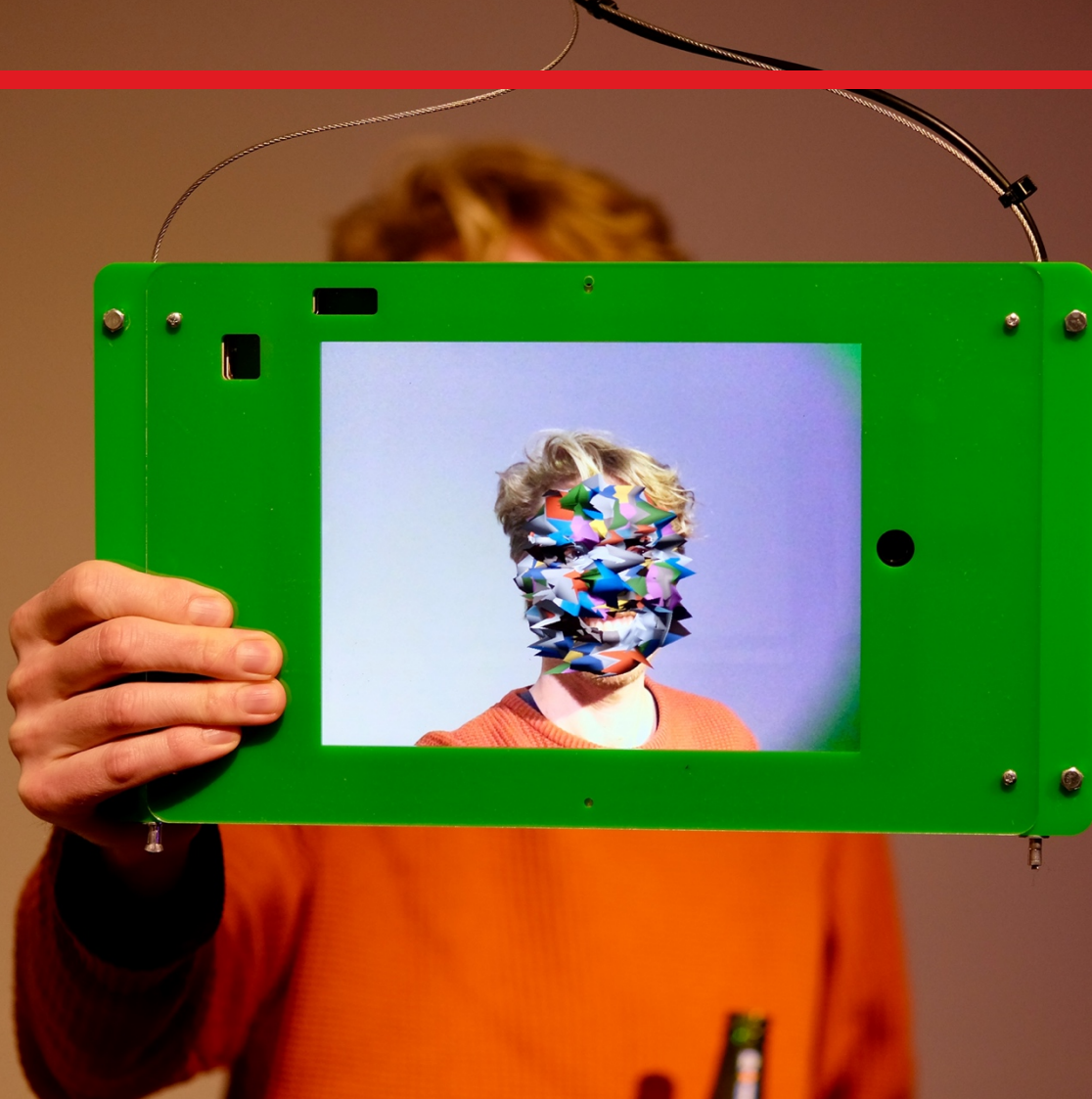




Een maatschappelijke onderzoeksagenda voor AI

Wat moet volgens de Nederlandse bevolking de focus
zijn in toekomstig onderzoek naar AI?

Waag Futurelab | Nederlandse AI Coalitie

Simone van der Burg en Jikke van den Ende

30 april 2024

waag  futurelab

NL  Coalitie

Samenvatting

Inleiding

In de afgelopen jaren hebben de ontwikkelingen in Artificial Intelligence (AI) elkaar in razend tempo opgevolgd. Daarmee heeft de technologie een grotere rol verworven in het dagelijks leven. Op school, op werk, thuis en in de publieke ruimte krijgen mensen steeds vaker te maken met nieuwe innovatieve toepassingen. Dit leidt tot interesse en inspiratie, maar ook tot vragen en zorgen onder de Nederlandse bevolking. Om te zorgen dat de bevolking invloed kan uitoefenen op de richting waarin AI zich ontwikkelt stelt de Nederlandse AI Coalitie de vraag: welke onderwerpen geven de inwoners van Nederland prioriteit als het gaat om de verdere ontwikkeling van AI? Om dit te onderzoeken heeft Waag Futurelab in samenwerking met Ipsos I&O een vragenlijst uitgezet voor een representatieve steekproef (n=1.327). Op basis van hun antwoorden hebben we een onderzoeksagenda gevormd.

Hoofdconclusies

Het overgrote deel van de Nederlandse bevolking geeft aan bekend te zijn met AI. Meer dan 70% van de deelnemers aan het onderzoek heeft er een betekenisvolle associatie bij, noemt concrete toepassingen of kenmerken of reflecteert op de waarde van AI ten opzichte van mens en maatschappij. Op basis van deze achtergrond hebben ze hun voorkeuren voor een onderzoeksagenda voor AI kenbaar gemaakt uit een lijst met 12 thema's. Hoewel veel mensen positieve mogelijkheden zien voor de ontwikkeling van AI (vooral medische toepassingen worden positief gewaardeerd) lijken zorgen over AI te overheersen als het gaat om de keuze voor onderwerpen op de onderzoeksagenda. **Figuur 1 toont de thema's op volgorde van prioriteit. 58% van de Nederlands bevolking geeft aan dat het thema "Nepnieuws, nepfoto's en polarisatie" hoog op de onderzoeksagenda geplaatst moet worden.**



Figuur 1. De twaalf thema's en de prioritering

Hoewel er wel verschillen te zien zijn tussen mensen van verschillende leeftijdscategorieën en etniciteit, komt het thema op de eerste plaats voor alle subgroepen. Mensen van alle leeftijden, genders en etniciteit, verdeeld over heel Nederland en van alle opleidingsniveaus kiezen veruit het vaakst voor dit thema. Mensen geven aan zich vooral veel zorgen te maken over de invloed en het mogelijke misbruik van *deep fakes*, (de verspreiding van) nepnieuws en het onvermogen om nep van echt te onderscheiden. Polarisatie komt in de reflecties van deelnemers niet direct terug. Zorgen lijken eerder samen te hangen met een afnemend gevoel van controle op informatie in hun eigen leven en sociale omgeving.

**“Geen grip meer te hebben op mijn eigen leefwereld.
Steeds moeilijker te onderscheiden wat echt is en nep.”**

Daarnaast geeft 37% van de Nederlandse bevolking prioriteit aan het thema machtsverhoudingen en machtsconcentraties. Mensen beschrijven hun zorgen over de oncontroleerbaarheid van AI voor de gewone mens. Wetenschappers, bedrijven en overheden ontwikkelen en gebruiken de technologie, maar hebben deze partijen wel altijd het belang van de samenleving voor ogen? Doordat AI moeilijk te controleren is voor mensen zonder specialistische kennis, kunnen kwaadwillende actoren gemakkelijk macht verkrijgen en AI ontwikkelen die hun eigen belang dient. Opvallend is dat mannen dit thema net iets vaker selecteren dan vrouwen, hoogopgeleiden wat vaker dan laagopgeleiden, en westerse migranten vaker dan niet-westerse migranten. Maar de verschillen zijn niet groot.

**“Het wordt nu ontwikkeld door partijen
die niet het algemeen belang in acht nemen.”**

Het thema **toezicht op mensen en vrijheid is gekozen door 33%** van de respondenten. Mensen uiten hun zorgen over het inleveren van hun data en het verlies van controle over wie daarvan gebruik maakt. Veel deelnemers maken zich zorgen over mogelijk misbruik van data; manipulatie, misdaden (stelen, afpersing, fraude) worden vaak genoemd, maar ook een algemeen gevoel van onveiligheid doordat er met behulp van AI derden kunnen meekijken in het alledaagse bestaan. Surveillance door bedrijven of overheden (China, Rusland, Nederland) wordt door sommige mensen naar voren gebracht als onwenselijk risico. Ook brengen veel respondenten in hun antwoorden op de open vragen naar voren dat ze vrezen dat mensen ‘dommer’ worden als ze hun eigen capaciteiten om te denken en keuzes te maken niet langer trainen, omdat ze taken delegeren aan AI. Uiteindelijk leidt dit ook tot een verlies aan vrijheid bij de vormgeving aan het eigen leven en aan de samenleving.

