



Mensgerichte Artificiële Intelligentie

*Een hernieuwde oproep voor zinvolle
en verantwoorde toepassingen*

Augustus 2023

Het vitale belang van verantwoorde AI

Artificiële Intelligentie (AI) is zonder enige twijfel de meest invloedrijke technologie van dit moment. Van de rechtspraak tot het onderwijs, van de zorg tot het openbaar bestuur, en van ons dagelijks internetgebruik tot de journalistiek: overal in de samenleving speelt AI als systeemtechnologie een steeds grotere rol in ons doen en laten. Die invloed wordt zowel bejubeld als gevreesd. AI biedt veel kansen en kan een positieve invloed hebben op de maatschappij en de economie. Denk aan het versneld ontwikkelen van medicijnen voor zeldzame ziekten, het faciliteren van verduurzaming en het ondersteunen van werknemers, zodat er meer tijd vrijkomt voor handen aan het bed en in het klaslokaal. AI heeft echter ook een keerzijde. (Menselijke) vooroordelen in het ontwerp hoe data wordt ingezet, hebben door het grote bereik van AI meer impact. Het is zaak dat deze op tijd worden gezien en opgelost om te voorkomen dat op grote schaal mensen worden buitengesloten. Kunnen mensen nog wel overzien hoe AI-systemen werken, en verantwoordelijkheid nemen voor beslissingen en acties die ze nemen op basis van AI?

Deze grote positieve en negatieve impact van AI op mens en maatschappij brengt een bijzondere verantwoordelijkheid met zich mee. Zowel voor de ontwerpers en ontwikkelaars van AI-systemen als voor de gebruikers ervan. Hoe kunnen we AI-systemen ontwerpen en gebruiken om de vruchten ervan verantwoord te kunnen plukken? Hoe kunnen we bepalen of deze systemen verantwoord data gebruiken? En hoe kunnen we ze op een goede manier maatschappelijk inbedden en gebruiken? Het belang van deze vragen wordt breed herkend. Het jaar 2019 is de boeken ingegaan als het jaar van de AI-ethiek.¹ Sinds die periode is er veel gebeurd. Er is Europese wetgeving in ontwikkeling en er is een start gemaakt om burgers en medewerkers bewust te maken en te betrekken bij AI via media, educatie en evenementen. Er zijn raamwerken en methoden ontwikkeld, waarmee organisaties aan de slag zijn gegaan om ethische principes te vertalen in hun organisatie en toepassingen. Er wordt ervaring opgedaan binnen organisaties met het inrichten van governance rond AI. En sinds het beschikbaar komen van generatieve AI heeft de aandacht voor en populariteit van AI een enorme vlucht genomen. Er zijn vele nieuwe groepen gebruikers en toepassingen van AI. Er wordt gesproken over de mogelijkheden en ook over de zorgen. Bijvoorbeeld over of het toestaan van AI in het onderwijs de kwaliteit schaadt? Of door AI gemaakte kunst wel kunst is? Maakt AI mijn beroep overbodig? Deze vragen zijn heel terecht en een logisch gevolg van de enorme schaal waarop AI nu een rol speelt in ons dagelijks leven.

Ook in de internationale verhoudingen is het streven naar 'verantwoorde AI' van cruciaal belang. In de mondiale concurrentiestrijd profileert Europa zich immers steeds nadrukkelijker als 'derde weg', naast de kapitalistische benadering van 'Big Tech' in het Westen en de centralistische benadering van staatscontrole in het Oosten. In deze derde weg staat de mens centraal bij het ontwikkelen en toepassen van AI. De ambitie om AI te verbinden met menselijke waarden is de drijvende kracht. Mensgerichte AI die AI ten dienste van de gehele mensheid wil stellen, verbindt Oost en West met elkaar.

