand van een systeem dat boten of gezichten kan leren herkennen.



Wat kun je nu met AI? AI kan ons volgens De Rijke helpen beslissen onder condities van onzekerheid, vooral wanneer grote hoeveelheden (historische) data beschikbaar zijn. Goede en succesvolle voorbeelden zijn de prognose van longkanker, het inschatten van het toekomstige succes van startups of schrijven van weerberichten en jaarrapportages. Belangrijk hierbij is dat mens en machine intensief leren samenwerken, want alleen zo komen in de praktijk betere beslissingen tot stand. De Rijke

spreekt van hybride intelligentie. Voor deze vorm van intelligentie zijn behalve kennisontwikkeling, innovatie en onderwijs ook nieuwe businessmodellen nodig, evenals aandacht voor de kwaliteit van data, ethiek en de rol van de overheid als mogelijke 'early adaptor' van AI-systemen.

De maatschappelijke impact van AI

Inleiding door Catelijne Muller (lid High-Level Expert group on AI voor de Europese Commissie en rapporteur AI & Society voor het Europees Economisch en Sociaal Comité)

De tweede lezing van de Hollands Spoorbijeenkomst richtte zich op de vraag wat de maatschappelijke impact is van AI. Hierover is veel discussie en bijna dagelijks verschijnen er in het nieuws berichten over de betekenis van AI, zowel in positieve als in negatieve zin. Muller identificeerde in haar lezing de belangrijkste impactgebieden.

De impact van AI loopt volgens Muller uiteen van bijvoorbeeld arbeid, onderwijs, gelijkheid, veiligheid, democratie, ethiek, superintelligentie tot wet- en regelgeving. Het komt bij AI vooral aan op de kwaliteit van de data en de algoritmes enerzijds en een verantwoord gebruik daarvan anderzijds. AI systemen hebben nu vaak ongewenste gevolgen, bijvoorbeeld omdat de kwaliteit niet hoog genoeg is of de data rommelig, onvolledig of 'biased'. Ter illustratie verwijst Muller naar een in de Verenigde Staten gehanteerd (en weer afgeschaft) systeem dat over recidive risico adviseerde en 'biases' bleek te bevatten ten opzichte van zwarte verdachten.



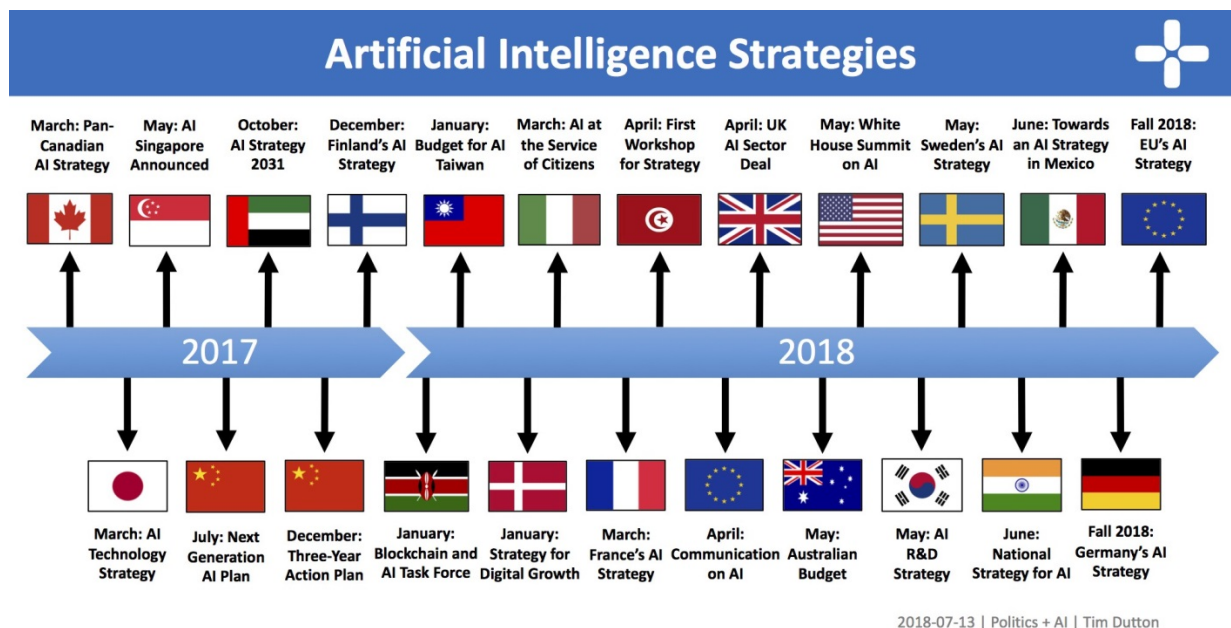
Ook speelt de vraag hoe ervoor te zorgen dat AI op een goede manier op de werkplek landt en wat mensen nodig hebben om hun werk beter te kunnen doen. Die vraag is belangrijker dan hoeveel banen er door AI verloren zullen gaan. Daarover zijn inmiddels zeer uiteenlopende voorspellingen gedaan, die vooral duidelijk maken dat we niet precies weten hoe het zit. Ten aanzien van wet- en regelgeving keert Muller zich tegen een vorm van rechtspersoonlijkheid van AI-systemen. AI en de vragen rondom aansprakelijkheid die daarbij spelen, zijn uitstekend in te passen binnen het huidige, sterke, rechtssysteem. Bovendien zou een dergelijke juridische constructie aldus Muller de preventieve werking van het aansprakelijkheidsrecht ondermijnen. Een van de belangrijkste opgaven is volgens Muller dat we met ons allen het gesprek voeren over de impact van AI op de maatschappij – er moet vanuit verschillende kanten naar worden gekeken. AI overkomt ons niet, zo besluit zij. Mensen blijven te allen tijde verantwoordelijk voor AI-systemen in de technische zin van het woord en voor beslissingen over hoe we AI willen inzetten.

