



Artificiële Intelligentie | Human-in-Command

Position Paper ten behoeve van de Tweede Kamer Commissie
Economicsche Zaken en Klimaat

Mr. Catelijne Muller

Artificiële Intelligentie | Human-in-Command

ALLAI Position Paper t.b.v. de Tweede Kamer Commissie Economische Zaken en Klimaat
Mr. Catelijne Muller¹

Artificiële Intelligentie (AI) maakt al enige tijd een groot aantal belangrijke ontwikkelingen door en heeft inmiddels zeer veel toepassingen in de maatschappij.² AI-systemen hebben daarmee een steeds grotere en diepere impact op ons dagelijks leven. En hoewel AI grote maatschappelijke voordelen zou kunnen hebben³, brengt het ook risico's en complexe beleidsuitdagingen met zich mee op gebieden als veiligheid en controleerbaarheid, social-economische aspecten, ethiek, betrouwbaarheid, wet- en regelgeving, etc. Dat vraagt om kritische reflectie en politieke keuzes maar ook om samenwerking tussen makers, gebruikers en de maatschappij als geheel. ALLAI is ervan overtuigd dat de voordelen van AI alleen duurzaam kunnen worden bereikt als ook de uitdagingen rondom AI adequaat en tijdig worden geadresseerd.

1. Human-in-Command⁴

Voorzitter van ALLAI, Catelijne Muller (tevens schrijver van dit Paper), roept al geruime tijd op tot een 'human-in-command' benadering van AI waarbij machines machines blijven en mensen te allen tijde de controle over deze machines behouden. Niet alleen op technisch vlak, maar met name ook voor wat betreft de keuze of en zo ja, onder welke condities het AI-systeem in een bepaalde situatie wordt gebruikt. Dit vereist dat de algehele impact van het AI-systeem (economisch, maatschappelijk, juridisch en ethisch) in acht wordt genomen.⁵

2. Nederland, Europa en de wereld

Met name op EU niveau, maar ook internationaal en in een groot aantal EU-lidstaten AI allang op de agenda. In 2017 signaleerde het EESC (Europees Economisch en Sociaal Comité) in haar baanbrekend rapport *AI & Society*⁶ dat AI een technologie was om serieus rekening mee te houden. Het rapport benoemde impact domeinen van AI⁷ en riep de Europese Commissie en maatschappelijke spelers op om gezamenlijk de (rand)voorwaarden voor de verdere verantwoorde ontwikkeling en inzet van AI te bepalen en hierin mondiaal een voortrekkersrol te spelen. Deze oproep heeft op EU-niveau ruimschoots zijn gehoor gekregen in de vorm van:

¹ Voorzitter ALLAI, lid van de High Level Expert Group on AI van de Europese Commissie, Voorzitter Thematic Study Group on AI van het Economisch en Sociaal Comité (EESC), Rapporteur AI voor het EESC.

² "The AI Index 2018 Annual Report"

³ Duurzamere landbouw, veiliger verkeer, veiliger financieel systeem, betere productieprocessen, betere geneeskunde, veiliger werk of oplossingen voor klimaatverandering en armoedebestrijding.

⁴ [VPRO Tegenlicht, Mens in de Machine](#), met Catelijne Muller over *human-in-command*, 42:20'; EESC INT/806, Artificial Intelligence & Society, 2018, Rapporteur: Catelijne Muller (schrijver van dit Position Paper).

⁵ De human-in-command benadering van AI is overgenomen in de EU Ethics Guidelines for Trustworthy AI van de AI HLEG.

⁶ EESC 2017 INT/806 Artificial Intelligence & Society.

⁷ Het EESC heeft 11 domeinen aangeduid waarbinnen AI voor maatschappelijke uitdagingen zorgt: ethiek, veiligheid, transparantie, privacy, arbeid, educatie, (on)gelijkheid,, wet- en regelgeving, democratie, oorlogsvoering en superintelligentie.