Nederland heeft de ambitie om een kernspeler te zijn met mensgerichte AI. Het wetenschappelijk onderzoek op AI-gebied is van hoog niveau en Nederland kent veel succesvolle AI-gerelateerde bedrijven. Het poldermodel van de Nederlandse economie vormt bovendien een uitstekende basis om overheid, wetenschap, bedrijfsleven en de samenleving met elkaar te verbinden. Daarnaast heeft het Nederlandse onderzoek op het gebied van Maatschappelijk Verantwoord Innoveren (MVI) en de ethiek

¹ UNESCO heeft een aantal van deze initiatieven verzameld in dit overzicht:
<https://unesdoc.unesco.org/ark:/48223/pf0000369756>

van technologie een mondiale topositie.² In de Nederlandse aanpak wordt ethiek sterk verbonden met technologisch ontwerp: ethische reflectie mondt niet alleen uit in mooie beschouwingen maar slaat ook neer in de technologie zelf. Deze pragmatische aanpak biedt een solide basis om niet alleen ethisch te *praten* over AI, maar om AI ook ethisch vorm te geven. Daarbij gaat het niet alleen om de algoritmen, maar ook om de vraag hoe, waar en door wie ze worden toegepast. Gebeurt dat rechtvaardig, met respect en met solidariteit? Algoritmen worden bovendien altijd getraind met data. Eerlijke algoritmen vergen dan ook representatieve trainingsdata.³ Bij veel toepassingen waarin AI zit, is het niet duidelijk met welke databronnen het is getraind. Als je dit niet (geheel) kunt achterhalen, maakt dat de toepassing ongeschikt voor professionele doeleinden. Transparantie is dus erg belangrijk, het geeft je een indicatie van hoe betrouwbaar de uitkomsten van de AI toepassing zijn. Maar er is meer nodig. Ook als de toepassing in gebruik is genomen, is het belangrijk dat je heel regelmatig blijft monitoren of het nog steeds (goed) doet waarvoor het is ontworpen. En dat je kunt ingrijpen als dat niet het geval is. Het inrichten van een goede governance waarin dit soort zaken zijn vastgelegd, is erg belangrijk. Met de juiste begeleiding wordt het mogelijk om op een verantwoorde manier AI in steeds meer toepassingen een rol te laten spelen en daarmee menselijke activiteiten aan te vullen.

Het wordt dan ook tijd om deze unieke kracht van Nederland optimaal te benutten voor mensgerichte Artificiële Intelligentie. Om dat duidelijk te maken laten we hieronder allereerst zien op welke manier AI de samenleving aan het veranderen is en wat daarin de uitdagingen zijn: we zitten in de 'vierde industriële revolutie', die tot de 'vijfde samenleving' leidt. Wij pleiten vervolgens voor een benadering waarin AI-systemen worden ontwikkeld vanuit een *ethisch verantwoord en maatschappelijk zinvol* perspectief. Tot slot verbinden we die benadering met de rollen die alle partijen in de 'quadrupel helix' kunnen spelen – overheid, bedrijfsleven, wetenschap en de burger – om zodoende een duurzame en verantwoorde basis voor welvaart en welzijn te realiseren.

Irvette Tempelman, voorzitter werkgroep Mensgerichte AI, over de noodzaak van mensgerichte AI: "De komst van ChatGPT is in korte tijd uitgegroeid tot een ware hype. Veel docenten zijn in paniek, omdat hun leerlingen opeens AI inschakelen voor het schrijven van huiswerkopdrachten. En het zet de brede samenleving aan tot nadenken over de kansen en de risico's van AI. Ondertussen gaat het om een ontwikkeling die we al veel langer zagen aankomen. Daarbij is het ook niet zo dat er nu pas nagedacht wordt over hoe AI-technologie op een mensgerichte manier kan worden ingezet. Al zolang ik me met dit vakgebied bezighoud, is dat onderwerp van discussie. Het is dus niet zo dat we nu opeens denken: o ja, de mens!"

"Het is belangrijk ervoor te zorgen dat mensen die met AI-technologie werken goed weten wat de aandachtspunten en beperkingen zijn. Dat betekent niet alleen dat er goede ethische en juridische kaders moeten komen, maar ook dat die ondersteund moeten worden met instructies, cursussen en opleidingen. Als Europa lopen we qua AI-ontwikkeling achter op de VS en China. Daarom is het ook zo belangrijk te beseffen dat wij maar een beperkte mogelijkheid hebben om vol op de rem te gaan trappen zoals soms wordt gesuggereerd. We kunnen echter wel vol inzetten op verantwoorde, mensgerichte AI".

De impact van AI: van 'Industry 4.0' naar 'Society 5.0'?

Artificiële Intelligentie is onderdeel van wat wel de 'vierde industriële revolutie' wordt genoemd. In deze revolutie wordt digitale technologie *fysiek*: het internet wordt een internet-der-*dingen*, en de

² Zie: Jeroen van den Hoven en Peter-Paul Verbeek, 'Nederland is ethisch gidsland in mondiale hightech' (<https://fd.nl/morgen/1251701/nederland-is-ethisch-gidsland-in-mondiale-hightech>).