AI-overheidsbeleid. Domme data, slimme computers en wijze mensen

Inleiding door Bennie Mols (wetenschapsjournalist AI & Robot en co-auteur Hallo Robot)

Nederland heeft tot op heden geen nationale AI-strategie, in tegenstelling tot een groeiend aantal andere landen. Welke prioriteiten stellen deze landen in hun strategieën? Welke impactgebieden keren daarin terug? Mols richtte in zijn lezing de aandacht op AI-overheidsbeleid en gaf daarbij een overzicht van enkele kenmerkende elementen in de door diverse landen inmiddels gepresenteerde AI-strategieën.

Mols opende met een overzicht van landen die afgelopen jaren een AI-strategie hebben geformuleerd. Centraal in deze strategieën staat het maximaliseren van de opbrengst van AI en het minimaliseren van de ermee gepaard gaande kosten en risico's. China wil in 2030 wereldleider zijn in AI-innovaties en overheid en industrie werken daartoe nauw samen. Ook de VS ambiëren leiderschap en geven onder Trum ruimte aan de vrije markt voor innovatie. De EU komt eind 2018 met een strategie. Er ligt een mooie kans voor Europa om, in samenwerking tussen EU-landen, een middenweg te vinden tussen China (overheid-industrie) en de VS (vrije markt). Er ligt al een voorstel voor een groot gezamenlijk onderzoeksinstituut (Claire). Duitsland is van plan zich met AI op de industrie te richten. Het Verenigd Koninkrijk en Frankrijk hebben onderzoekscentra van grote tech-bedrijven aan zich weten te binden.



Welk verhaal zou de overheid moeten vertellen over AI? Volgens Mols biedt AI vooral mogelijkheden om ons leven te verbeteren als we denken in termen van mens én machine in plaats van mens versus machine. Omdat AI-effecten ontstaan in relatie tot de bredere omgeving, moeten ook niet-technenuten bij de ontwikkeling ervan worden betrokken. Bovendien is democratische controle gewenst om te bepalen welke doelen we willen realiseren. AI helpt ons bij het optimaliseren, maar het is de mens die bepaalt wat we willen optimaliseren en wat als optimaal geldt.

Discussie onder leiding van WRR-secretaris Frans Brom

Wat de overheid kan doen

Uit de discussie komt naar voren dat het van groot belang is om de burger mee te krijgen bij de toepassing van AI. Hierbij moet het doel centraal staan en niet het middel: hoe breng je AI in de praktijk (AI als 'capability') op zo'n manier dat de techniek bruikbaar is? Hier ligt een taak voor de overheid. Verder gaat het erom dat de media

de burger een gebalanceerde kijk bieden op AI: wat kan de technologie wel en wat niet? Ook verhalen van de werkvloer zijn essentieel om te vertellen omdat ze de maatschappelijke impact van AI concreet maken.