“Vrijheid van mensen. Zie de ontwikkelingen in China.”

Aanbevelingen

De uitkomsten van dit onderzoek vragen om een reactie van iedereen die een rol heeft in de ontwikkeling van AI in Nederland. Vooral is het nodig om een passende respons te bieden op de zorgen die ten grondslag liggen aan de keuzes van inwoners van Nederland. Wij doen daarvoor de volgende aanbevelingen:

1. Onderzoek en innovatie rond AI stimuleren die de onderwerpen op de onderzoeksagenda van de samenleving adresseert; om te beginnen moet er concrete actie worden ondernomen om nepnieuws en nepbeelden tegen te gaan; door te investeren in strategische projecten én open calls waarbij ook partijen uit de samenleving zelf kunnen deelnemen.
2. De machtsverhoudingen kantelen door projecten en calls te ontwikkelen die een maatschappelijk doel dienen en waarbij de samenleving de opdrachtgever is. Niet alleen de overheid, wetenschap of bedrijven mogen bepalen wat voor AI we ontwikkelen, maar ook de samenleving.
3. Kaders (regels) opstellen samen met de samenleving die aangeven waaraan AI moet voldoen; daarin uitwerken wat uitlegbaarheid en transparantie van AI minimaal betekent en wat het betekent om publieke waarden te waarborgen in ontwerp zoals vrijheid, democratie, inclusiviteit, privacy en veiligheid
4. Substantieel werk maken van vertrouwen in AI en in het maken van AI die vertrouwen verdient, bijvoorbeeld door:
 - a. Mensen een kans te geven (middels co-creatie) om invloed uit te oefenen op de ontwikkeling van AI en hun waarden te gebruiken als basis voor ontwerp
 - b. Kennis van en ervaring met AI vergroten bij de samenleving middels onderwijs en experimenten voor iedereen
 - c. Realiseren van een organisatie of meldpunt dat AI controleert en waar mensen naartoe kunnen om informatie te krijgen of problemen en/of misstanden te melden

Vervolg

De conclusies en aanbevelingen die voortkomen uit de onderzoeksresultaten scherpen de inhoudelijke koers aan van het programma Societal vanuit de werkgroepen Mensgerichte AI en Publieke Diensten van de Nederlandse AI Coalitie (NL AIC). Daarnaast wordt het belang bevestigd van het actief betrekken van inwoners bij de impact van AI op de samenleving. De NL AIC werkt hier samen met haar partners aan vanuit drie programmalijnen. Deze zijn AI-onderzoek, AI-wijsheid en het nationale AI-gesprek. Kennisdeling, bewustwording, transparantie en zeggenschap om vertrouwen in AI-toepassingen te vergroten, staan hierin centraal. De samenwerking met de NL AIC werkgroep Cultuur en Media wordt geïntensiveerd aangezien het thema nepnieuws als meest belangrijk werd bevonden.

Inhoudsopgave

EEN MAATSCHAPPELIJKE ONDERZOEKSAGENDA VOOR AI	1
SAMENVATTING.....	2
INLEIDING	6
ACHTERGROND	7
LIJST MET MAATSCHAPPELIJKE THEMA'S	8
METHODE EN OPZET.....	10
RESULTATEN.....	14
BEVINDINGEN UIT DE OPEN VRAGEN	14
BEVINDINGEN UIT DE VRAGEN MET ANTWOORDOPTIES	24
CONCLUSIE EN REFLECTIE	36
BIJLAGE 1: VRAGENLIJST.....	43
BIJLAGE 2: VOORBEELDEN VAN AI EN HUN BELANG OP VIJFPUNTSSCHAAL.....	45
COLOFON	46

Inleiding

In de afgelopen jaren hebben de ontwikkelingen in Artificial Intelligence (AI) elkaar in razend tempo opgevolgd. Zowel op school, op werk als in de publieke ruimte krijgen mensen steeds vaker te maken met nieuwe innovatieve toepassingen, waarvan AI een onderdeel is. Dit leidt tot interesse en inspiratie, maar ook tot vragen en zorgen onder de Nederlandse bevolking. Denk bijvoorbeeld artikelen in de krant over de inzet van AI in het onderwijs, omdat scholieren hun huiswerk laten schrijven door ChatGPT of de nieuwsberichten over *deep fakes*, waarbij met behulp van AI een video is gemanipuleerd waardoor het lijkt of een beroemde popster heeft meegedaan aan een pornofilm.

In de toekomst zullen de ontwikkelingen op het gebied van AI naar verwachting vermeerderen en versnellen. Om in Nederland tot verantwoorde en mensgerichte AI te komen, is het van belang dat de bevolking invloed kan uitoefenen op de richting waarin AI zich ontwikkelt en de toepassingen die er moeten komen. Daarom is het belangrijk om te weten welke onderwerpen de inwoners van Nederland op de agenda zouden zetten als het gaat om onderzoek naar- en ontwikkeling van AI. Welke toepassingen vinden zij interessant en belangrijk en zouden ze graag verder ontwikkeld zien worden? Welke onderwerpen zouden zij prioriteit geven als het gaat om onderzoek bij universiteiten, hogescholen en publieke en private onderzoekcentra?

Met behulp van een maatschappelijke onderzoeksagenda is het mogelijk om onderzoekers bewuster te maken van de interesses, vragen en zorgen die er bij de bevolking leven. Bovendien wordt het mogelijk om vragen vanuit de samenleving te adresseren in nieuwe onderzoeks- en innovatieprojecten rond AI.

Voorafgaand stellen we vast hoe we in dit rapport over AI praten. We vertrekken vanuit het uitgangspunt dat het ontwikkelen en implementeren van AI mensenwerk is. Daarom onderzoeken we ook wat de samenleving erover denkt, want het maakt uit wie er over de ontwikkeling van AI mag meebeslissen en meedenken. Maar het is ook kenmerkend aan AI-technologie dat het meer en meer autonoom functioneert. Een chatbot geeft bijvoorbeeld antwoord op vragen, zonder dat er steeds een persoon naast zit die monitort of die chatbot het goed doet. AI wordt ingezet, zonder dat er steeds een persoon naast zit die evalueert en besluiten neemt. In onze formuleringen in dit rapport komt die autonomie van AI soms ook naar voren. We schrijven dan niet 'mensen hebben AI geprogrammeerd om a en b te doen', maar 'AI doet a en b'. Dit betekent niet dat we vergeten dat AI door mensen is gemaakt, maar dat we willen aangeven dat AI geprogrammeerd wordt om taken uit te voeren zonder het toeziend oog van mensen.

Achtergrond

Dit onderzoek wortelt in gedachtegoed over de relatie tussen wetenschap en samenleving met een geschiedenis die teruggaat tot op de Tweede Wereldoorlog (WO2). Hoewel voor WO2 de grondgedachte was dat wetenschap leidde tot kennis en innovaties die de samenleving ten goede kwamen, ontstond na de bommen op Hiroshima en Nagasaki steeds vaker debat over die waarde¹. Dit had onder meer effect op de wijze waarop onderzoek werd georganiseerd en gefinancierd. Onderzoekers werden steeds vaker al aan het begin van een onderzoek gevraagd om expliciet te maken wat de maatschappelijke waarde is. Hoewel die waarde eerst voornamelijk in economische termen (in cijfers) werd uitgedrukt, is er na verloop van tijd steeds vaker een complexere gelaagde discussie gevoerd die ook de invloed beschouwt op menselijk leven, gezondheid, welzijn, (sociale) relaties of verhouding tot dieren en de natuur².

Deze discussie over de maatschappelijke waarde van wetenschap, kan op verschillende momenten worden gevoerd: voorafgaand aan het wetenschappelijk onderzoek om te bepalen waar het onderzoek over moet gaan³, tijdens het onderzoek of naderhand als de resultaten worden opgeleverd⁴. Steeds vaker gaat men ervan uit dat het beter is om die discussie al vroeg te voeren, want dan is er nog ruimte om iets bij te sturen aan de wetenschap en innovatie. Als de maatschappelijke feedback laat komt dan is dat meestal weinig constructief: tijd, geld en moeite zijn al besteed aan het onderzoek en eventuele negatief commentaar vanuit de maatschappij is dan vooral frustrerend en demotiverend. Het is dus beter om de maatschappij al vroeg te betrekken, want dan kunnen wetenschappers en innovators ook echt nog iets doen met de voorkeuren van de samenleving.