³ <https://www.cbs.nl/nl-nl/over-ons/innovatie/project/position-paper-cbs-en-artificial-intelligence>

computer krijgt een lichaam als *robot*. Door deze technologische ontwikkelingen beweegt onze samenleving zich volgens sommigen in de richting van 'Society 5.0': de vijfde samenleving na de vierde revolutie.⁴ Na de tijd van de jagers-verzamelaars bracht de ploeg een overgang naar de agrarische samenleving, waarna de stoommachine de industriële samenleving mogelijk maakte en de computer tot de informatiesamenleving leidde. De vierde revolutie leidt nu tot 'de digitale samenleving', in Japan ook wel aangeduid als de 'super smart society'. Als volgende stap moeten we antwoord geven op de vraag hoe AI invloed zal hebben op de relatie tussen professionals en hun organisaties, en op het ontwikkelen van nieuwe productiviteitsconcepten zoals 'een leven lang leren' en participatieve arbeidsmarkt vaardigheden.

De kern van de maatschappelijke impact van AI ligt in haar invloed op het *cognitieve functioneren* van de mens: AI verandert het menselijke denken en waarnemen. Historisch gezien is dat niet voor het eerst. Al eerder had het schrift grote invloed op hoe ons geheugen werkt: Plato was al ongerust dat we niets meer zouden kunnen onthouden als we alles zouden gaan opschrijven, en de boekdrukkunst leidde tot een democratisering van kennis en legde de basis voor de wetenschappelijke revolutie. Een revolutie waarbij je de boot niet wilt missen, zoals het Ottomaanse rijk in de tijd van de introductie van de boekdrukkunst. Het is belangrijk om de kansen te benutten die de nieuwe technologie biedt, maar wel op een bewuste en verantwoorde manier. Wat is de nieuwe stap in ons denken die AI teweeg gaat brengen? Wat gaat er gebeuren met ons begrip van de wereld en van onszelf, onze morele en politieke opvattingen en besluiten, onze existentiële vragen en kwetsbaarheden? En hoe zal AI invloed hebben op het functioneren van onze democratie, op de relatie tussen burgers en overheid, en op mogelijkheden tot participatie aan de samenleving?

Er kunnen drie niveaus van impact van AI worden onderscheiden. Op *individueel* niveau kunnen AI-systemen de keuzes, gedragingen, en interpretatiekaders van individuele gebruikers beïnvloeden: met behulp van zoekmachines zoeken en ontvangen mensen informatie waarmee hun wereldbeeld wordt opgebouwd, expertsystemen geven artsen en rechters een kader voor een diagnose of een vonnis. Op *sociaal* niveau werkt deze invloed door in een grote variëteit van domeinen waaronder de zorg (zoals het detecteren van kankercellen in beelden), het onderwijs (adaptief lesmateriaal) en de rechtspraak (expertsystemen), verkeer en vervoer (via zelfrijdende voertuigen), en journalistiek ('*automated journalism*'). Op maatschappelijk niveau, tenslotte, beïnvloedt AI de organisatie van de samenleving: '*predictive policing*' verandert de aard van de handhavende taak van de politie, zoekmachines beïnvloeden politiek engagement en algoritmen zijn betrokken in militaire besluitvorming.⁵

Verantwoord omgaan met deze impact van AI vereist dat we nieuwe denkkaders ontwikkelen. Tot nu toe zagen we technologie vaak als 'passief' en de mens als 'actief', maar AI bezit een technologische vorm van 'autonomie'. Sommige algoritmen kunnen leren, en daardoor ontwikkelen ze zich op manieren die hun programmeurs niet altijd hadden kunnen voorzien. Ze zijn in staat om mensen te adviseren over zaken die het menselijk brein minder goed en minder snel kan overzien om zelfstandig beslissingen te nemen in bepaalde situaties.