- Een Europese AI strategie, AI for Europe⁸
- Het *Coordinated Plan on AI*⁹, getekend door de EU lidstaten
- De *High Level Expert Group on AI* van de Europese Commissie (AI HLEG)¹⁰
- Ethische Richtlijnen voor Betrouwbare AI van de AI HLEG¹¹
- Investerings- en Beleidsaanbevelingen voor Betrouwbare AI van de AI HLEG¹²
- Nieuwe Voorzitter van de Europese Commissie Ursula von der Leyen: "Tijdens mijn eerste 100 dagen in functie zal ik wetgeving voorstellen voor een gecoördineerde Europese aanpak van de menselijke en ethische implicaties van kunstmatige intelligentie."¹³
- Whitepaper on AI van de Europese Commissie (wordt gepubliceerd op 19 februari 2020).

Nederland heeft inmiddels een ook een AI-strategie. Het AI-beleid dat daarin wordt omschreven, is echter zeer gefragmenteerd en versnipperd over diverse ministeries, waardoor het risico bestaat dat het onvoldoende samenhangt, niet beklijft of zelfs vervliegt. Deze strategie biedt bovendien onvoldoende visie op de grote kansen maar ook uitdagingen die AI biedt. Naar de mening van ALLAI is het Nederlandse kabinet op beide punten te 'nonchalant'. De kansen worden onvoldoende ondersteund met financiële middelen en de uitdagingen worden benoemd, maar niet daadwerkelijk onderkend of geadresseerd. Exemplarisch voor dat laatste is de recente SyRI-uitspraak, waarin de Rechtbank heeft bepaald dat de SyRI-wet in strijd is met art. 8 EVRM.¹⁴

3. Wat is AI?

AI is een containerbegrip voor een groot aantal (sub)domeinen zoals *machine learning*, *deep learning*, *natural language processing*, etc. Het doel van AI-onderzoek en -ontwikkeling is de automatisering van intelligent gedrag.¹⁵ AI valt grofweg te onderscheiden in *narrow AI* (in staat om zeer specifieke taken te verrichten) en *general AI* (in staat iedere intellectuele taak te verrichten die een mens kan verrichten¹⁶). *General AI* bestaat niet. De huidige AI-systemen zijn voorbeelden van *narrow AI* met belangrijke karaktersitien:

- Doelgericht: AI-systemen definiëren niet hun eigen doelen, maar gebruiken technieken om de optimale oplossing voor een door mensen gesteld doel te bereiken.
- Big Data: AI-systemen hebben veel data nodig om goed te presteren en de kwaliteit, het volume en de inhoud van de data beïnvloedt de werking van het AI-systeem¹⁷

⁸ Artificial Intelligence for Europe, COM(2018) 237 final, gebaseerd op 3 pijlers: (i) technologische en industriële AI capaciteit van de EU; (ii) oog voor sociaal-economische veranderingen door AI; en (iii) een passend ethisch en wettelijk kader voor AI.

⁹ Coordinated Plan on Artificial Intelligence (COM(2018) 795 final), ondertekend door de lidstaten om tot een gezamenlijke Europese benadering van AI te komen in lijn met de de Europese strategie.

¹⁰ Onafhankelijke adviesgroep van 52 experts uit het maatschappelijk middenveld, de wetenschap en het bedrijfsleven; Schrijver van dit Paper, Catelijne Muller, is lid van de AI HLEG.

¹¹ Ethics Guidelines for Trustworthy AI, High Level Expert Group on AI, 2018.

¹² Policy and Investment Recommendations for Trustworthy AI, High Level Expert Group on AI, 2019.

¹³ Von der Leyen: "A Union that strives for more: My agenda for Europe".

¹⁴ ECLI:NL:RBDHA:2020:865

¹⁵ De AI HLEG heeft een Definitie voor AI opgesteld: A Definition of AI: Main Capabilities and Disciplines, 2018.

