



Europees Economisch en Sociaal Comité

**INT/806
Kunstmatige intelligentie**

ADVIES

Europees Economisch en Sociaal Comité

**Kunstmatige intelligentie - De gevolgen van kunstmatige intelligentie voor de (digitale)
eengemaakte markt, de productie, consumptie, werkgelegenheid en samenleving
(initiatiefadvies)**

Rapporteur: **Catelijne MULLER**

Besluit van de voltallige vergadering	22/09/2016
Rechtsgrondslag	Artikel 29, lid 2, van het reglement van orde Initiatiefadvies
Bevoegde afdeling	Interne Markt, Productie en Consumptie
Goedkeuring door de afdeling	04/05/2017
Goedkeuring door de voltallige vergadering	31/05/2017
Zitting nr.	526
Stemuitslag (voor/tegen/onthoudingen)	159/3/14

1. Conclusies en aanbevelingen

- 1.1 Artificiële Intelligentie (AI) maakt op dit moment een aantal belangrijke ontwikkelingen door en vindt in snel tempo toepassingen in de maatschappij. De omvang van de markt voor AI bedraagt circa USD 664 miljoen en zal naar verwachting toenemen tot USD 38,8 miljard in 2025. Omdat AI zowel positieve als negatieve impact op de maatschappij kan hebben, heeft het EESC zich voorgenomen om de ontwikkelingen rondom AI nadrukkelijk te volgen, niet alleen vanuit technisch maar uitdrukkelijk ook vanuit ethisch, veiligheids- en maatschappelijk oogpunt.
- 1.2 Het EESC zal, als vertegenwoordiger van het Europees maatschappelijk middenveld het **maatschappelijk debat rondom AI** in de komende periode vormgeven, centraliseren en aanjagen en daarbij alle relevante stakeholders betrekken: beleidsmakers, de industrie, sociale partners, consumenten, NGO's, onderwijs- en zorginstellingen, experts en academici vanuit diverse disciplines (waaronder AI, veiligheid, ethiek, economie, arbeidswetenschappen, rechtswetenschappen, gedragswetenschappen, psychologie, filosofie).
- 1.3 Hoewel belangrijk, domineert de discussie over super-intelligentie op dit moment en overschaduwde deze het debat over de impact van de huidige toepassingen van AI. Taak en doelstelling van dit proces zal daarom onder meer zijn om kennis over AI te vergroten en verbreden en zo bij te dragen aan een geïnformeerd en gebalanceerd debat zonder doemscenario's of extreme relativisering. Het EESC zal zich daarbij inzetten voor de bevordering van de ontwikkeling van AI ten goede van de mensheid. Niettemin zal een belangrijke taak en doelstelling van dit proces ook zijn om disruptieve ontwikkeling in en rondom de ontwikkeling van AI te herkennen, benoemen, en monitoren, teneinde deze adequaat en tijdig te kunnen adresseren. Dit zal leiden tot grotere maatschappelijk(e) betrokkenheid vertrouwen, en draagvlak ten aanzien van de verdere duurzame ontwikkeling en inzet van AI.
- 1.4 De impact van AI is grensoverschrijdend, dus zullen ook supra-nationaal beleidskaders moeten worden vastgesteld. Het EESC beveelt aan dat de **EU mondiaal een voortrekkersrol** neemt in de vaststelling van wereldwijde, eenduidige beleidskaders voor AI, in lijn met Europese waarden en fundamentele rechten. Het EESC kan en wil daaraan haar bijdrage leveren.
- 1.5 Het EESC signaleert op dit moment **11 domeinen** waarbinnen AI voor maatschappelijke uitdagingen zorgt: ethiek; veiligheid; privacy; transparantie en verklaarbaarheid; arbeid; educatie en vaardigheden; (on)gelijkheid en inclusiviteit; wet- en regelgeving; bestuur en democratie; oorlogsvoering; superintelligentie. Het EESC geeft de volgende aanbevelingen.
- 1.6 Het EESC pleit voor een **human-in-command** benadering van AI met als randvoorwaarden de verantwoordelijke, veilige en nuttige ontwikkeling van AI, waarbij machines machines blijven en mensen te allen tijde de controle over deze machines zullen behouden.
- 1.7 Het EESC pleit voor een **ethische code** voor de ontwikkeling, inzet en het gebruik van AI, zodat AI-systemen gedurende hun gehele operationele proces compatibel zijn met de beginselen van

menselijke waardigheid, integriteit, vrijheid, privacy en culturele- en genderdiversiteit en met de fundamentele mensenrechten.