Om wetenschap structureel te verbinden aan maatschappelijke vraagstukken, zijn er het laatste decennium verschillende initiatieven genomen in Nederland. Zo is bijvoorbeeld de Nationale Wetenschapsagenda opgesteld in 2015 in opdracht van het Ministerie van Onderwijs, Cultuur en Wetenschap met als doel wetenschappelijke, economische en

¹ Vannevar Bush had in de VS bijvoorbeeld veel invloed op de verandering van deze science-society verhouding, door zijn publicatie *The Endless Frontier* waarin hij meer openheid bepleitte over wetenschap richting de samenleving. Bush, V. (2020). *Science, the endless frontier*. Princeton University Press. Zie voor latere invloed op financiering van onderzoek in de VS: Holbrook, J. B. (2005). Assessing the science-society relation: The case of the US National Science Foundation's second merit review criterion. *Technology in society*, 27(4), 437-451.

² van der Burg, S. (2009). Taking the "soft impacts" of technology into account: broadening the discourse in research practice. *Social Epistemology*, 23(3-4), 301-316; van der Burg, S. (2010).

³ Abma, T. A., Pittens, C. A., Visse, M., Elberse, J. E., & Broerse, J. E. (2015). Patient involvement in research programming and implementation: a responsive evaluation of the dialogue model for research agenda setting. *Health Expectations*, 18(6), 2449-2464.

⁴ Fisher, E., Mahajan, R. L., & Mitcham, C. (2006). Midstream modulation of technology: governance from within. *Bulletin of Science, Technology & Society*, 26(6), 485-496.

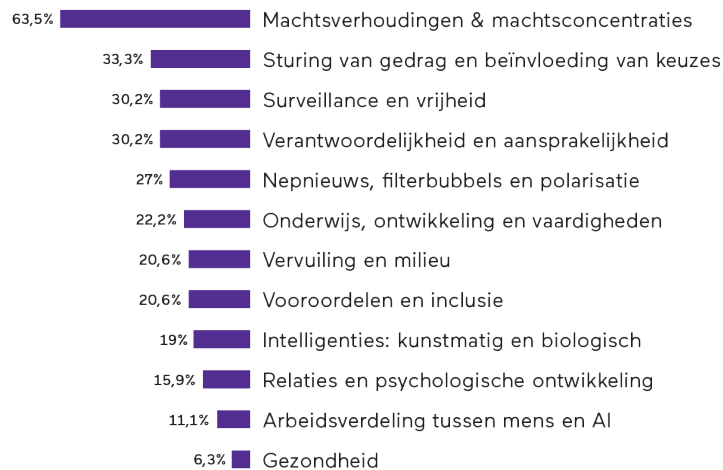
maatschappelijke vraagstukken bij elkaar te brengen. Ook heeft de Nederlandse AI Coalitie als doel om mensgerichte AI te realiseren, en zorgt de coalitie actief voor het bij elkaar brengen van wetenschap en maatschappij. Deze beide initiatieven staan aan de basis van dit onderzoek.

Lijst met maatschappelijke thema's

Oorspronkelijk is de Nationale Wetenschapsagenda (NWA) in 2015 opgesteld door vragen op te halen bij een breed publiek, om prioriteiten voor onderzoek te vormen. Deze vragen zijn ondergebracht in 140 clustervragen verdeeld over 25 verschillende routes die samen een wetenschappelijke agenda vormen, die richting geeft aan onderzoek en innovatie van NWO. Samen met de routemanagers Tibor Bosse (Radboud Universiteit) en Freek Bomhof (TNO) heeft Waag gewerkt aan een thematische focus voor de route over Big data, waar vanaf 2023 het thema AI aan werd toegevoegd.

In het kader van de route *Waardecreatie door verantwoorde Artificial Intelligence en Big Data* van de Nationale Wetenschapsagenda (NWA), heeft Waag Futurelab in het voorjaar van 2023 een klein onderzoek uitgevoerd. Het onderzoek dat leidde tot deze 12 thema's had twee stappen.

Eerst heeft Waag toegankelijke *burgerpanels* georganiseerd waarbij de 32 deelnemers werden bevraagd op hun interesses in toekomstig onderzoek naar AI, en ook op hun zorgen. Omdat we niet wisten of mensen al bekend waren met AI startten we de bijeenkomst met een introductielesing, en gingen we daarna aan de slag met voorbeelden van AI die we verspreid over de tafel neerlegden. Deelnemers konden zo zelf voorbeelden kiezen en vertellen of ze daarin interesse hadden bij vervolgonderzoek, of niet. Op basis van deze gesprekken zijn er 12 maatschappelijke thema's geïdentificeerd. In stap twee hebben we deze thema's middels een online enquête bij een grotere groep van 64 mensen neergelegd. Deze groep werd middels open vragen gevraagd per thema toe te lichten hoe ze hierover dachten en hoe ze de thema's zouden prioriteren. Ook werd er gevraagd of er nog thema's misten. Deze resultaten, gebaseerd op de input van 96 mensen in totaal, zijn vervolgens geanalyseerd en daaruit volgde een uiteindelijke lijst van 12 thema's en een eerste prioritering, te zien in Figuur 2. De onderzoeksmethodiek en resultaten van het voorgaande onderzoek zijn gepubliceerd in het eindrapport [Maatschappelijke Input Thematische Focus NWA-ORC 2024](#).



Figuur 2. De twaalf thema's en de prioritering volgens het onderzoek in 2023

De formulering van de thema's en de prioritering kwam tot stand binnen de groep van 96 respondenten (32 burgers in het panel en 64 deelnemers aan de online vragenlijst), die de moeite namen om hun rijke reflecties over de thematiek met ons te delen. Maar omdat dit onderzoek grotendeels is uitgevoerd binnen het netwerk van Waag Futurelab, waren de resultaten geen weerspiegeling van de prioriteiten van de gehele Nederlandse bevolking. Om te komen tot een meer representatieve agenda hebben we vervolgonderzoek gedaan, waarvan het resultaat in dit rapport is opgeschreven.

Methode en opzet

Opzet vragenlijst

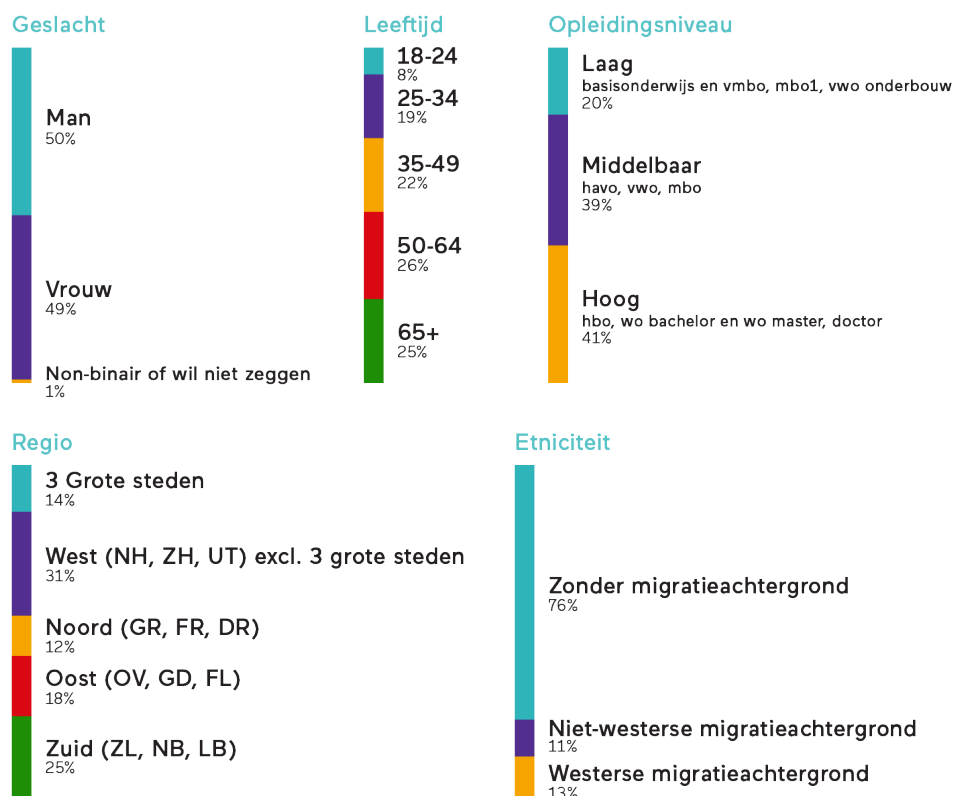
De vragenlijst voor het huidige onderzoek heeft Waag Futurelab opgesteld in samenwerking met de Nederlandse AI Coalitie en Ipsos I&O. De twaalf thema's uit het eerdere onderzoek waren hierin het uitgangspunt. We wilden het publiek niet zomaar met die thema's confronteren, omdat we niet zeker wisten hoeveel het Nederlandse publiek al weet over AI. Om die reden start de vragenlijst met drie open vragen, die nagaan wat voor associaties respondenten hebben bij AI, wat zij interessant vinden aan AI en wat zij zorgelijk vinden aan de ontwikkeling van AI. Ook hebben we gevraagd waar respondenten hun kennis over AI vandaan halen. De volledige vragenlijst is terug te zien in Bijlage 1.