Computers kunnen inmiddels dingen die vroeger alleen mensen konden, en omdat ze dat vaak beter en sneller kunnen, ervaren sommigen deze ontwikkeling als een bedreiging voor de mensheid. Vooral nog geldt dat mensen eindverantwoordelijkheid dragen. Maar de relatie tussen mens en

⁴ Zie: https://www8.cao.go.jp/cstp/english/society5_0/index.html

⁵ Een korte video met daarin een kernachtig beeld van de ethische uitdagingen van AI is te vinden op: <https://www.youtube.com/watch?v=TIihPx4z0Nw&t=14s>.

technologie verandert snel, en er worden steeds nieuwe vaardigheden van mensen gevraagd. Als computers sommige vormen van kanker al beter en/of sneller kunnen herkennen dan pathologen, wat gaat dat dan betekenen voor de rol van de arts? En wat is de toekomst van de journalistiek als nieuwssites niet alleen gevuld worden door journalisten maar ook door algoritmen? Hoe kunnen we controleren of het systeem de juiste analyse maakt en niet bevooroordeeld is, en kunnen we überhaupt wel verantwoordelijkheid nemen voor adviezen die tot stand komen op basis van AI?

De praktijk ligt echter aanmerkelijk genuanceerder dan een simpel streven naar vervanging van de mens door AI. Uiteindelijk gaat het steeds om het ontstaan van nieuwe *verbindingen* van de mens met AI. Artsen stellen diagnoses *in interactie met AI*, net zoals journalisten het nieuws maken *in samenwerking met* algoritmen. De cruciale vraag is niet of AI de mens wel mag vervangen, maar hoe we de mens en AI op een goede manier kunnen verbinden en met elkaar laten samenwerken, zodat we op een verantwoorde manier meer kunnen bereiken. AI kan dingen vergemakkelijken (bijvoorbeeld door een tekst voor je samen te vatten), versnellen (bijvoorbeeld door geautomatiseerd code te genereren), je helpen zoeken (snel tekst doorzoeken), navigeren (de snelste route vinden) of je bijvoorbeeld gepersonaliseerd iets aanleren op een manier die past bij jouw kennis en vaardigheid (op school of op het werk). Hierdoor kan er tijd vrijkomen om andere dingen te doen, zoals meer menselijke interactie met de patiënt of leerling. Hoe kan AI onze dagelijkse bezigheden optimaal ondersteunen op een verantwoorde manier?

Verantwoorde AI: van principes naar praktijk

In de veelheid aan ethische analyses en kaders rondom Artificiële Intelligentie komt stevast een aantal kernthema's terug. Allereerst draait de discussie vaak om *eerlijkheid en rechtvaardigheid*: algoritmen kunnen bevooroordeeld zijn, door hun eigen ontwerp of door de data waarmee ze getraind zijn, en een onrechtvaardige impact op de samenleving hebben. Een tweede centraal thema is *aansprakelijkheid*: in hoeverre is het mogelijk mensen aansprakelijk te houden voor het functioneren en de impact van AI-systemen, bijvoorbeeld voor besluiten die mede genomen zijn op basis van AI? Daarnaast zijn *transparantie en uitlegbaarheid* belangrijke ethische issues: in welke mate kan een systeem uitleggen hoe het tot zijn conclusies is gekomen, zodat mensen kunnen overzien op grond waarvan ze het systeem en de onderliggende data kunnen vertrouwen, ook als die data zelf onvolledig en onbetrouwbaar zijn? En tenslotte wordt ook de *veiligheid* van de systemen vaak benoemd: in hoeverre zijn ze betrouwbaar en bestand tegen vijandig gebruik ('dual use') en is de mens vaardig genoeg om goed met AI-systemen om te gaan?

Sinds generatieve applicaties zoals ChatGPT, Bing en Midjourney voor eenieder gratis bruikbaar werden, zijn de vragen over de voor- en nadelen van generatieve AI toegenomen. Generatieve AI (AI die teksten, audio of beelden kan maken) biedt volop kansen om mensen te ondersteunen in hun dagelijkse bezigheden, zowel in hun vrije tijd als in hun werk. Zo kan generatieve AI ondersteuning bieden bij het maken van begrijpelijke samenvattingen van complexe teksten; of suggesties doen voor het opzetten van een presentatie. Er zijn echter ook risico's aan verbonden. Zo is de output sterk afhankelijk van de data waarmee het model getraind is. Waar het systeem niet mee getraind is, kan het systeem ook geen rekening mee houden. Net als bij andere AI-systemen, spelen hier de vragen over eerlijkheid en betrouwbaarheid, aansprakelijkheid, transparantie en uitlegbaarheid en veiligheid een rol.