- 1.8 Het EESC pleit voor de ontwikkeling van een **normeringssysteem** ter verificatie, validatie en controle van AI-systemen, op basis van een breed spectrum aan standaarden op het gebied van veiligheid, transparantie, begrijpelijkheid, verklaarbaarheid en ethische waarden.
- 1.9 Het EESC pleit voor een **Europese AI-infrastructuur**, bestaande uit opensource en privacy respecterende leeromgevingen, *real life* testomgevingen en datasets van hoge kwaliteit voor ontwikkeling en training van AI-systemen. Het EESC wijst op het (concurrentie)voordeel dat de EU kan behalen op de mondiale markt, door de ontwikkeling en promotie van 'Europees verantwoorde AI-systemen', voorzien van Europese AI-certificering en labelling.
- 1.10 De EU, nationale overheden, en sociale partners moeten gezamenlijk in kaart brengen welke **arbeidsmarktsectoren** in welke mate en op welke termijn beïnvloed zullen worden door AI en oplossingen zoeken om de gevolgen voor de werkgelegenheid, de inhoud van het werk, sociale stelsels en (on)gelijkheid te adresseren. Daarnaast dient geïnvesteerd te worden in arbeidsmarktsectoren waarop AI weinig tot geen invloed heeft.
- 1.11 Het EESC beveelt deze stakeholders aan om gezamenlijk in te zetten op **complementaire AI-systemen en co-creatie** daarvan op de werkplek, zoals mens-machine teams, waarbij AI de prestaties van de mens aanvult en verbetert. Ook moeten de stakeholders investeren in **formeel en informeel leren, onderwijs, en opleidingen** voor iedereen om te kunnen werken met AI, maar ook om vaardigheden te ontwikkelen die niet door AI zullen of kunnen worden overgenomen.
- 1.12 Een **concrete beoordeling van de EU wet- en regelgeving** binnen de 6 door STOA (Scientific Foresight Unit) benoemde aandachtsgebieden die mogelijk herzien of aangepast moeten worden, is nu al noodzakelijk. Het EESC kan en wil bij dit beoordelingsproces een rol spelen. Het EESC is tegen invoering van een vorm van rechtspersoonlijkheid voor robots of AI. De preventief corrigerende werking van het aansprakelijkheidsrecht wordt daarmee uitgehold, er ontstaat een moreel risico bij zowel de ontwikkeling als het gebruik van AI en het biedt mogelijkheden tot misbruik.
- 1.13 Ontwikkeling van **AI-toepassingen die de maatschappij dienen**, inclusiviteit bevorderen en het leven van mensen verbeteren, moet actief bevorderd en gepromoot worden, zowel publiek als privaat. De Europese Commissie dient binnen haar programma's onderzoek te financieren naar de maatschappelijke impact van AI alsmede van door de EU gefinancierde AI-innovaties.
- 1.14 Het EESC ondersteunt de oproep van Human Rights Watch c.s. tot een **verbod op autonome wapensystemen**. Het EESC is ingenomen met het aangekondigde overleg hierover in de VN, maar meent dat ook de toepassingen van AI in cyberoorlogsvoering onderwerp van dit VN-overleg dienen te zijn.

2. **Artificiële Intelligentie**

- 2.1 Er is geen eenduidig geaccepteerde, vastomlijnde definitie van AI. AI is een containerbegrip voor een groot aantal (sub)domeinen zoals: cognitive computing (algoritmes die redeneren en begrijpen op een hoger (menselijker) niveau), machine learning (algoritmes die zichzelf taken aanleren), augmented intelligence (samenwerking tussen mens en machine), AI-robotica (AI ingebed in robots). Het centrale doel van AI-onderzoek en -ontwikkeling is echter de automatisering van intelligent gedrag, zoals redeneren, informatie vergaren, plannen, leren, communiceren, manipuleren, signaleren en zelfs creëren, dromen en percipiëren.
- 2.2 AI valt grofweg te onderscheiden in *narrow* AI en *general* AI. *Narrow* AI is in staat om specifieke taken te verrichten. *General* AI is in staat iedere intellectuele taak te verrichten die een mens kan verrichten.
- 2.3 Op het gebied van *narrow* AI is de afgelopen tijd grote vooruitgang geboekt, met name door de groei van *computer processing power*, de beschikbaarheid van grote hoeveelheden data en de ontwikkeling van *machine learning* (ML). ML behelst algoritmes die zichzelf specifieke taken kunnen aanleren, zonder daarvoor geprogrammeerd te zijn. Deze methode berust op het verwerken van 'trainingsdata' op basis waarvan het algoritme leert patronen te herkennen en regels op te stellen. *Deep learning* (DL), een vorm van ML, gebruikt structuren (*neural networks*) die losjes gebaseerd zijn op het menselijk brein, en die leren door middel van oefening en feedback. Het resultaat van deze ontwikkelingen is dat AI-systemen (door middel van algoritmes) inmiddels zelflerend, autonoom en adaptief kunnen zijn.
- 2.4 De belangrijkste focus van onderzoek en ontwikkeling in AI ligt al enige tijd op de elementen redeneren, kennisvergaring, planning, communicatie en perceptie (visueel, auditief en sensorisch). Dit heeft geleid tot een groot aantal AI-toepassingen: virtuele assistenten, zelfrijdende auto's, automatische nieuwsaggregatie, spraakherkenning, vertaalprogramma's, *text-to-speak* programma's, geautomatiseerde financiële handel, juridische *e-discovery*, etc.
- 2.5 Het EESC signaleert dat het aantal toepassingen en investeringen in AI recent exponentieel is gestegen. De omvang van de markt voor AI bedraagt op dit moment circa USD 664 miljoen en zal naar verwachting toenemen tot USD 38,8 miljard in 2025.