De eerste vier vragen zijn aan elke respondent voorgelegd. Bij vraag 5 kreeg elke respondent 10 willekeurige voorbeelden van AI te zien, uit de totale lijst van 20 voorbeelden. Bij de voorbeelden werd een afbeelding getoond, vergezeld door een korte toelichting. Respondenten werden bij vraag 5 eerst gevraagd of ze reeds bekend waren met het voorbeeld. Ook vroegen we ze op een vijf-puntschaal aan te geven in welke mate ze het belangrijk vinden voor de samenleving dat deze toepassing van AI verder wordt ontwikkeld of beschikbaar is in de toekomst. De voorbeelden dienden een dubbel doel: we wilden weten in hoeverre mensen bekend zijn met deze voorbeelden, en in welke mate ze er belang aan hechten, maar we wilden ook zorgen dat ze wisten waarover ze spraken voordat ze begonnen aan vraag 6 waarin we vroegen welke drie thema's uit de lijst met 12 thema's zij het belangrijkste vinden om onderzoek naar te doen. Bij vraag 6 werd er een willekeurige volgorde getoond van de lijst met 12 thema's met een korte uitleg, waaruit respondenten drie thema's konden kiezen.

Selectie respondenten

Om de vragenlijst af te kunnen nemen hadden we respondenten nodig die representatief zijn voor de Nederlandse bevolking. Voor het bereiken van deze doelgroep is gebruik gemaakt van het ISO-gecertificeerd I&O Research Panel. Dit panel telt 40.000 leden. De panelleden worden actief geworven door Ipsos I&O via aselechte steekproeven uit (gemeentelijke) bevolkingsregisters of omvangrijke adresbestanden. Panelleden kunnen zich niet zelf aanmelden, omdat dan het risico zou kunnen zijn dat er alleen mensen deelnemen die echt geïnteresseerd zijn om aan onderzoek mee te doen, of dat respondenten zich aanmelden die er een hobby of beroep van maken om enquêtes in te vullen. Dit zou de resultaten kunnen beïnvloeden en minder betrouwbaar kunnen maken.

Het veldwerk door Ipsos I&O is uitgevoerd tussen 4 en 10 maart 2024. Er is aan ongeveer 3.000 panelleden een uitnodiging verstuurd. Uiteindelijk is er een steekproef gerealiseerd van $n=1.327$. De verdeling van deze steekproef is te zien in Figuur 3.



Figuur 3. De verdeling van deze steekproef

Analyse van data

Na afronding van het veldwerk is er een weging toegepast op de verkregen data: voor het uitvoeren van representatief opinieonderzoek is dit doorgaans noodzakelijk. De steekproef is na statistische weging representatief voor opleiding, leeftijd, geslacht, regio (Nielsen-districten) en etniciteit. Voor de populatiecijfers en het samenstellen van de wegingsfactoren maakt Ipsos I&O gebruik van de ideaalcijfers van de Gouden Standaard (op basis van data van het CBS). De verhoudingen van de kenmerken opleiding, leeftijd, geslacht, regio en etniciteit in de respons zijn vergeleken met de populatieverhoudingen van deze kenmerken. Groepen die ondervertegenwoordigd zijn in de steekproef kregen een groter gewicht en zijn daarom zwaarder meegeteld in het eindresultaat (weegfactoren lagen tussen de 0,5 en 2). Oververtegenwoordigde groepen hebben juist een kleiner gewicht gekregen.

De resultaten van het veldwerk zijn door Ipsos I&O opgeleverd in een tabellenboek in Excel. Per gesloten vraag is er een uitsplitsing gemaakt naar de hiervoor genoemde kenmerken. Significante afwijkingen tussen subgroepen zijn daarin expliciet aangegeven. Significantie is in dit onderzoek getoetst op het 95% betrouwbaarheidsinterval, wat betekent dat de kans kleiner is dan 5% dat de gevonden verschillen berusten op toeval. Waag heeft vervolgens de interpretatie van de uitkomsten gedaan en het rapport geschreven.

Voor de antwoorden op de eerste drie open vragen is er een analyse uitgevoerd volgens een *grounded theory* methode (Glaser and Strauss 1967; Lingard et al. 2008⁵). Eerst hebben we een semantische analyse uitgevoerd, waarbij we keken naar de inhoud van de data. We hebben stukken tekst uitgelicht en gecodeerd op basis van inhoud. Door vervolgens te kijken naar de overeenkomsten tussen de gecodeerde passages, hebben we de antwoorden onderverdeeld in categorieën en subcategorieën. Bijvoorbeeld: we kwamen in de antwoorden die respondenten gaven op vraag 1 waarin we vroegen waar mensen aan denken bij AI, vaak associaties tegen van AI met zelf-denkende computers of algoritmes. Het woord 'denken' werd daardoor een code, maar ook woorden die een vergelijkbaar idee uitdrukken zoals 'verzinnen', 'redeneren', 'slim' zijn, een 'brein' hebben of 'intelligentie'. Ook kwamen we vaak tegen dat respondenten een 'lerend' vermogen toeschrijven aan AI, of benadrukken dat AI 'zelfstandig' is, 'beslissingen neemt', 'kiest' of 'handelt'. Dit werden allemaal codes en we hebben tekstpassages geselecteerd die hierbij hoorden. Vervolgens hebben we de gecodeerde tekstpassages gegroepeerd: alles wat met denken te maken heeft bij elkaar, en ook alles wat met handelen/doen, kiezen of beslissen, zelfstandigheid en autonomie, werden allemaal groepjes. Uiteindelijk hebben we deze aldus gegroepeerde codes gerangschikt onder het thema 'AI en menselijke capaciteiten', want er komt allemaal een associatie in naar voren met capaciteiten die gewoonlijk aan mensen worden toegeschreven. Zo zijn we de open vragen doorgegaan en hebben we alle onderwerpen die we tegenkwamen gecodeerd, gerangschikt en onder thema's gehangen. Aan de hand van de themalist, zijn we uiteindelijk nogmaals de hele lijst met antwoorden langsgelopen om een beter beeld te krijgen van de grootte van de groepen die antwoorden gaven die onder elk thema thuishoorden. We hebben de categorieën die we tegenkwamen dus ook gekwantificeerd en samengevat in tabellen.

Deze analyse middels codes en thema's hebben we uitgevoerd, waarbij we steeds met twee onderzoekers naar de interpretatie van codes keken en gezamenlijk beslissingen namen over hoe ze te interpreteren en te rangschikken onder thema's. Het aantal antwoorden op de open vragen is wel lager dan het totaal aantal respondenten (per vraag n=908-1082 respondenten) omdat niet iedere deelnemer de moeite nam om een antwoord te typen. Maar de mensen die wel een antwoord gaven formuleerden vaak uitgebreide antwoorden, en adresseerden soms meerdere thema's in één antwoord. We hebben het aantal keren dat een thema voorkwam in een antwoord geteld, om een indruk te geven van het aantal keren dat dit voorkwam. Voor de resultaten van de open vragen betekent dit dat de tellingen afwijken van de totale steekproefgrootte en ook van het aantal mensen dat een antwoord formuleerde op de vraag. De gesloten vragen met voorgeprogrammeerde antwoordopties waren wel verplicht en daar is het aantal antwoorden wel gelijk aan n=1327.

⁵ Glaser B, Strauss A (1967) The discovery of grounded theory: strategies for qualitative research. Sociology Press, Mill Valley; Lingard, L., Albert, M., & Levinson, W. (2008). Grounded theory, mixed methods, and action research. *Bmj*, 337.

De analyse van de antwoorden die respondenten gaven op de open vragen, heeft ons erg geholpen om te begrijpen waarom respondenten thema's op de onderzoeksagenda willen zetten. De antwoorden op de open vragen zorgen dat we een verhaal konden vertellen bij de keuzes voor drie thema's die mensen maken. Dat verhaal is geïnformeerd door de antwoorden die mensen zelf geven op de open vragen.

Resultaten

De antwoorden die mensen gaven op de vragenlijst gaf ons een rijk en gevarieerd beeld van de bekendheid is van AI onder de Nederlandse bevolking, en van hun prioriteiten voor de onderzoeksagenda. We presenteren de resultaten in twee onderdelen: eerst de bevindingen uit de antwoorden op de open vragen, daarna uit de antwoorden op de vragen met gestructureerde antwoordopties.

Bevindingen uit de open vragen

Waar denken mensen aan bij het woord AI?

Allereerst werd de respondenten gevraagd waar ze aan denken bij het woord AI ofwel Artificiële Intelligentie. Slechts 15 respondenten geven aan geen enkel idee te hebben. 59 andere respondenten hebben brede, algemene associaties bij AI zoals 'computers', 'algoritmen', 'het internet', 'apps', of herhalen het woord, zeggend dat ze bij AI denken aan:

"Precies wat het woord aangeeft. Kunstmatige intelligentie."