¹⁶ Wetenschappers zijn het er niet over eens of en zo ja, wanneer *general AI* zal worden bereikt..

¹⁷ Data is niet neutraal, het zijn geen feiten, kan rommelig zijn, onvolledig of te omvangrijk en het kan vooroordelen bevatten.

- Correlatie: AI-besluiten berusten veelal slechts op statistische verbanden opgemaakt uit veel data (correlatie). Het systeem legt (of ziet) geen causaal verband tussen een gebeurtenis of handeling en een beslissing, maar doet slechts een voorspelling.
- Black-boxes: veel AI-systemen zijn zogenaamde *black boxes*, waarbinnen (beslis)processen plaatsvinden die niet kunnen worden doorgrond.
- Gezond verstand: AI-systemen hebben geen gezond verstand.
- Algoritmes zijn slechts een onderdeel van AI: een AI-systeem bestaat uit (vaak meerdere) complexe algoritmes, data en modellen.

4. AI impact domeinen

Hieronder worden de belangrijkste impact domeinen van AI beschreven:

- Veiligheid: om veilig te kunnen opereren moeten AI-systemen robuust zijn, bestand zijn tegen hacken en veilig opereren onder onbekende, kritieke of onvoorspelbare situaties.
- Begrijpelijkheid, controleerbaarheid en uitlegbaarheid: inmiddels wordt er alom gepleit voor transparante, begrijpelijke, uitlegbare en controleerbare AI-systemen waarvan de werking ook achteraf verklaarbaar is.¹⁸ Dit is cruciaal en vaak verplicht¹⁹, maar niet voldoende. Belangrijk is om ook vast te stellen welke beslisprocessen of -momenten aan AI-systemen kunnen worden overgelaten en welke niet en wanneer menselijke interventie gewenst of vereist zou moeten zijn (human-in-command).²⁰
- Ethiek: AI roept veel ethische vragen op. Wat is de invloed van autonome (zelflerende) AI op onze persoonlijke integriteit, autonomie, waardigheid, zelfstandigheid, gelijkheid, veiligheid, keuzevrijheid? Hoe zorgen we ervoor dat onze fundamentele normen, waarden en mensenrechten geëerbiedigd en gewaarborgd blijven? De AI HLEG heeft inmiddels Ethische Richtlijnen voor Verantwoorde AI gepresenteerd, waarbij de grondrechten als basis voor AI-ethiek zijn genomen.²¹ De richtlijnen beschrijven 7 vereisten voor betrouwbare (ethische, wettige en robuuste) AI²²:
 1. Menselijke controle en menselijk toezicht
 2. Technische robuustheid en veiligheid
 3. Privacy en data-governance
 4. Transparantie
 5. Diversiteit, non-discriminatie en rechtvaardigheid
 6. Maatschappelijk en milieuwelzijn
 7. Verantwoording
- Werk: AI kan grote voordelen hebben wanneer het wordt ingezet bij gevaarlijk, zwaar, uitputtend, vuil, onaangenaam, repetitief of saai werk, maar ook dataverwerking en -analyse, planning en predictie, kan in toenemende mate door AI-systemen worden

¹⁸ EESC opinie AI & Society (INT/806), 2017; Ethics Guidelines for Trustworthy AI van de HLEG AI, 2018; Artificial Intelligence for Europe, COM(2018) 237 final; AI for Humanity, France, 2018, a.o.

¹⁹ Dan de situatie waar een overheid een geautomatiseerd beslissysteem inzet om te bepalen of iemand gaat frauderen (SyRi). De overheid zal onder het motiveringsbeginsel en zorgvuldigheidsbeginsel haar besluiten moeten kunnen uitleggen. Met een *black box* AI systeem kan dat niet (zie ook Raad van State, Kamerstukken II 2017/18, 26643, nr. 557).