3. **Kansen en bedreigingen van AI**

- 3.1 Dat AI grote maatschappelijke voordelen kan hebben is vrijwel onbetwist: denk aan toepassingen voor duurzame landbouw, veiliger verkeer, een veiliger financieel systeem, milieuvriendelijkere productieprocessen, betere geneeskunde, veiliger werk, persoonlijker onderwijs, betere rechtspraak, een veiliger maatschappij. Het kan mogelijk zelfs bijdragen aan de uitbanning van ziekte en armoede. Ook kan AI een belangrijke bijdrage leveren aan de groei van de industrie en de verbetering van de concurrentiekracht van de EU.
- 3.2 Zoals met elke disruptieve technologie, brengt AI echter ook risico's en complexe beleidsuitdagingen met zich mee op gebieden als veiligheid en controleerbaarheid, sociaaleconomische aspecten, ethiek en privacy, betrouwbaarheid, etc.

3.3 We staan op een cruciaal punt voor het bepalen van de (rand)voorwaarden voor de huidige en verdere ontwikkeling en inzet van AI. Deze voordelen van AI kunnen alleen duurzaam worden bereikt als ook de uitdagingen rondom AI adequaat worden geadresseerd. Hiervoor dienen politieke keuzes te worden gemaakt.

a) Ethiek

3.4 De ontwikkeling van 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?

3.5 Daar komt bij dat de ontwikkeling van AI op dit moment plaatsvindt binnen een homogene omgeving van voornamelijk jonge, blanke mannen, waardoor (al dan niet bewust) cultuur- en genderdispariteiten in AI worden ingebed, onder meer omdat AI-systemen leren op basis van trainingsdata. Deze data dienen juist te zijn en daarnaast van goede kwaliteit, gevarieerd, voldoende diepgaand en onbevooroordeeld te zijn. Er is een algemene tendens om te geloven dat data per definitie objectief zijn, maar dat is een misvatting. Data zijn makkelijk te manipuleren, kunnen biased zijn, culturele, gender en andere vooroordelen en voorkeuren weerspiegelen en fouten bevatten.

3.6 De AI-systemen die nu worden ontwikkeld bevatten geen ethische waarden. Wij mensen moeten de AI-systemen en de omgeving waarin zij worden ingezet daarvan voorzien. Ontwikkeling, inzet en gebruik van AI-systemen (zowel publiek als commercieel) dient plaats te vinden binnen de grenzen van onze fundamentele normen, waarden, vrijheden en mensenrechten. Het EESC roept daarom op tot de ontwikkeling en vaststelling van een uniforme globale ethische code voor de ontwikkeling, de inzet en het gebruik van AI.

b) Veiligheid

3.7 Het gebruik van AI in de fysieke wereld leidt onbetwist tot veiligheidsvraagstukken. Een onderscheid kan worden gemaakt tussen interne en externe veiligheid.

- *Interne veiligheid:* is het AI-systeem robuust genoeg om goed te (blijven) functioneren? Is het algoritme goed geprogrammeerd? Crasht het niet, is het bestand tegen *hacken*? Is het effectief? Is het betrouwbaar?
- *Externe veiligheid:* is het AI-systeem veilig in gebruik in de maatschappij? Handelt het niet alleen veilig in normale, maar ook in onbekende, kritieke of onvoorspelbare situaties? Wat is de invloed van het zelflerend vermogen op de veiligheid, ook als het na ingebruikname verder leert?