De rol van Europa

De aanwezigen zien een belangrijke rol voor Europa weggelegd. De VS hechten groot belang aan de Europese markt: een grote markt met kapitaalkrachtige consumenten. Een markt waar de VS als belangrijke speler wil worden erkend. Dit biedt de overheid (EU) de kans om regels te stellen en af te dwingen dat bepaalde waarden worden geborgd. De kracht van Europa ligt op het gebied van juist die publieke waarden. Pak vraagstukken rond die waarden dan ook zoveel mogelijk internationaal op. Publieke waarden verschillen echter per land. Zorg daarom wel dat afspraken ook nationaal kunnen landen. Wel moet de EU zich realiseren dat ze de competitie verliest als de innovatie elders plaats vindt. Overigens is een aantal bedrijven al volop bezig met AI. Gezamenlijk optreden van die bedrijven is van belang om verdere stappen te kunnen zetten. Voor startups geldt dat zij behoefte hebben aan een goede infrastructuur om op het gebied van AI iets te kunnen gaan betekenen.

Mens en machine

AI is volgens de deelnemers een containerbegrip, dat raakt aan data, algoritmes, presentatie en interactie, mensen en maatschappij. AI die aspecten verdienen de aandacht. De deelnemers beschouwen AI bovendien vooral als een hulpmiddel om slimmer, beter en productiever te worden. AI heeft echter ook beperkingen: bij 'medical imaging' scoort het systeem weliswaar beter dan de gemiddelde mens, maar het kan niet uitleggen waarom het tot zijn beslissing is gekomen. Daarom moet het bij AI altijd gaan om de combinatie van mens en machine: hybride AI. Die combinatie is overigens niet overal wenselijk. In de rechtspraak is een robotrechter bijvoorbeeld ondenkbaar. Een machine kan niet beslissen over partijen, onder andere vanwege het vertrouwen dat mensen moeten hebben in de rechtspraak. Dit laatste hangt samen met de empathie die een menselijke rechter kan uiten. AI creëert bovendien een dubbele trainingsbehoefte. Enerzijds zullen mensen met AI moeten leren werken. Anderzijds dreigt het gevaar dat nieuwe generaties niet langer in staat zijn om AI te beoordelen, omdat ze nooit zonder hebben gewerkt. Dit aspect verdient aandacht.

Publieke waarden

AI brengt ook vragen over bijvoorbeeld privacy en ethiek met zich mee. AI is bijvoorbeeld goed in personalisatie. Heb je dan het recht om niet-gepersonaliseerde informatie te ontvangen? AI geeft ook ruimte voor gedragsbeïnvloeding op een heel ander niveau dan tot nu toe. Wat doet dit met de balans in de samenleving? Dergelijke zaken behoeven transparantie. In meer algemene zin geldt dat controle nodig is op AI-activiteiten, bijvoorbeeld om vast te stellen hoe goed een algoritme moet zijn om te kunnen vertrouwen en vast te stellen of een algoritme wel de manier is om selecties te maken. Dit speelt als het gaat om de toewijzing van kinderen aan scholen (zoals in Amsterdam). Ook de rechtspraak krijgt meer en meer te maken met het beoordelen van besluiten die zijn gebaseerd op AI-applicaties. De vraag die daarbij speelt, is hoe de beslissing tot stand is gekomen. Wat zijn de gevaren als te snel op (slechte) algoritmes wordt vertrouwd? Ook kunnen fouten worden gemaakt bij het invoeren en verwerken van data. In het internationaal strafrecht speelt dit soort lastige kwesties bijvoorbeeld bij oorlogsmisdaden via geautomatiseerde wapensystemen: hoe zit het met de verantwoordelijkheid en de besluitvorming bij dergelijke systemen? Zijn de gebruikte systemen voldoende betrouwbaar? Moeten we zeggen dat AI niet in oorlogsvoering mag worden gebruikt?

Een ecosysteem voor talent

Zorg (à la Canada) dat mensen aan boeiende vraagstukken kunnen werken, dan blijven zij behouden voor Nederland. Universitair onderzoek en opleiding spelen hierbij een belangrijke rol. Begin bijvoorbeeld met jongeren uit te leggen wat AI is (in navolging van Finland). De binnenkort te lanceren alliantie ALLAI Nederland gaat bijvoorbeeld (in navolging van Groot-Brittannië) weekend- en zomerscholen verzorgen waar jongeren de kans krijgen om in teams te werken aan de oplossing van een (sociaal) probleem. Of een onderzoeksinstituut à la CERN (CLAIRE) gaat werken, is de vraag. Het fundamentele onderzoek naar elementaire deeltjes binnen CERN, heeft geen economische betekenis. Dit in tegenstelling tot de spin-offs van het onderzoek. Op het gebied van AI zijn er daarentegen wel enorme economische verwachtingen.