²⁰ Denk aan het gebruik van automatische beslissystemen of risico-analyses in de rechtspraak. Rechtspreken bestaat uit zoveel meer dan alleen het ontdekken van patronen in historische data. Typisch menselijke eigenschappen als compassie, intuïtie, voorstellingsvermogen en redelijkheid zijn evenzo belangrijk. Deze eigenschappen kunnen niet door AI worden overgenomen. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

²¹ EU Ethics Guidelines for Trustworthy AI, AI HLEG 2018.

²² De Europese Commissie heeft deze vereisten onderschreven in haar Communicatie, Building Trust in Human-Centric Artificial Intelligence (COM(2019) 168 final).

verricht.²³ AI-systemen worden steeds meer ingezet voor het volgen en monitoren van werkers, waardoor de autonomie en privacy in het geding komen. Werk wordt meer en meer bepaald en verdeeld door algoritmes, zonder menselijke tussenkomst, wat de aard van het werk en de werkomstandigheden beïnvloedt. En er bestaat het risico op verschraving van werk en het verleren van vaardigheden.²⁴

- Educatie en vaardigheden: het vanaf jonge leeftijd bijhouden of aanleren van digitale (en AI) vaardigheden en AI-begrip is nodig, maar ontwikkeling van (typisch menselijke) vaardigheden die niet door AI-systemen zullen worden bedreigd is net zo belangrijk.
- Bedrijfsmacht: het overgrote deel van de ontwikkeling van AI en alle daarbij behorende elementen (ontwikkelplatforms, data, kennis, expertise) is in handen van de 5 grote Amerikaanse en 2 Chinese technologiebedrijven (Amazon, Apple, Facebook, Google, Microsoft, Alibaba, Tencent).
- Maatschappelijke (on)gelijkheid: technologische veranderingen kunnen leiden tot (inkomens)ongelijkheid tussen mensen. AI kan deze trends verder versterken.²⁵
- Democratie: Het gebruik van AI-systemen voor nieuwsaggregatie kan leiden tot vernauwing van informatie ('*filter bubbles*' en '*nepnieuws*'). AI-systemen zijn gericht ingezet om stemgedrag van mensen te beïnvloeden.²⁶ Dit zijn bedreigingen voor een eerlijke en open democratie en kunnen voor ontwrichting van de maatschappij zorgen.²⁷
- Oorlogsvoering en dual-use: Een onbekend aantal landen en sectoren houdt zich bezig met onderzoek naar en ontwikkeling van dodelijke autonome wapensystemen, met het risico op een oncontroleerbare wapenwedloop, militaire situaties waarbij er vrijwel geen sprake meer is van menselijke controle, 'oneindige oorlogen' en catastrofaal systeemfalen.²⁸ AI is bovendien een *dual-use* technologie die makkelijk is in te zetten voor criminele, maatschappij-ontwrichtende of terroristische activiteiten.

5. Verdieping: AI & wet- en regelgeving

Bij technologische vooruitgang komt altijd de vraag op of en zo ja, hoe we de nieuwe disruptieve technologie moeten reguleren. Zo ook bij AI. AI is echter geen volledig ongeregelde technologie. We beschikken over vele wettelijke kaders die net zo goed op AI van toepassing zijn als op iedere andere technologie. Daarmee is niet gezegd is dat bestaande wet- en regelgeving in alle gevallen *fit for purpose* is voor AI en voldoende bescherming biedt. Zowel het EESC, HLEG AI en ALLAI pleiten daarom voor een AI wet- en regelgeving *stress-test* en nieuwe regelgeving voor intrusieve of ontwrichtende AI.

²³ ['A white-collar sweatshop': Google Assistant contractors allege wage theft](#); Gray en Suri: Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass.