3.8 Het EESC vindt dat AI systemen alleen in gebruik mogen worden genomen indien deze voldoen aan specifieke eisen voor interne en externe veiligheid. Deze eisen dienen door beleidsmakers,

AI- en veiligheidsspecialisten, bedrijven en maatschappelijke organisaties gezamenlijk te worden bepaald.

c) *Transparantie, begrijpelijkheid, controleerbaarheid, verklaarbaarheid*

- 3.9 Acceptatie en duurzame ontwikkeling en toepassing van AI hangt samen met de mogelijkheid om de werking, acties en beslissingen van AI-systemen te begrijpen, controleren en verklaren, met name ook achteraf.
- 3.10 Acties en beslissingen van AI-systemen (door slimme algoritmes) grijpen in toenemende mate in het leven van mensen. Voorbeelden zijn het gebruik van AI bij informatiegestuurd politiewerk, bij de beoordeling van hypotheekaanvragen of bij het toelatingsproces tot verzekeringen. Begrijpelijkheid, controleerbaarheid en verklaarbaarheid van het beslisproces van een AI-systeem is hierbij cruciaal.
- 3.11 Op dit moment zijn veel AI-systemen al zeer moeilijk te doorgronden voor gebruikers. In toenemende mate geldt dit echter ook voor de makers van de systemen. Met name *neural networks* zijn vaak '*black boxes*' waarbinnen (beslis)processen plaatsvinden die niet langer kunnen worden doorgrond, en waarvoor geen uitlegmechanismes bestaan.
- 3.12 Het EESC pleit voor transparante, begrijpelijke en controleerbare AI-systemen waarvan de werking ook achteraf verklaarbaar is. Daarnaast moet worden vastgesteld welke beslisprocessen aan AI systemen kunnen worden overgelaten en welke niet en wanneer menselijke interventie gewenst of vereist zou moeten zijn.

d) *Privacy*

- 3.13 De privacy van AI-systemen is een onderwerp van zorg. Veel (consumenten)producten beschikken nu al over ingebodde AI: huishoudelijke apparaten, kinderspeelgoed, auto's, gezondheidstrackers, smartphones. Al deze producten communiceren (vaak persoonlijke) data met de cloudplatforms van de producenten. Zeker nu de handel in data *booming* is, wat betekent dat de gegenereerde data niet bij de producent blijven, maar worden doorverkocht aan derden, is een punt van zorg of privacy wel voldoende gewaarborgd blijft.
- 3.14 AI is bovendien in staat om, door middel van analyse van zeer veel (vaak) persoonlijke data, de keuzes van mensen op vele terreinen (van commerciële beslissingen tot verkiezingen en referenda) te beïnvloeden. Een bijzonder kwetsbare groep zijn kinderen. Het EESC is bezorgd over AI toepassingen die zich expliciet richten op beïnvloeding van het gedrag en de wensen van kinderen.
- 3.15 Voorkomen moet worden dat de toepassing van AI op persoonlijke data de werkelijke of gepercipieerde vrijheid van mensen inperkt. De EU General Data Protection Regulation (GDPR) voorziet in belangrijke mate in de privacybescherming van digitaal aangeleverde persoonsinformatie. Of het recht van mensen op geïnformeerde instemming en keuzevrijheid bij afgifte van data, maar ook op toegang tot, aanpassing van en controle van afgegeven data in de

praktijk voldoende is gewaarborgd moet, in het licht van de ontwikkelingen van AI, goed worden gemonitord.

e) Normen, standaarden en infrastructuur

- 3.16 Nieuwe normeringssystemen ter verificatie en validatie van AI-systemen, op basis van een breed spectrum aan standaarden dienen te worden ontwikkeld om de veiligheid, transparantie, begrijpelijkheid, verklaarbaarheid en ethische verantwoordelijkheid van AI-systemen te kunnen beoordelen en controleren.
- 3.17 Het EESC pleit ervoor dat de EU een eigen AI-infrastructuur ontwikkelt, bestaande uit open source en privacy respecterende leeromgevingen en datasets van hoge kwaliteit voor ontwikkeling en training van AI-systemen. De EU kan hier bovendien een (concurrentie)voordeel behalen op de mondiale markt, door het promoten van Europese verantwoorde AI-systemen'. Het EESC beveelt aan daarbij de mogelijkheden van Europese AI-certificering en labelling te onderzoeken.