Deze groep heeft duidelijk minder precieze associaties of zoekt nog naar de juiste woorden om er uitdrukking aan te geven. De overige 893 respondenten die deze vraag hebben beantwoord hebben wel een betekenisvolle associatie bij AI en velen van hen nemen de moeite om deze in een uitgebreid verhaal uiteen te zetten als antwoord de vraag. De associaties die mensen hebben bij AI variëren wel sterk. In de analyse hebben we de antwoorden onderverdeeld in twee categorieën. Enerzijds zijn er associaties die te maken hebben met concrete toepassingen van AI of kenmerken van de technologie. Ook de antwoorden van respondenten die kenmerken van AI associëren met menselijke eigenschappen behoren hiertoe. De andere categorie bevat associaties van respondenten die gaan over de gevolgen van AI voor mens en maatschappij en bevatten doorgaans evaluaties. Er zijn tellingen gemaakt van het aantal keer dat het type associatie voorkwam in de antwoorden van de respondenten. Deze worden weergegeven in figuur 4.

Een groep van 171 respondenten noemt concrete toepassingen van AI. Ze denken aan ChatGPT (74), robots (59), chatbots (of chatboxen) (22) en *smart assistants* (6), medische toepassingen (6), LLMs (5), 'smart' of zelfrijdende auto's (4), militaire toepassingen (2), neural networks (2) of GenAI (1).

Thema's: soort associatie bij AI



Figuur 4. Associaties met het woord AI en het aantal keer dat het benoemd wordt

Respondenten beantwoordden de vraag ook door omschrijvingen te geven van typerende technische eigenschappen. We kwamen 75 associaties tegen; zoals respondenten die denken aan computers die veel data of informatie kunnen verwerken of analyseren (43), die goed statistiek kunnen doen (3), die snel patronen in data herkennen (10) of processen die door computers aangestuurd worden (15). Ook *machine learning* (4) komt in deze antwoorden naar voren als kenmerkende eigenschap van AI. Voorbeelden van dit soort antwoorden zijn:

“Algoritmen die getraind zijn om patronen in taal, beelden etc. te herkennen en hierop te reageren.”

Een grotere groep van 317 respondenten beschrijft ook kenmerkend functioneren van AI, maar doet dat door metaforen te gebruiken die verwijzen naar menselijke eigenschappen. Sommigen geven aan dat AI een machine of computer is die ‘zelf denkt’ (66), dat het slimme computers zijn (52), computers die menselijk denken aanvullen (6), of overnemen (20), computers die intelligent *lijken* of die een andere (niet-menselijke) vorm van intelligentie hebben (36) of dat AI bewustzijn (2) of kennis heeft (6). Voorts vallen onder deze groep ook respondenten die AI associëren met een zelflerend vermogen (62), die aangeven dat het een computer is die ‘getraind’ moet worden om goede resultaten te leveren (12). Deze respondenten geven bijvoorbeeld aan dat zij bij AI denken aan:

“Een computer die zelf denkt en reageert op menselijke input. Vaak in combinatie met machine learning zodat de computer leert van zichzelf en het eigen functioneren analyseert.”

Andere menselijke eigenschappen worden ook vaak genoemd, zoals het vermogen van AI om zelf beslissingen te nemen (8), om zelfsturend of zelfstandig te zijn (17), om te

reageren op hun omgeving (2), om zelf te handelen of handelingen over te nemen (17), om te maken/creëren (6) of AI die dingen beter kan dan mensen (5).

Anders dan de respondenten die concrete toepassingen beschrijven of eigenschappen die kenmerkend zijn voor AI, is er de groep respondenten die AI vooral associeert met gevolgen voor mens en maatschappij. Het is opvallend een groot deel van deze groep zowel goede als slechte gevolgen noemen. Uit de antwoorden van 51 respondenten blijkt ook expliciet die ambiguïteit, zoals bijvoorbeeld blijkt uit de volgende uitspraken:

“Robots en automatisering, voelt heel dubbel, handig en goed voor de vooruitgang maar ook bedreigend, kunnen we als mens nog wel de controle behouden?”

Er is ook een groep van 276 respondenten die AI voornamelijk associeert met positieve termen. Soms zijn de gedachten heel algemeen en noemen mensen eenvoudigweg ‘vooruitgang’ of ‘toekomst’ die AI mogelijk maakt of ze noemen niet gespecificeerde ‘mogelijkheden’ die AI biedt (41). Anderen denken bij AI aan ‘gemak’ (12) en wijzen op de capaciteit van deze technologie om het leven van mensen eenvoudiger te maken, bijvoorbeeld omdat het ‘denkwerk’ zou overnemen of omdat ‘werk makkelijker wordt en informatie sneller en beter’. Anderen gebruiken het woord ‘handig’ als ze spreken over AI of ‘hulp’ (38) omdat ze bij AI denken ‘aan hulp bij het verwoorden en oplossen van vraagstukken’ of aan ‘een onafhankelijke, ontastbare entiteit die ondersteuning biedt aan de mensheid’ of aan technologie die concrete ‘hulp biedt bij diverse vragen’.

“Dat een robot taken van een mens overneemt, inclusief het denken.”

Onder de respondenten die AI associëren met ‘hulp’. Of ‘gemak’ bevinden zich ook mensen die bij AI denken aan concrete taken die AI voor mensen kan uitvoeren, zoals vertalingen bieden (3), afbeeldingen maken of bijstellen (41), teksten schrijven (46), informatie geven (26), voorspellen (9), stem- en gezichtsherkenning (3). Citaten waarin respondenten dit soort taken noemen, drukken vaak een zekere bewondering uit voor wat de technologie kan, omdat mensen niet goed kunnen begrijpen/controleren hoe het werkt:

“Door computers vervaardigde teksten en/of beelden op basis van door mensen gegeven opdrachten of vragen, waarbij de computer zijn data ontleent wat er ooit door mensen is ingebracht en dat dan op basis van voor gewone stervelingen nauwelijks te bevatten algoritmes.”

Verder associëren respondenten AI met positieve veranderingen op het werk, zoals de automatisering van taken, het efficiënter maken van taken zoals het invullen van formulieren, het voorbereiden van producten of het doen van complexe analyses of berekeningen (36). Sommigen geven aan dat AI het werk van mensen zal overnemen. Soms wordt dit gezien als iets positiefs, zoals wanneer het gaat om eentonig repetitief

werk, of huiswerk (12). Er wordt echter ook gewaarschuwd dat het kan leiden tot het verlies van banen en werkloosheid (9):

**“Welke veranderingen zal AI meebrengen in het verrichten van arbeid?
Hoeveel werkloosheid zal het gevolg zijn van AI?”**

Respondenten hebben ook negatieve associaties bij AI. Voor 88 van hen hebben deze te maken met een verlies aan controle van mensen over machines. Sommigen vrezen dat computers mensen gaan controleren of ‘de baas’ zijn (18). Hoewel slimme toepassingen het werk vergemakkelijken, snel rekenwerk verrichten of voorspellingen kunnen doen is er volgens sommigen ook een gevaar als we te veel taken delegeren. Dit zou kunnen leiden tot ‘(..) AI die het gezag overneemt van mensen’. Sommige respondenten hebben associaties met films waarin de wereld ‘wordt overgenomen door robots’, anderen redeneren heel concreet dat mensen controle verliezen omdat zij niet zelf kunnen nagaan hoe AI werkt, waardoor het moeilijker wordt voor mensen om fouten te ontdekken en AI bij te sturen of te corrigeren (8). Sommigen geven ook aan dat als mensen AI inzetten om voor hen na te denken, zij op den duur zelf de capaciteit zullen kwijtraken die nodig is om kritisch te kunnen zijn en te kunnen ingrijpen als dat nodig is (6). Mensen worden dan afhankelijk van een computer die ze niet begrijpen.

**“Makkelijker werk, maar minder voor mensen. Het leven
gemakkelijker maken en de samenleving afhankelijker maken.”**

Verlies aan controle speelt ook een grote rol in uitspraken van respondenten die AI zien als een bedreiging of gevaar (56). Sommigen associëren AI met ‘gevaarlijke gemakzucht’, met ‘angst’, ‘een bedreigende technologie’ en luiden de noodklok ‘code rood’, waarbij soms wordt gerefereerd aan Hollywood films en soms aan technologie waarvan gebruikers de werking niet goed kunnen doorzien.

Negatieve associaties bij het woord AI hebben voor 157 respondenten te maken met de capaciteit van AI om iets ‘echts’ na te bootsen. Een deel van deze groep denkt bij het woord AI aan iets ‘kunstmatigs’ of iets wat is ‘nagemaakt’ (64), waarbij het soms gaat om AI zelf (‘alles wat je ziet is nep, werkt als een menselijk brein’) en soms over teksten of beelden die door AI zijn gemaakt (‘Met computer iets nagemaakt wat net echt is’). Sommigen geven aan dat AI hen doet denken aan valse informatie of ‘fake news’ (33) of aan valse beelden of films (26) en aan *deep fake* (8). Misleiding wordt onder dit kopje ook genoemd als eerste associatie met AI (26). Deze respondenten noemen AI ‘eng’ omdat het tekst, beelden en filmpjes genereert die een gewoon mens niet van echt kan onderscheiden.