Ethische vraagstukken als deze zijn sterk gericht op het beschermen van de menselijke autonomie tegen de nieuwe vorm van 'autonomie' die de techniek ontwikkelt: is er nog voldoende menselijke controle mogelijk, en sluit het systeem de mens of groepen mensen niet te veel buiten in bepaalde gevallen? Er is echter ook een hele categorie ethische vragen die hierbuiten valt. Die richten zich

minder op het individu, en meer op de gemeenschap. *Inclusie* is zo'n thema: hoe waarborgen we participatie en diversiteit in onze samenleving? *Welzijn* is een ander voorbeeld: hoe ontwikkelen we AI die een positieve bijdrage levert aan het floreren van de gehele samenleving? Ook deze thema's verdienen expliciete ruimte in het debat. We moeten in gesprek met de burger die zijn stem moet laten horen. De invloed van de mens op dit proces is van belang zodat we samen leren leven met AI op een mensgerichte wijze.

Na een periode van het formuleren van kaders, principes en uitgangspunten is het nu tijd om de ethiek, de op handen zijnde Europese AI verordening (AI-Act) en verantwoord onderzoek concreet te verbinden met de praktijk. Dat vereist een andere invulling van ethiek dan het klassieke beeld van de 'ethische toetsing'. AI te vaak wordt ethiek geassocieerd met een soort externe beoordeling, zoals een medisch-ethische toetsingscommissie beoordeelt of een bepaald onderzoek wel of niet uitgevoerd mag worden. Maar de ethiek van Artificiële Intelligentie moet zich juist met de techniek *verbinden*. Het gaat erom het samenspel tussen AI, data, mens en maatschappij op een goede manier te begeleiden, en in een sterk participatief en interactief proces te zoeken naar zinvolle en verantwoorde toepassingen. Door te identificeren welke impact concrete AI-systemen en de onderliggende data kunnen hebben op mens en maatschappij, en de waarden die daarbij in het geding zijn een plek te geven in het ontwerp, de maatschappelijke inbedding en het gebruik van die systemen.⁶

Op deze manier blijft het ontwikkelen van mensgerichte en maatschappelijk verantwoorde AI op basis van eerlijke data niet bij het formuleren van algemene principes, maar wordt ethiek concreet verbonden met de technologische praktijk: het leidt tot een ethiek 'van binnenuit' en niet alleen 'van buitenaf'. Bovendien krijgt ethiek hierdoor niet alleen een beperkende, maar ook een stimulerende rol: naast het formuleren van de grenzen van wat we *niet* willen, richt ze zich op het formuleren van de voorwaarden/handelingsopties om te bereiken wat we *wel* willen.

In de AI Act staat mensgerichte AI centraal

De AI Act kent een risicogebaseerde insteek:

- a. Sommige toepassingen worden ronduit verboden (denk aan het gebruik van 'remote biometric identification' in publieke openbare ruimten of AI-systemen voor 'social crediting')
- b. Hoog risicosystemen worden zwaar gereguleerd.
- c. Voor een aantal lagere risico AI-systemen gelden transparantie-eisen (dat u bijvoorbeeld op de hoogte gesteld moet worden dat u met een deepfake te maken hebt, of dat u met een chatbot communiceert in plaats van met een mens).

Hoog-risico AI-systemen worden aan uitgebreide assessments onderworpen om te zorgen dat alleen betrouwbare AI systemen op de markt worden gebracht en toegepast. Maar een regelgevend kader alleen is niet genoeg. Nederland zal daarnaast stevig moeten investeren in mensgerichte AI-innovatie.

Naar ethisch verantwoorde en maatschappelijk zinvolle toepassingen: een oproep

Om AI en de onderliggende data te verankeren in ethisch verantwoorde en maatschappelijk zinvolle toepassingen is een nieuwe aanpak nodig, en nauwe afstemming tussen alle betrokken partijen. We stellen daarom de volgende acties voor:

⁶ Zie: P.P. Verbeek en D. Tijink, *Aanpak Begeleidingsethiek*. Den Haag: ECP, 2019:

<https://ecp.nl/wp-content/uploads/2019/11/060-001-Boek-Aanpak-begeleidingsethiek-240165-binnenwerk-digitaal.pdf>

1. Stel het streven naar maatschappelijke waarden centraal

Om AI-toepassingen ethisch verantwoord en maatschappelijk zinvol te maken is het belangrijk om uit te gaan van een goede identificatie van de sociaal-maatschappelijke problemen en uitdagingen die een brede welvaart in de weg staan. Voorbeelden hiervan zijn armoede, onveiligheid, kans-ongelijkheid, arbeidsmarkt-uitsluiting, eenzaamheid, lichamelijk ongemak en onderwijs-barrières. Verzamel data, kennis en inzichten in deze onderliggende problematiek en valideer deze zodat ze erkend worden. Stel de inzichten ter beschikking in een open context om partijen die aan innovaties werken de kans te geven zinvolle oplossingsrichtingen te exploreren en te exploiteren.