f) Impact op werk, werkgelegenheid, arbeidsomstandigheden en sociale stelsels

- 3.18 Over de snelheid waarmee en de mate waarin dit het geval zal zijn verschillen de meningen, maar het is evident dat AI impact zal hebben op de omvang van de werkgelegenheid en op de aard en het karakter van veel werk en daarmee ook op sociale stelsels.
- 3.19 Brynjolfsson en McAfee van MIT noemen de huidige technologische ontwikkelingen (waaronder AI) het tweede machinetijdperk. Er zijn echter twee belangrijke verschillen: (i) de 'oude' machines vervangen voornamelijk spierkracht, terwijl de nieuwe machines denkkracht en cognitieve vaardigheden vervangen, waardoor niet alleen laaggeschoolden (*blue collar workers*) middelbaar- en hooggeschoolden (*white collar workers*) getroffen worden en (ii) AI is een *general purpose technology*, die impact heeft op vrijwel alle sectoren tegelijkertijd.
- 3.20 AI kan grote voordelen hebben wanneer het wordt ingezet bij gevaarlijk, zwaar, uitputtend, vuil, onaangenaam, repetitief of saai werk. Maar ook routiniseerbaar werk en dataverwerking en -analyse of werk waarbij planning of predictie een grote rol speelt, werk dat vaak door hooggeschoolden wordt gedaan, kan in toenemende mate door AI-systemen worden verricht.
- 3.21 De meeste banen bestaan echter uit diverse werkzaamheden. De kans dat alle werkzaamheden die iemand vervult door AI of robots worden overgenomen, lijkt klein. Vrijwel iedereen zal echter geconfronteerd worden met de automatisering van delen van zijn functie. De tijd die daarmee vrijkomt kan gebruikt worden voor andere taken, mits daarop door overheden en sociale partners wordt ingezet. Het is noodzakelijk daarbij oog te hebben de impact die deze ontwikkelingen op professionals en managers kunnen hebben en hun betrokkenheid daarbij te promoten, zodat zij de ontwikkelingen de baas blijven en geen slachtoffer ervan worden.
- 3.22 Daarnaast zullen er, ook nieuwe banen bijkomen. Niemand kan echter voorspellen welke dat zijn, hoeveel dat er zijn en hoe snel dit gebeurt. Bedrijven zoals Google en Facebook slagen erin om een gigantische waarde te ontwikkelen met een relatief klein aantal werknemers.

Bovendien zijn de nieuwe banen niet altijd goede banen. Zorg is dat met de verdere ontwikkeling van AI er straks alleen nog laagbetaalde mini-taken overblijven voor een groeiende groep flexwerkers.

- 3.23 AI zal niet alleen gevolgen hebben voor de hoeveelheid beschikbaar werk, maar ook voor de aard van het bestaande werk. AI-systemen bieden steeds meer mogelijkheden tot het volgen en monitoren van werkers, waardoor de autonomie en privacy in het geding komen. Werk wordt nu al vaak 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 de werkzaamheden en het verleren van belangrijke vaardigheden bij gebruik van AI-systemen.
- 3.24 Feit blijft echter dat technologie ons niet 'overkomt'. Overheden en sociale partners hebben de mogelijkheid om de manier waarop AI verder wordt ontwikkeld en wordt toegepast in het arbeidsproces te bepalen en moeten deze mogelijkheid ook met beide handen aanpakken. Het is daarbij belangrijk om niet alleen te focussen op waar AI toe in staat is, maar juist ook op waar mensen toe in staat zijn (creativiteit, empathie, samenwerken), op wat we willen dat mensen blijven doen en om te zoeken naar mogelijkheden om mens en machine beter samen te laten werken (complementariteit).
- 3.25 Augmented intelligence (complementariteit), waarbij mens en machine samenwerken en elkaar versterken, is de interessantste toepassing van AI omdat het gaat om mens mét machine, in tegenstelling tot mens in plaats van machine. Co-creatie is echter van groot belang: werkers moeten bij de ontwikkeling van dergelijke complementaire AI-systemen worden betrokken om ervoor te zorgen dat de AI-systemen bruikbaar zijn en dat de werker voldoende autonomie en controle (human-in-command), voldoening en plezier in het werk behoudt.