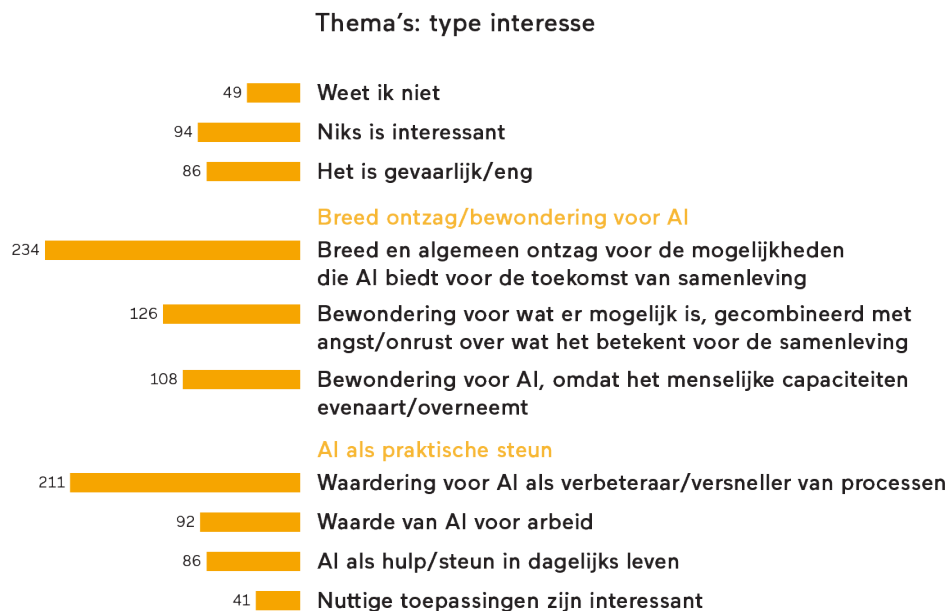
**“Een groot gevaar doordat echt en nep niet meer te onderscheiden
zijn en mensen hun hersenen niet meer gebruiken.”**

Wat vinden mensen interessant aan AI?

Respondenten werd ook gevraagd wat zij interessant vonden aan AI. De verschillende antwoorden die respondenten gaven hebben we opnieuw gecategoriseerd op basis van type interesse. De categorieën staan samen met het aantal keer dat ze genoemd zijn weergegeven in figuur 5. In totaal beantwoordden 909 mensen deze vraag.

Opvallend was dat naast een kleine groep van 49 mensen die aangaf het niet te weten, ook 94 respondenten aangaven eigenlijk niets interessant te vinden aan AI. Daarnaast gaven 86 respondenten aan AI in plaats van interessant eigenlijk 'eng' of 'gevaarlijk' te vinden.

“Eigenlijk ben ik een beetje bezorgd in plaats van geïnteresseerd.”



Figuur 5. Interesses in AI en het aantal keer dat het benoemd wordt

Er zijn ook 234 respondenten die op de vraag wat ze interessant vinden aan AI, vooral een breed en algemeen ontzag uitdrukken voor de technologie. Sommigen gebruiken algemene termen: ze noemen AI 'de toekomst' (12), 'verrassend' (4), een 'vooruitgang' (8), of een belangwekkende 'noviteit' of 'innovatie' (9), een 'snelle' of 'bijzondere ontwikkeling' (38) 'waarvan het einde nog niet in zicht is' of ze drukken brede belangstelling uit voor de 'vele mogelijkheden' (146) die AI biedt voor de toekomst van mensen, de samenleving of de wetenschap:

“De mogelijkheden die het biedt om zaken aan te pakken, te verscherpen, dingen te bedenken die ik nog niet heb bedacht.”

Ook zijn verschillende respondenten onder de indruk van wat AI allemaal kan en uiten ze bewondering voor de mensen die deze vernieuwende technologie ontwikkelen (17). Wat wel opvalt is dat in 126 deze uitingen van ontzag voor AI ook een waarschuwing naar voren komt. Veel respondenten die zeggen geïnteresseerd te zijn in de 'ongekende mogelijkheden' die AI biedt, waarschuwen tegelijk voor de gevaren, of geven aan dat hoewel de technologische ontwikkeling 'indrukwekkend is', deze nu ver genoeg is gegaan en dat die slimme mensen die AI hebben gemaakt er nu beter aan doen om 'er weer mee te stoppen'.

"Eigenlijk vind ik het vooral eng. Ik vind het vooralsnog nog niet interessant behalve dan dat ik het wel bewonder dat het mogelijk is."

Ook verwachten sommige respondenten dat AI impact zal hebben op de samenleving die weliswaar positief kan zijn, maar waarover toch ook goed zal moeten worden nagedacht omdat er gemakkelijk fouten kunnen worden gemaakt die in de toekomst moeilijk te herstellen zullen zijn, als AI eenmaal is ingebed in de samenleving. Het lijkt alsof het gebruik van AI wordt vergeleken met het openen van Pandora's doos, die naderhand lastig weer kan worden gesloten. Het is dus belangrijk om heel behoedzaam met de ontwikkeling van AI om te gaan:

"AI biedt ontzettend veel mogelijkheden en manieren het leven te verbeteren. Ik maak er nu al gretig gebruik van, al vind ik het vergelijkbaar met ontwikkelingen op het gebied van nucleaire energie. Iets, waarvan, als ik de keuze had gehad, ik had gewild dat het er nooit zou zijn geweest omdat de gevaren vele malen groter zijn dan wat het ooit aan goeds op kan leveren. De menselijke droom om onze kennis uit te breiden trekt ook aan, maar zou nooit meer mogen wegen dan het gevaar waarmee het hand in hand gaat."

Een andere groep van 108 respondenten vindt het interessant dat AI capaciteiten van mensen ondersteunt of overneemt, zoals het aanscherpen/ondersteunen van de denkcapaciteit (22), denken overnemen (32), kennis bieden (8), het zelflerende vermogen (13), het overnemen van handelen (6), het overnemen van vaardigheden (8). Ook zeggen sommige mensen het interessant te vinden dat AI lijkt op echte mensen (19). Ook deze groep heeft een brede belangstelling voor AI en kijkt met bewondering naar wat de technologie vermag.

Een andere groep respondenten kijkt meer naar de praktische kant van AI. Een grote groep van 211 respondenten is geïnteresseerd in AI als verbeteraar van processen. Verbeteren betekent dan dat processen met AI kunnen worden versneld (49) of efficiënter kunnen worden gemaakt (23), of handelswijzen kunnen worden verbeterd (30). Ook geven sommige respondenten aan dat er specifiek verbeteringen zijn te verwachten in het versnellen en nauwkeuriger maken van berekeningen (6) of het vereenvoudigen van complexe processen (26). Sommigen specificeren dat ze

verwachten dat dit voordelen biedt voor de wetenschap (34), die door AI meer data kan verwerken en patronen daarin kan herkennen en daarvoor meer kennis kan genereren dan tot dusverre mogelijk was, betere voorspellingen kan doen of problemen kan oplossen (8). Anderen verwachten verbeteringen in de geneeskunde (32), waar AI zou kunnen helpen om kennis te vergroten rond complexe ziekten en diagnoses te verbeteren of ze sneller te kunnen geven. Ook is er een klein aantal respondenten die verwachten dat AI meerwaarde heeft voor het interpreteren en toegankelijk maken van oude teksten (1) of voor vervoer en mobiliteit en water en wegenbouw (2).

Rond arbeid zijn 92 respondenten geïnteresseerd in de mogelijkheden die AI biedt voor 'het verlichten van werklast van de arbeider' (13), het overnemen van taken of werk (39). Sommigen vinden het aantrekkelijk dat AI eenvoudig, saai en repetitief werk overneemt of gevaarlijk werk (8). Maar er zijn ook respondenten die aangeven dat AI juist ingezet moet worden om ingewikkeld werk voor mensen te doen of precies werk (12), of huiswerk van scholieren (3). Sommigen zijn geïnteresseerd in de mogelijkheden die AI biedt om het werk van mensen te veranderen of te verbeteren (8) en voorzien veranderingen op de arbeidsmarkt, specifiek voor het vervangen van personeel of als oplossing voor het personeelstekort (5), anderen waarschuwen tegelijk dat dit kan leiden tot werkloosheid en dat werk van mensen moet blijven (4).

Er is ook een groep van 86 respondenten die AI waardeert in het alledaagse leven omdat het gemak oplevert of ondersteuning biedt met 'dingen waarin ik niet zo goed ben'. AI wordt gewaardeerd als vraagbaak die 'antwoord weet op de meest uiteenlopende vragen' en als informatieverschaffer of schrijver van teksten, samenvattingen en brieven.