²⁴ Overheid en sociale partners moeten gaan bepalen hoe en wanneer AI verder wordt ontwikkeld en ingezet in het arbeidsproces. Focus daarbij juist ook op waar mensen toe in staat zijn (creativiteit, empathie, samenwerken), op wat we willen dat mensen blijven doen op mogelijkheden om mens en machine beter samen te laten werken (complementariteit); voor complementariteit zie: De Robot de Baas, WRR, Robert Went *et al.*

²⁵ Cathy O'Neill, Weapons of Match destruction

²⁶ De recente Netflix documentaire "[The Great Hack](#)" geeft inzicht in de wijze waarop Cambridge Analytica met behulp van grote hoeveelheden data en AI in staat was kiezers te profileren en deze te proberen met fake news te beïnvloeden (nudgen).

²⁷ [VPRO Tegenlicht, Mens in de Machine](#), 30:36'

²⁸ Het Europees Parlement heeft opgeroepen tot de urgente ontwikkeling van een gemeenschappelijk, juridisch bindend standpunt ten aanzien van dergelijke systemen IResolutie 2018/2752(RSP) De HLEG AI heeft verklaard achter deze resolutie te staan en alle verwante inspanningen daaromtrent te ondersteunen. Ook ALLAI steunt deze resolutie. Resolutie 2018/2752(RSP)

5.1 AI opereert niet in een wetteloze wereld

Hieronder een aantal belangrijke bestaande wettelijke kaders relevant voor AI.

- **Grondrechten.** De volgende 'families van grondrechten' zijn in het bijzonder relevant voor de ontwikkeling en het gebruik van AI²⁹:
 - (i) Respect voor de menselijke waardigheid³⁰: in het kader van AI betekent dit dat alle mensen moeten worden behandeld als morele *subjecten*, en niet als ontleedbare, becijferbare, sorteerbare, conditioneerbare, manipuleerbare *objecten*.
 - (ii) Vrijheid van het individu³¹: in het kader van AI betekent dit dat mensen vrij zijn van (in)directe illegale dwang, bedreigingen voor de mentale autonomie en geestelijke gezondheid, ongerechtvaardigd toezicht, misleiding en oneerlijke manipulatie.³²
 - (iii) Respect voor democratie, justitie en de rechtsstaat³³: democratische processen en de beginselen van de rechtsstaat mogen niet worden ondermijnd door AI.
 - (iv) Gelijkheid, non-discriminatie en solidariteit: AI-systeem mogen geen *biased* uitkomsten hebben. Dit betekent dat zowel de data die wordt gebruikt om AI te trainen, als de samenstelling van de ontwikkelteams zo inclusief en divers mogelijk moeten zijn.
 - (v) Rechten van burgers: burgers hebben daarnaast vele overige rechten, zoals het kiesrecht, het recht van vereniging en het recht op openbaarheid van bestuur, inzage in overheidsdocumenten en bezwaar.
- **Aansprakelijkheidsrecht.** ALLAI is fel gekant tegen enige vorm van rechtspersoonlijkheid voor AI-systemen (net als het EESC en de AI HLEG). Van het aansprakelijkheidsrecht gaat een preventieve werking uit, die verdwijnt zodra de ontwikkelaar niet langer het aansprakelijkheidsrisico draagt.³⁴ Bovendien zijn aansprakelijkheidsregels omtrent hoofdelijke en mede-aansprakelijkheid, risico-aansprakelijkheid, productaansprakelijkheid, draagplicht, eigen schuld en regres ook van toepassing op AI³⁵ en kan aansluiting gezocht worden bij het leerstuk van gevaarstelling.
- **GDPR/AVG.** De GDPR (*General Data Protection Regulation*, in Nederland omgezet in de AVG (Algemene Verordening Gegevensbescherming) biedt enkele kaders voor het recht op uitleg van AI-besluiten (in de overwegingen)³⁶ en het recht op menselijke tussenkomst. Dit is echter onvoldoende om adequate transparantie in automatische besluitvorming te waarborgen. De voorwaarde van een volledig geautomatiseerd besluit

²⁹ Ethics Guidelines for Trustworthy AI of the High Level Expert Group on AI to the European Commission, 2018