g) *Educatie en vaardigheden*

- 3.26 Het bijhouden of aanleren van digitale vaardigheden is nodig om mensen de mogelijkheid te geven om zich aan de snelle ontwikkelingen op het gebied van AI aan te passen. De Europese Commissie zet sterk in op de ontwikkeling van digitale vaardigheden met haar *Digital Skills and Jobs Coalition*. Niet iedereen zal echter in staat of geïnteresseerd zijn om te coderen en programmeur te worden. Beleid en financiële middelen zullen daarom ook gericht moeten worden op educatie en ontwikkeling van vaardigheden op terreinen die niet door AI-systemen zullen worden bedreigd (denk aan taken waarbij menselijke interactie voorop staat, waarbij mens en machine samenwerken, of waarvan we willen dat mensen die blijven doen).
- 3.27 Wanneer wordt ingezet op complementariteit van mens en AI (*augmented intelligence*) zal voor iedereen en vanaf jonge leeftijd educatie in de omgang en het werken met AI-systemen noodzakelijk zijn om te zorgen dat mensen autonomie en controle kunnen behouden in hun werk (*human-in-command*). Daarbij is met name ook onderwijs in ethiek en privacy van belang, omdat AI op die terreinen grote impact heeft.

h) Toegankelijkheid, maatschappelijke (on)gelijkheid, inclusiviteit, verdeling

- 3.28 Het overgrote deel van de ontwikkeling van AI en alle daarbij behorende elementen (ontwikkelpatforms, data, kennis, expertise) is in handen van de 5 grote technologiebedrijven (Amazon, Apple, Facebook, Google, Microsoft). Hoewel deze bedrijven voorstander zijn van de open ontwikkeling van AI, en sommigen hun AI-ontwikkelpatforms als open-source aanbieden, is de volledige toegankelijkheid van AI-systemen hiermee niet gewaarborgd. De EU, internationale beleidsmakers en maatschappelijke organisaties hebben hier een belangrijke rol te vervullen door ervoor te zorgen dat AI-systemen voor iedereen toegankelijk zijn, maar ook dat de ontwikkeling ervan in een open omgeving plaatsvindt.
- 3.29 Kapitaal bevoordelende technologische veranderingen, waarbij innovaties vooral voordelen bieden aan de eigenaars van die innovatie, verslechtert de positie van arbeid ten opzichte van die van kapitaal. Technologische veranderingen kunnen ook leiden tot (inkomens)ongelijkheid tussen mensen (zowel lokaal als regionaal en mondiaal). AI kan deze trends verder versterken.
- 3.30 Belangrijk is om deze trends nadrukkelijk te volgen en adequaat te adresseren. Er wordt al gepleit voor een AI-belasting, AI-dividend of gedeeld eigendom van AI-systemen door werknemers en werkgevers. Ook wordt meer en meer gesproken over de noodzaak van een onconditioneel basisinkomen.
- 3.31 Het EESC heeft in een eerdere opinie¹ de mogelijkheid van een digitaal dividend en de gelijke verdeling ervan teneinde positieve groeieffecten te bereiken benoemd. Het EESC hecht belang aan onderzoek naar al deze oplossingen, waarbij een juiste balans moet worden gevonden tussen de ontwikkeling van AI die de mens ten goede komt en eventuele belemmerende effecten als gevolg van die oplossingen. Ook moet het morele risico, waarbij de verantwoordelijkheid voor AI-systemen wordt afgeschoven op een entiteit die niet effectief verantwoordelijk kan worden gehouden, worden voorkomen.

i) Wet- en regelgeving

- 3.32 De implicaties van AI voor bestaande wet- en regelgeving zijn groot. De STOA van het Europees Parlement heeft in juni 2016 een overzicht uitgebracht van de Europese wet- en regelgeving die geraakt zal worden door de ontwikkelingen op het gebied van robotica, cyber-physical systems en AI. Het STOA benoemt 6 aandachtsgebieden - transport, dual-use systemen, burgerlijke vrijheden, veiligheid, gezondheid, energie - waarbinnen maar liefst 39 EU Verordeningen, Richtlijnen, Verklaringen, Mededelingen en het Europees Handvest voor Fundamentele Rechten mogelijk herzien of aangepast moeten worden. Deze evaluatie dient snel en voortvarend te worden opgepakt en het EESC kan en wil bij dit proces een rol spelen.
- 3.33 Er is veel discussie over de vraag wie aansprakelijk kan worden gehouden wanneer een AI-systeem schade veroorzaakt. Zeker wanneer het AI-systeem zelflerend is en na ingebruikname verder leert. Het Europees Parlement heeft aanbevelingen opgesteld voor burgerlijke wetgeving voor robotica, met daarin het voorstel om een zogenaamde 'e-personality' voor robots,