"Het kan een handig hulpmiddel zijn met bijvoorbeeld een brief maken naar de verzekering."

Tot slot is er een groep die AI waardeert om de nuttige toepassingen (41). Hier worden geen specifieke toepassingen genoemd, maar wel dat elke technologie uiteindelijk alleen maar kan worden gewaardeerd voor wat je er mee kunt 'doen'. Ook wordt benoemd dat AI net als andere technologieën een 'instrument' is dat zowel tot goede en slechte zaken in staat is, afhankelijk van de doelen waarvoor het wordt gebruikt of de intentie waarmee het wordt ingezet:

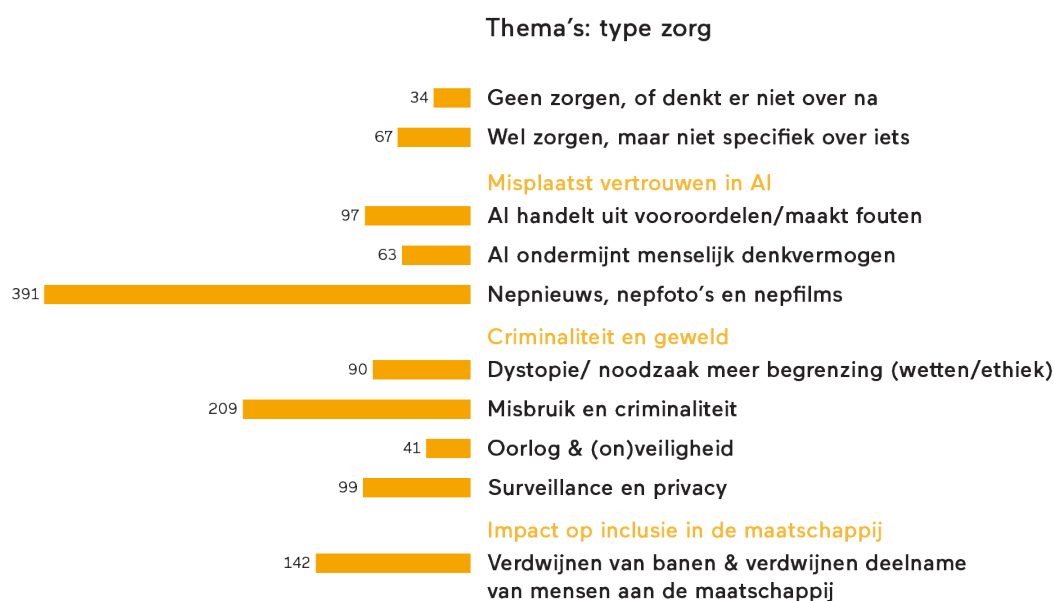
"Het is een Tool zoals een hamer dat ook is. Met een hamer kan je een raam kapotslaan maar ook een kast bouwen. Ik denk dat AI voor veel goeds gebruikt kan worden, maar dat er goede wetten en een bepaalde controle moet komen. Wij mogen het echt niet aan de Elon Musk's overlaten wat ermee gedaan wordt."

Waar maken mensen zich zorgen over bij AI?

De derde open vraag die we stelden luidde: zijn er ook dingen waar u zich zorgen over maakt bij AI? Opnieuw hebben we de respons gecategoriseerd en geteld. Het

overzicht is terug te vinden in figuur 6. In totaal beantwoordden 1081 mensen deze vraag.

Slechts 34 respondenten zeggen geen zorgen te hebben over AI, waarbij sommigen aangeven dat ze zich 'nog geen' zorgen maken, dat ze 'zich er niet in hebben verdiept' of 'dat het gezien hun gevorderde leeftijd eerder jongeren betreft'. 67 anderen geven aan zich wel zorgen te maken, maar specificeren niet waar die zorg over gaat. Ze geven aan zich zorgen te maken over 'alles', zich 'zeker' zorgen te maken, het eng/gevaarlijk te vinden of onzeker te zijn over de toekomst met AI.



Figuur 6. Zorgen rondom AI en het aantal antwoorden per categorie

Een groep van 97 respondenten maakt zich zorgen over de mogelijkheid dat AI verkeerde beslissingen neemt of handelt vanuit vooroordelen die zijn ingebed in het algoritme van AI of op basis van onbetrouwbare informatie/data (38). Sommigen wijzen op het gevaar dat partijen willens en wetens informatie verdraaien, omdat ze daarbij belang hebben. Dat kunnen grote bedrijven zijn met een commercieel belang (14), of regeringen met politieke belangen (27). Dit is volgens respondenten vooral gevaarlijk als mensen vertrouwen op de informatie die AI biedt of de beslissingen die AI neemt en dit 'zelf niet meer controleren'. Ook zijn er respondenten die opmerken (18) dat het niet helpt als het algoritme niet transparant is en je niet kunt zien (of controleren) hoe het functioneert en of het wellicht fouten of vooroordelen bevat.

"(...) Daarnaast zie je dat als er geen controle is op de trainingsdata of dat er geen verantwoording is over hoe beslissingen worden gemaakt er foute beslissingen worden genomen die moeilijk terug te draaien zijn als de beslissing weerslag heeft op de werkelijkheid."

Een groep van 63 respondenten maakt zich zorgen over het groter wordende denkvermogen van AI, en het kleiner wordende denkvermogen van mensen. Zij vrezen dat als mensen eraan gewend raken om op AI te vertrouwen, ze zelf op den duur lui worden en niet langer controleren of AI het bij het rechte eind heeft. Een aantal geeft aan (33) dat mensen op den duur ook niet meer daartoe in staat zijn: ze worden zelf minder slim, als ze hun denkvermogen niet trainen. Het verlies aan creatief denken als er te veel taken aan AI worden gedelegeerd wordt hierbij door sommige respondenten ook specifiek naar voren gebracht (12). Zij wijzen op het gebruik van AI door scholieren voor het maken van werkstukken en verslagen, waardoor ze van dit soort oefeningen niet meer leren. Volgens sommigen zal dit er uiteindelijk toe leiden dat AI 'te slim' wordt voor mensen, en 'de wereld overneemt' (18). Daarbij helpt het niet dat de systemen weinig toegankelijk zijn (black box) (6).

“Dat de mens niet bestand zal zijn tegen manipulaties en dat het denkvermogen wordt gereduceerd.”

Volgens 90 respondenten is het beter om eerst goed na te denken over de begrenzing die aan AI moet worden opgelegd. Sommigen van hen schetsen mogelijke dystopische gevolgen (27) die moeten worden voorkomen, anderen benoemen algemene (ethische) grenzen die aan AI moeten worden opgelegd (53). Respondenten wijzen erop dat het verkeer is dat 'we als samenleving totaal niet bezig zijn met de mogelijke gevolgen', dat 'de menselijke maat verdwijnt', en dat AI 'veel energie gebruikt en slecht is voor het milieu'.

Misbruik is een onderwerp dat respondenten bij eerdere vragen ook naar voren hebben gebracht. Onder het kopje 'zorgen rond AI' doen ze dat opnieuw. 209 respondenten noemen problemen gerelateerd aan kwaadwillend gebruik van AI. Misbruik (90) wordt veelvuldig genoemd en sommigen specificeren wat voor soort misbruik ze verwachten, zoals: van misleiding (met valse informatie) (48), oplichting (12), chantage, criminaliteit in het algemeen (19), diefstal van kunst (6), fraude (10) en plagiaat (12).

“(..) Niet alleen over de privacy, maar ook over hoe kunst "gestolen" wordt van echte kunstenaars en hoe beelden soms zo echt kunnen lijken dat het lijkt alsof mensen iets gedaan/gezegd kunnen hebben, terwijl ze dat niet echt hebben, maar daardoor wel in de problemen kunnen komen (door bijvoorbeeld een extremistische mening of porno/naaktbeelden die van hen gemaakt worden).”

Over veiligheid en oorlog maken 41 respondenten zich zorgen. Ze vrezen dat AI wordt ingezet om oorlogsdoeleinden te dienen (22), dus dat AI wordt ingezet om doelwitten op te sporen, wapens te besturen en dat er op een afstandelijke manier oorlog wordt gevoerd zonder menselijke tussenkomst. Sommigen maken zich zorgen over oorlogen die louter door robots worden gevoerd (4). Ook noemen respondenten veiligheid (security) als belangrijk thema (15). Dit relateert nauw aan het thema surveillance en

privacy waarover 120 respondenten zich zorgen maken. Zij wijzen erop dat partijen zoals Rusland of China, de Nederlandse overheid, of grote techbedrijven AI kunnen gebruiken om burgers te bespioneren (58), bijvoorbeeld via chatbots, smart homes en smart assistants. Ook maken (41) mensen zich zorgen over verlies aan privacy, doordat AI aan de haal gaat met foto's en films, of andere private data op straat legt.