³⁰ C. McCrudden, Human Dignity and Judicial Interpretation of Human Rights, *EJIL*, 19(4), 2008. Zie voor een vergelijkbare interpretatie van "menselijke waardigheid" E. Hilgendorf, Problem Areas in the Dignity Debate and the Ensemble Theory of Human Dignity, in: D. Grimm, A. Kemmerer,

C. Möllers (red.), *Human Dignity in Context. Explorations of a Contested Concept*, 2018, blz. 325 ff.

³¹ Waaronder het recht op fysieke en geestelijke vrijheid, de vrijheid van meningsuiting, het recht op een privéleven en privacy, de vrijheid van vergadering en vereniging, de vrijheid van ondernemerschap, kunsten en wetenschap.

³² Hieronder valt ook het recht op een persoonlijke levenssfeer. Bij privacyvraagstukken rondom AI gaat niet alleen om data die wordt gegenereerd en verwerkt door AI-systemen, maar met name ook om de impact op de persoonlijke levenssfeer die het AI-systeem kan hebben doordat het ons kan bespieden, af luisteren, beoordelen, profileren, manipuleren etc.

³³ Waaronder o.m. het recht op een eerlijk proces (*fair trial*).

³⁴ De vergelijking met de beperkte aansprakelijkheid van vennootschappen gaat hierbij mank, omdat daarbij altijd een natuurlijk persoon de ultieme aansprakelijke is.

³⁵ De Europese Commissie onderzoekt thans aansprakelijkheidsvraagstukken rondom AI m.m.v. de Expert Group on liability.

³⁶ Lokke Moerel en Marijn Storm, Law and Autonomous Systems Series, "Automated Decisions Based on Profiling - Information, Explanation or Justification, That is the Question!", Oxford University, Faculty of Law.

is eenvoudig te omzeilen en ook een besluit dat op niet-persoonlijke gegevens is gebaseerd kan grote impact hebben.

- **Consumentenrecht.**³⁷ Bestaande consumentenwetgeving is van toepassing, maar beschermt niet voldoende tegen oneerlijke, misleidende of manipulatieve praktijken (gepersonaliseerde prijsvorming, profilering, *nudging*).
- **Overige bestaande wet- en regelgeving en zelfregulering.** Er bestaan er talloze sector-, domein- en gebruikersspecifieke wetten en regels die relevant zijn voor AI. Denk aan het arbeidsrecht, bestuursrecht, intellectueel eigendomsrecht, vervoersrecht, gezondheidsrecht, consumentenrecht en regels voor financiële producten en diensten. Ten slotte is de praktijk van zelfregulering en standaardisering relevant voor AI.³⁸

5.2 AI wet- en regelgeving *stress-test*³⁹

Of bestaande wettelijke kaders voldoende waarborgen bieden moet door middel van een AI wet- en regelgeving *stress-test* worden onderzocht. Hierbij moet een 'risicobenadering' worden gekozen, uitgaande van het type en de mate van impact van de AI-toepassing. Voor specifieke AI-toepassingen die "onaanvaardbare" risico's genereren of substantiële bedreigingen vormen moet een op het voorzorgsbeginsel gebaseerde aanpak worden gevolgd.⁴⁰ Bij het ontwikkelen en aanpassen van wet- en regelgeving voor AI systemen is het van belang om het (gewenste) niveau van autonomie van het AI-systeem in acht te nemen. Daarbij moet ook worden gekeken en of de kaders voor monitoring, informatievergaring en handhaving van de wetgeving in staat om effectief toezicht te bieden.