¹ [PB C 13 van 15.01.2016, blz. 161.](#)

inhoudende dat deze civielrechtelijk aansprakelijk kunnen worden gehouden voor schade die zij veroorzaken, te onderzoeken. Het EESC is tegen een vorm van rechtspersoonlijkheid voor robots of AI(-systemen), omdat dit een onaanvaardbaar moreel risico in zich draagt. Van het aansprakelijkheidsrecht gaat een preventieve gedragscorrigerende werking uit, welke kan verdwijnen zodra de maker niet langer het aansprakelijkheidsrisico draagt, omdat dit is verlegd naar de robot (of het AI-systeem). Bovendien is er risico op oneigenlijke toepassing en misbruik van een dergelijke rechtsvorm. De vergelijking met de beperkte aansprakelijkheid van vennootschappen gaat hierbij mank, omdat daarbij altijd een natuurlijk persoon de ultieme aansprakelijke is. Onderzocht moet in dit verband worden in hoeverre de huidige nationale en EU wet- en regelgeving en jurisprudentie op het gebied van (product- en risico)aansprakelijkheid en eigen schuld voldoende antwoord op deze vraag geven en zo niet, welke juridische oplossingen dan geboden zijn.

- 3.34 Voor een juiste benadering van wet- en regelgeving in het kader van AI, zal bovendien een goed begrip van wat AI kan, niet kan, en op korte, middellange en lange termijn mogelijk zal kunnen, cruciaal zijn.
- 3.35 AI beperkt zich niet tot grenzen. Daarom is het belangrijk om de noodzaak van mondiale regelgeving te onderzoeken, omdat regionale regelgeving onvoldoende zal zijn en zelfs onwenselijke effecten zal hebben. Gezien haar beproefde stelsel van product- en veiligheidsnormen, de hang naar protectionisme op andere continenten, het hoge niveau van kennis binnen Europa, het stelsel van Europese fundamentele rechten en sociale waarden en de sociale dialoog, beveelt het EESC de EU aan een voortrekkersrol te nemen in de vaststelling van wereldwijde, eenduidige beleidskaders voor AI, en dit proces mondiaal aan te jagen.

j) Bestuur en democratie

- 3.36 AI toepassingen kunnen een bijdrage leveren aan betere betrokkenheid van mensen bij overheidsbeleid en transparantere bestuurlijke besluitvorming. Het EESC roept de EU en nationale overheden op om AI hiervoor in te zetten.
- 3.37 Het EESC is bezorgd over het gericht gebruik van AI-systemen (in de vorm van slimme algoritmes) voor nieuwsaggregatie op bijvoorbeeld social media, welke tot vernauwing van informatie en verdere splijting van de maatschappij lijken te hebben geleid (denk aan *'filter bubbles'* en *'nepnieuws'* op Twitter en Facebook tijdens de Amerikaanse verkiezingen).
- 3.38 Het EESC is ook bezorgd over de signalen dat AI-systemen zouden zijn ingezet om het (stem)gedrag van mensen te beïnvloeden. Door middel van slimme algoritmes zouden voorkeuren en gedrag van mensen zijn voorspeld en gericht beïnvloed. Dit is een bedreiging voor een eerlijke en open democratie. In het huidige tijdperk van polarisering en ontmanteling van internationale instituties, kan de precisie en kracht van dergelijke propagandatechnologie snel voor verder ontwrichting van de maatschappij zorgen. Dit is een van de redenen waarom standaarden ten aanzien van transparantie en controleerbaarheid van (slimme) algoritmes noodzakelijk zijn.

k) Oorlogsvoering

- 3.39 De United Nations Convention on Conventional Weapons heeft besloten om in 2017 experts met elkaar te laten spreken over de implicaties van autonome wapens. Het EESC is hiermee ingenomen en ondersteunt de oproep van Human Rights Watch c.s. tot een verbod op autonome wapensystemen. Het EESC is van mening dat een dergelijk verbod serieus moet worden onderzocht en overwogen. Dit is echter niet voldoende om de mogelijkheden van AI in oorlogs- en conflictsituaties adequaat te adresseren. Ook de toepassingen van AI in cyberoorlogsvoering dienen onderwerp van dit VN-overleg te zijn.
- 3.40 Daarnaast dient voorkomen te worden dat AI in handen valt van mensen of regimes, die als doel hebben het in te zetten voor terroristische activiteiten.

l) Superintelligentie

- 3.41 Tot slot is de vraag aan de orde wat de mogelijkheden en risico's zijn van de ontwikkeling van super-intelligentie. Volgens Stephen Hawking zal de ontwikkeling van *general* AI al het eind van de mensheid betekenen. Op dat moment zal AI zich verder ontwikkelen in een tempo dat voor mensen niet is bij te houden, Aldus Hawking. Er zijn daarom deskundigen die opteren voor een zgn. 'kill-switch' of 'reset-button', waarmee we een op hol geslagen of super-intelligent AI-systeem kunnen deactiveren of resetten.
- 3.42 Het EESC pleit voor een human-in-command benadering met als randvoorwaarden de verantwoordelijke en veilige ontwikkeling en inzet van AI, waarbij machines machines blijven en mensen te allen tijde de controle over deze machines zullen kunnen blijven behouden. De discussie over super-intelligentie is daarbij belangrijk, maar overschaduwde op dit moment het debat over de impact van de huidige toepassingen van AI.