2. Het bedrijfsleven neemt medeverantwoordelijkheid voor maatschappelijke innovatie

Om deze gezamenlijke agenda te realiseren is het van belang dat het bedrijfsleven zich committeert aan 'Responsible Innovation'. Innovatie is dan niet alleen technologie-gedreven maar ook maatschappij-gedreven. Daarin verdienen juist publieke waarden een centrale rol: digitale technologie is gaan behoren tot de vitale infrastructuur van samenlevingen overal ter wereld. AI in het ontwerp en de ontwikkeling van AI-systemen dient ethiek een rol te spelen. Om dit te kunnen realiseren is nauwe samenwerking nodig met wetenschappers op het gebied van AI, data en ethiek en met maatschappelijke organisaties uit de sectoren waarvoor AI-systemen ontwikkeld worden. Ethiek van binnenuit, niet van buitenaf. Dit staat een gezond businessmodel voor bedrijven geenszins in de weg.

3. Kennisinstellingen zorgen voor verantwoord AI onderwijs en onderzoek

De wetenschap en de kennisinstellingen moeten zich richten op verantwoorde onderwijs- en onderzoekspraktijken: omdat AI zowel is gebaseerd op wetenschappelijk onderzoek als ook een belangrijke rol speelt in de hedendaagse wetenschappelijke praktijk, dienen de kaders voor wetenschappelijke integriteit en kwaliteit te worden bijgesteld in interactie met de nieuwe mogelijkheden die AI biedt. Datasets moeten op een eerlijke en transparante manier worden samengesteld en algoritmen moeten ze op een uitlegbare manier bewerken. Van waarde vrije wetenschap naar waardevolle wetenschap. De [ELSA Labs](#) die voortgekomen zijn vanuit de inzet van de werkgroep Mensgerichte AI zijn hier een goed voorbeeld van.

4. Introduceer ethische begeleiding en regulering

Voor het verantwoord inbedden van AI-systemen in de samenleving draagt de overheid een bijzondere verantwoordelijkheid. Wet- en regelgeving dient afgestemd te zijn op de nieuwe uitdagingen waar AI de samenleving voor stelt, en dat vergt ruimte voor experiment en dynamiek. Daarnaast vraagt mensgerichte AI om het stimuleren van normatieve analyse en reflectie in de praktijken zelf waar AI-systemen en de onderliggende data worden ingebed, zodat ethiek concreet wordt ingezet voor herontwerp, maatschappelijke implementatie en verantwoord gebruik. AI en data hebben bovendien invloed op het functioneren van de overheid zelf: AI-systemen helpen de overheid de samenleving te analyseren en beslissingen te nemen. Dat vergt bijzondere aandacht voor de impact van AI en de kwaliteit van data op het functioneren van de democratie, en op de manier waarop politieke kwesties worden benaderd, gerepresenteerd en geanalyseerd. Waarbij ethiek van A tot Z wordt ingebed in een organisatie.

5. Organiseer ethiek vanuit burgers en betrokkenen

De stem van de burger is onvoldoende aanwezig in debatten over AI en data. In het maatschappelijke debat zijn het vooral maatschappelijke organisaties en intellectuelen die het woord nemen, en niet de burger en de professional die de impact van AI in het dagelijks leven en werken heel direct en concreet ervaart. Deze stem van burgers en betrokkenen zou expliciet moeten worden verbonden met die van wetenschap, overheid en bedrijfsleven. Zoals 'co-design' of 'design thinking' heel gebruikelijk is geworden in het ontwerpveld (waarom zou de ontwerper beter weten wat de gebruiker wil dan de

gebruiker zelf?), en zoals in de wetenschap 'burgerwetenschap' sterk in opkomst is (waarom zouden wetenschappers een monopolie hebben op het doen van wetenschappelijk onderzoek?), zo zou de ethiek zich expliciet moeten richten op 'burgerethiek', waar het AI betreft. Zoals het aangaan van een ethische dialoog tussen zorgprofessionals en patiënten, tussen docenten en leerlingen, tussen overheid en burgers, tussen werkgever en werknemer, enzovoort. Welke concrete ethische vragen en zorgen hebben burgers ten aanzien van concrete toepassingen van AI en data? En hoe kunnen we die verbinden met onderzoek, innovatie, beleid en hun/ons eigen gedrag? Een ethiek van onderop, niet van bovenaf.