5.3 Nieuwe wet- en regelgeving voor (gebruik van) bepaalde AI-toepassingen

Een aantal AI-toepassingen is dermate intrusief of (potentieel) ontwrichtend dat nieuwe wetgeving op zijn plaats is. Onderstaande (niet limitatieve) aanbevelingen zijn gelijk aan de aanbevelingen van de AI HLEG⁴¹. In de toekomst zullen er echter mogelijk nieuwe AI-toepassingen worden ontwikkeld waarvoor wetgeving moet worden overwogen:

- **Massa-surveillance:** onevenredige en massale surveillance van personen met gebruik van AI (bijvoorbeeld gezichtsherkenning), zowel door publieke als private partijen, moet worden voorkomen. Dergelijke systemen niet alleen een grote inbreuk op de privacy, maar hebben ze ook een 'chilling effect' op burgers en kunnen fundamentele rechten ondermijnen.⁴²
- **Biometrische herkenning:** personen mogen niet worden onderworpen aan ongerechtvaardigde persoonlijke, fysieke of mentale tracking of identificatie, profilering

³⁷ Commerciële AI-toepassingen zijn wijdverbreid: IoT producten met AI (stemherkenning (Google Home), objectherkenning (Roomba, Ring), klimaatbeheersing (Nest), *recommender*-systemen (Google Search, Booking.com, newsfeeds, etc.)).

³⁸ ISO (m.m.v. CEN-CENELEC en NEN) ontwikkelt op dit moment standaarden voor AI.

³⁹ De Nederlandse wetgever dient aansluiting bij (lopende en aanstaande) *stress-test* activiteiten van de Europese Commissie en te zoeken, om zo dubbel werk en divergerende resultaten te voorkomen.

⁴⁰ Regelgevende instanties moeten voorzorgsmaatregelen nemen bij een gevaar voor het milieu, de volksgezondheid of andere ernstige maatschappelijke bedreiging (zoals het democratisch proces)

⁴¹ Policy and Investment Recommendations for Trustworthy AI, AI HLEG, 2019.

⁴² Gezichtsherkenning wordt in een aantal Amerikaanse steden al verboden of beperkt. De positieve applicaties, zoals het vinden van een vermist kind of een terrorist, zijn potentieel aanwezig, ongereguleerd en indiscriminatoir verzamelen van gezichten in een database kan leiden tot een 'chilling effect', waarbij burgers zich constant bespied voelen en zich daar naar gaan gedragen. Gezichtsherkenning kan het recht op vereniging, staking of demonstratie ondermijnen, en gezichtsherkenning is niet perfect. Het herkent blanke mannen beter dan blanke vrouwen en een lichte huidskleur beter dan een donkere.

en nudging via AI-methoden voor biometrische herkenning zoals: emotionele tracking, empathische media, DNA, iris- en stemidentificatie, affectherkenning, gezichtsherkenning en de herkenning van micro-expressies.

- Social scoring: massale sociale scoring (zoals dat in delen van China wordt getest) moet worden voorkomen, maar ook beperkte scoring mechanismes (bijv. bij leningen) kunnen onze waarden en autonomie in gevaar brengen. Een volledig transparante procedure, voorafgaande toestemming (*opt-in*) of *opt-out* zijn onvoldoende waarborgen.
- Verplichte zelfidentificatie van AI-systemen: een AI-systeem mag zich niet heimelijk voordoen als mens.⁴³
- Autonome wapens: het Europees Parlement heeft opgeroepen tot de urgente ontwikkeling van een juridisch bindend standpunt ten aanzien van autonome wapens⁴⁴. De AI HLEG beveelt aan dat de ontwikkeling van geautomatiseerde dodelijke wapens moet worden bewaakt en beperkt.⁴⁵
- Schone data-lei voor kinderen: de invoering van een wettelijke leeftijd waarop kinderen een 'schone datalei', geschoond van openbare of particuliere data die aan hen als kinderen zijn gerelateerd, krijgen, moet worden overwogen.