4. AI ten dienste van de mensheid

- 4.1 Grote commerciële partijen hebben inmiddels diverse initiatieven ontplooid voor de open, veilige en maatschappelijk verantwoorde ontwikkeling van AI (zoals OpenAI). Beleidsmakers kunnen dit echter niet aan bedrijven overlaten en hebben hier een rol te vervullen. Gerichte maatregelen en steun voor onderzoek naar maatschappelijke uitdagingen rondom AI en naar de ontwikkeling van veilige en robuuste AI-systemen zijn noodzakelijk.
- 4.2 EU-programma's, waaronder Horizon 2020 zijn bij uitstek geschikt om deze uitdaging te adresseren. Het EESC heeft geconstateerd dat financiering onder met name de pijler *Societal Challenges* van Horizon2020 'achterblijft' op die onder de twee andere pijlers *Excellent Science* en *Industrial Leadership* en wordt teruggeschroefd. Het EESC dringt erop aan dat onderzoek naar de brede maatschappelijke uitdagingen, alsmede naar de maatschappelijke toepassingen van AI binnen de pijler *Societal Challenges* een belangrijke plaats krijgt.
- 4.3 Ook moeten mogelijke *cross-cutting*-effecten van AI worden geadresseerd. Gelijktijdig met financiering van de ontwikkeling van disruptieve AI-innovaties dient financiering plaats te vinden van onderzoek naar en adressering van de maatschappelijke impact van die innovaties.

4.4 Onderzoek en ontwikkeling van AI ten dienste van de mensheid vereist bovendien een variëteit aan hooggekwalificeerde openlijk beschikbare trainings- en testdata en *real life* testomgevingen. Tot nu toe is de AI-infrastructuur en veel gekwalificeerde data alleen beschikbaar bij en voor een beperkt aantal private partijen en zijn er obstakels om AI in de openbare ruimte te testen, waardoor de toepassing van AI op andere terreinen wordt belemmerd. Ontwikkeling van openlijk beschikbare hooggekwalificeerde data en een Europese AI-infrastructuur zijn cruciaal om tot veilige, robuuste en nuttige AI te komen.

5. Monitoring en adressering

5.1 De brede maatschappelijke impact van AI kan nog niet volledig worden overzien. Dat de impact groot zal zijn is echter onbetwist. De ontwikkelingen op het gebied van AI volgen elkaar op dit moment in snel tempo op, wat vraagt om kritische monitoring vanuit een breed perspectief om belangrijke en disruptieve ontwikkelingen, zowel technische als maatschappelijke, in en rondom AI (*gamechangers*) adequaat en tijdig te kunnen adresseren.

5.2 Bij technische *gamechangers* kan gedacht worden aan opvallende of significante sprongen in de ontwikkeling van AI-vaardigheden, die een voorbode kunnen zijn van het bereiken van *general AI*. Bij maatschappelijke *gamechangers* kan gedacht worden aan aanzienlijk baanverlies zonder zicht op vervangende banen, onveilige situaties, systeemfalen, onverwachte internationale ontwikkelingen, etc.

5.3 Beleidsmakers, de industrie, sociale partners, consumenten, NGO's, onderwijs- en zorginstellingen en academici en specialisten uit diverse disciplines (waaronder (toegepaste) AI, ethiek, veiligheid, economie, arbeidswetenschappen, rechtswetenschappen, gedragswetenschappen, psychologie, filosofie) dienen gezamenlijk de ontwikkelingen rondom AI nauwgezet te volgen en deze *gamechangers* in kaart te brengen en up-to-date te houden, om zo op het juiste moment de juiste maatregelen te kunnen nemen, in de vorm van beleid, wet- en regelgeving, zelfregulering en sociale dialoog.

5.4 Het EESC zal, als vertegenwoordiger van het Europees maatschappelijk middenveld dit multi-stakeholder debat rondom AI in de komende periode vormgeven, centraliseren en aanjagen.

Brussel, 31 mei 2017

Georges DASSIS,
Voorzitter van het Europees Economisch en Sociaal Comité