6. Verbind AI met Duurzaamheid

Steeds vaker wordt een vijfde element toegevoegd aan de quadrupel helix: de natuur en de toekomst van onze planeet. Om op een goede manier vorm te geven aan de samenleving dient AI niet alleen mensgericht maar ook natuurgericht te zijn. Ecosystemen zijn kwetsbaar en de mensheid met hen. AI en de onderliggende data hebben een substantiële impact op die systemen, vanwege de grondstoffen waarmee computers worden gebouwd en de energie die AI-systemen en dataopslag consumeren, maar ook vanuit de manier waarop AI-systemen en data het mogelijk maken klimaatsystemen beter te begrijpen en klimaatproblemen te voorspellen, analyseren en tegemoet te treden. Van quadrupel helix naar quintupel helix, en naar duurzame, mensgerichte AI.

7. Ondersteun en stimuleer brede toegankelijkheid van verantwoorde en zinvolle oplossingen

Om de doelstelling van brede welvaart te realiseren, is het belangrijk dat de AI-oplossingen die bedacht en ontwikkeld worden leiden tot een verantwoord gebruik van AI ook in commerciële toepassingen en dat deze kunnen worden opgeschaald tot proporties die alle inwoners bereiken en insluiten. Iedereen moet erop kunnen vertrouwen dat AI-systemen verantwoord worden ontworpen en gereguleerd, en kunnen profiteren van de positieve bijdrage die AI-systemen kunnen leveren aan ons welzijn en onze welvaart. In een maatschappij waarin de burger in grote mate samenleeft met digitale technologie is het van belang dat ieder mens hiervan kan profiteren. De processen moeten begrijpelijk en transparant worden gemaakt. En met name het toekomstperspectief en de normen en waarden van de eindgebruiker moeten worden meegenomen in de ontwikkeling van AI-systemen. Met andere woorden, de burger moet invloed kunnen uitoefenen op hoe men wil samenleven met AI.

Met de bovenstaande zeven aanbevelingen zetten we in op Mensgerichte Artificiële Intelligentie voor een brede samenleving en welvaart. We nodigen u uit om ze kritisch onder de loep te nemen in het licht van snelle ontwikkelingen rond het gebruik van AI. Belangrijke vragen die u zich daarbij kunt stellen zijn: wat is AI? Wat vinden we er goed aan? Wat is mensgerichte AI en hoe komen we in de praktijk tot mensgerichte AI? Hoe ontwikkelen en gebruiken we AI op een verantwoorde wijze? Hoe benutten we de mogelijkheden van AI zonder de balans te verstoren tussen individuele vrijheid en collectieve waarden en rechten? Hoe reguleren we het gebruik van AI in toepassingen met de juiste rollen van bedrijven, overheden en burger?

Aangezien AI een nieuw interdisciplinair innovatiedomein is, dat ons dagelijks leven, zorgen, leren en werken in de samenleving ingrijpend zal veranderen, willen wij als Nederlanders graag zelf bepalen vanuit welke Nederlandse cultuur en waarden we willen omgaan met AI-ontwikkelingen en dit vraagt om een brede maatschappelijke dialoog. Deze dialoog gaat de werkgroep Mensgerichte AI aan door middel van het organiseren van projecten en activiteiten in samenwerking met de andere NL AIC werkgroepen, diverse afdelingen binnen ministeries en externe partijen die aansluiten bij het vormgeven van een welvarende samenleving. Een samenleving waarbij AI de mens dient.

Colofon

Nederlandse AI Coalitie, werkgroep Mensgerichte AI

Redactie:

- Irvette Tempelman, voorzitter werkgroep Mensgerichte AI
- Jos van der Wijst,
- Náhani Oosterwijk
- Sophie Kuijt

Met speciale dank aan Peter-Paul Verbeek, lid van de werkgroep Mensgerichte AI tot en met 2022.

Wilt u meer informatie over Mensgerichte AI en het ELSA concept, dan nodigen we u graag uit [onze website](#) te bezoeken.

Website: www.nlaic.com

E-mail: info@nlaic.com