6. Tot slot

Tot slot will ALLAI nog een aantal zaken meegeven voor beleidsbepaling rondom AI:

- Het wiel hoeft niet opnieuw uitgevonden: er is al bijzonder veel werk verzet en informatie beschikbaar over hoe de kansen van AI ten volle te benutten en de uitdagingen adequaat en tijdig het hoofd te bieden. Nederland hoeft het wiel dus niet opnieuw uit te vinden. Zij kan voortbouwen op het vele werk dat inmiddels met name op EU-niveau is verzet en gebruikmaken van de expertise die bijvoorbeeld bij ALLAI aanwezig is. Het is met name niet nodig om (publiek) geld en tijd te steken in het opstellen van (weer) een set ethische richtlijnen.
- Geen AI-race: de media, delen van het bedrijfsleven en beleidsmakers waarschuwen vaak voor een zogenaamde AI-race en het feit dat Nederland en Europa die race aan het verliezen zouden zijn. Naar de mening van ALLAI is dit discours zowel verkeerd als gevaarlijk.⁴⁶ Vaak wordt vergeten dat Europa de grootste markt.. Dit geeft de EU de mogelijkheid om de randvoorwaarden te stellen aan AI en daarmee ook wereldwijd invloed te hebben.⁴⁷
- Betrek de brede maatschappij: de impact van AI zo breed is dat er gekozen zou moeten worden voor een "full-helix-benadering", waarbij de maatschappij in de breedste zin is betrokken zoals bedrijven, onderzoekers uit diverse disciplines (waaronder AI-wetenschap, data-science, ethiek, filosofie, recht, economie en met name ook sociale wetenschappen (zoals culturele antropologie, pedagogiek, politicologie en psychologie, (de)centrale overheden en maatschappelijke organisaties, inclusief de sociale partners.⁴⁸

⁴³ "Google Duplex Uses AI to Mimic Humans (sometimes)", New York Times, 22 mei 2019.

⁴⁴ Resolutie 2018/2752(RSP) van het Europees Parlement.

⁴⁵ Policy and Investment Recommendations for Trustworthy AI, HLEG 2019.

⁴⁶ Virginia Dignum, "[There is no AI – race and if there is, it's the wrong one to run](#)".

⁴⁷ Een voorbeeld is de GDPR, waardoor bedrijven van buiten de EU zich aan strenge data privacy eisen moeten houden, maar die vindt inmiddels ook gehoor vindt bij beleidsmakers in andere delen van de wereld.

⁴⁸ <http://www.allai.nl/about-us>.

- Investeer slim en met visie: AI kan niet langer worden gezien als een technische of informatica-discipline, het is per definitie multidisciplinair, aldus Prof. Virginia Dignum. Onderzoek naar sociaal-maatschappelijke uitdagingen van AI, de ontwikkeling van veilige en robuuste AI-systemen alsook diversiteit binnen de wetenschappen om ervoor te zorgen dat AI niet-discriminerend en inclusief is zijn cruciaal. Investerings in AI moeten slim zijn, gericht op sectoren waar Nederland nu al sterk in is, waaronder 'Water en Maritiem', 'Agri en Food', 'Life Sciences & Health' en 'High-tech Systemen en Materialen'. In deze sectoren kan AI voor economische, sociaal-maatschappelijke en milieu-voordelen zorgen.⁴⁹

⁴⁹ AI kan een uitgelezen mogelijkheid bieden om oplossingen te vinden voor urgente uitdagingen waarmee de samenleving te maken heeft, zoals een vergrijzende bevolking, toenemende sociale ongelijkheid en milieuvuiling. Deze potentie is ook wereldwijd zichtbaar, bijvoorbeeld in de duurzameontwikkelingsdoelstellingen van de VN⁴⁹.

ALLAI.

ALLAI verwijst naar Stichting ALLAI Nederland, een stichting naar Nederlands recht. Geen enkele entiteit of persoon in het ALLAI-netwerk is verantwoordelijk of aansprakelijk voor enig direct of indirect verlies of enige schade geleden door een entiteit of persoon die op deze communicatie vertrouwt.

© 2019 ALLAI, Nederland

Herengracht 247
1016 BH Amsterdam
welkom@allai.nl
www.allai.nl