Nepnieuws en nepfoto's en films wordt door een groep van 391 respondenten naar voren gebracht als probleem. Een groot deel (128) van hen maakt zich zorgen over het feit dat AI informatie en foto's kan genereren die moeilijk van echt zijn te onderscheiden: 'Je weet niet of het echt is of gemaakt door een computer'. Dit merken velen (in verschillende bewoordingen) op. Ook schrijven respondenten veelvuldig dat het steeds moeilijker wordt 'te controleren' wat echt is en wat niet. Daarbij noemen sommigen (61) specifiek *deep fakes*, dus het gebruik van foto's en films voor nieuwe foto's en films van situaties die nooit hebben plaatsgevonden. Ook noemen velen (202) misleiding door valse nieuwsberichten of nepnieuws als groot probleem, omdat dit zorgt dat mensen geen inzicht meer hebben in de wereld waarin ze leven.

"Onderscheid tussen waar en onwaar, tussen echt en onecht vervaagt. Hoe betrouwbaar is de bron? Wat is de bron?"

Een laatste brede groep van 142 respondenten maakt zich zorgen over de overname van werk door AI en de bredere maatschappelijke gevolgen die dat heeft. Sommige respondenten (61) maakt zich zorgen over het verdwijnen van banen (of werkgelegenheid) als AI en robots taken overnemen. Zij wijzen op negatieve effecten die dat kan hebben voor de participatie van mensen in de maatschappij, als zij geen werk meer hebben. Respondenten (13) noemen het risico dat robots mensen gaan vervangen en een groep (58) schetst schrikbeelden van AI als 'overheerser van mensen'. Onder hen zijn mensen die verwijzen naar science fiction literatuur of films die een dergelijk scenario hebben uitgewerkt, anderen geven een andersoortige onderbouwing voor hun vrees: ze geven bijvoorbeeld aan dat er 'steeds minder contact is tussen mensen' als AI teveel overneemt, of vragen zich af of 'de mens nog wel beheer heeft over zijn eigen leven' als hij zelf niks meer kan. Ook waarschuwen sommige respondenten (10) voor bredere maatschappelijke gevolgen van de overname van werk en andere bezigheden door robots, namelijk 'dat mensen zich afkeren en niet meer mee willen doen in de toekomst', wat leidt tot existentiële vragen 'over waarom je eigenlijk leeft'.

"Dat ze in het dagelijks leven ervoor zorgen dat mensen niet meer nodig zijn."

Hoe komen mensen aan hun kennis over AI?

De vierde vraag ging over de bronnen waar respondenten hun kennis van AI vandaan halen. Deze vraag is door alle respondenten ingevuld. Figuur 6 laat zien waar wat de verdeling is onder de gehele steekproef. Driekwart van de respondenten geeft aan dat ze hun kennis via het nieuws hebben verkregen.

Daarnaast geeft 40% aan zijn of haar kennis te hebben vergaard door het kijken naar televisieprogramma's. Een groot deel haalt ook kennis uit sociale media. Voor

jongeren onder de 25 geldt zelfs dat dit de voornaamste bron van kennis is. Het verschil met oudere leeftijdscategorieën is duidelijk te zien: daar wordt de bron een stuk minder vaak geselecteerd. Ook voor niet-westerse migranten geldt dat sociale media vaker een bron van kennis over AI zijn. Respondenten geven aan dat zij minder vaak aan hun kennis komen via werk of familie en vrienden. Jongeren tussen de 18 en 24 en 25 en 34 geven deze bronnen wel vaker aan dan respondenten tussen de 50 en 64 en boven de 65.

Overige bronnen waar respondenten hun kennis vandaan halen zijn tijdschriften, het internet, (wetenschappelijke) artikelen, opleiding, boeken, films, podcasts, cursussen of de krant.

Hoe komt u aan uw kennis over AI?



Figuur 7. Bronnen van kennis over AI

Bevindingen uit de vragen met antwoordopties

Voorbeelden van toepassingen van AI

Na de open vragen hebben we vragen gesteld met voorgestructureerde antwoorden. Deze zijn door alle respondenten ingevuld. Vraag 5 bestond voor elke respondent uit 10 willekeurige voorbeelden van toepassingen van AI. Op de vraag of respondenten al eens eerder (voor dit onderzoek) van de toepassing gehoord hadden kwam veelal het antwoord "ja" terug. Het overgrote deel van de voorbeelden zijn onder een ruime meerderheid (tussen de 60% en 90%) van de bevolking bekend. Dit percentage ligt voor sommige voorbeelden zelfs boven de 90%, zoals voor de digitale assistent of gezichtsherkenning. Enkele voorbeelden, zoals AI voor planning en optimalisering of AI voor de beveiliging van computersystemen, zijn bekend onder ongeveer de helft van de respondenten. Slechts bij drie toepassingen geeft de meerderheid aan er nog niet bekend mee te zijn: AI voor de selectie van sollicitanten (66% is hier onbekend mee), seksbots (71%) en AI voor schadedetectie (73%).

Tabel 1 geeft de verdeling weer van de antwoorden op deel twee van de vraag bij ieder voorbeeld. Daarbij zijn de categorieën “belangrijk” en “heel belangrijk” samengevoegd. Datzelfde geldt voor “niet belangrijk” en “helemaal niet belangrijk”. De kleuren in de tabel geven de verdeling goed weer: hoe donkerder de achtergrond van de cel, hoe hoger het percentage respondenten dat deze antwoordoptie gekozen heeft. De verdeling op de antwoordschaal per voorbeeld buigt voor sommige voorbeelden één kant op, maar bij sommige voorbeelden is er meer verschil terug te vinden in het belang dat de respondenten eraan hechten en liggen de percentages op de schaal dichter tegen elkaar aan.

	(heel) belangrijk	niet belangrijk, niet onbelangrijk	(helemaal) niet belangrijk	weet niet/ geen mening
Medische technologie met AI	81%	9%	5%	4%
AI voor beveiliging van computersystemen	76%	9%	4%	10%
AI voor herkennen van schade aan wegen of bruggen	75%	14%	5%	5%
Robots in de voedingsindustrie	71%	18%	9%	3%
AI voor fraude en risico's door banken	70%	16%	7%	7%
Slimme vertaalmachines	70%	21%	6%	3%
AI voor planning en optimalisering	66%	22%	7%	4%
Militaire technologie met AI	65%	15%	12%	8%
Leerplatforms voor kinderen zoals sQula	62%	18%	11%	9%
Robots in de zorg	56%	22%	16%	6%
Sporthorloges zoals de Fitbit	41%	30%	24%	5%
Technologie voor gezichtsherkenning	39%	31%	26%	4%
Chatbots zoals ChatGPT	31%	25%	35%	9%
Zelfrijdende auto's	29%	23%	45%	2%
Chatbots voor digitale klantenservice	26%	30%	41%	3%
Digitale assistent als Google Home of Siri	23%	33%	39%	5%
AI voor de selectie van sollicitanten	12%	22%	57%	9%
AI voor aanbevelingen	12%	30%	54%	4%
Seksbots	9%	20%	64%	7%
Deepfake technologie	8%	16%	68%	8%

Tabel 1. Voorbeelden van AI en hun belang. “Hoe belangrijk vindt u het voor de samenleving dat [dit voorbeeld] beschikbaar is of verder ontwikkeld wordt in de toekomst?” (% genoemd in de categorie; n=1.327)

Kijkend naar het belang van het bestaan van de verschillende toepassingen en de verdere ontwikkeling daarvan, dan wordt deze voor medische technologie met AI het hoogst gewaardeerd. Dat betekent dat het overgrote merendeel (81% van de respondenten) het belangrijk of heel belangrijk vindt dat medische toepassingen verder worden ontwikkeld. Slechts 5% van de respondenten geeft aan het (helemaal) niet belangrijk te vinden. Mannen hechten er vaker belang aan dan vrouwen (86% t.o.v. 77%) en vrouwen staan er vaker neutraal in dan mannen (12% t.o.v. 6%). Deze verschillen zijn in soortgelijke mate ook terug te vinden in opleidingsniveau voor hoogopgeleiden in vergelijking met laag- en middelbaar opgeleiden. In de kwalitatieve analyse kwamen er nauwelijks zorgen rondom medische toepassingen van AI voorbij en bij de vraag naar de interesse werden de medische mogelijkheden vaker genoemd. Slechts een enkeling gaf aan zich er zorgen over te maken (“als er bijvoorbeeld door AI op medisch niveau uitspraken worden gedaan”) maar de algemene tendens is dat mensen de ontwikkelingen van AI binnen dit domein hoog in het vaandel hebben.

De daaropvolgende “belangrijkste” voorbeelden volgens de respondenten zijn AI voor de beveiliging van computersystemen of AI voor het herkennen van schade. Opvallend is dat bij beide voorbeelden een groot deel van de respondenten aangeeft ze niet te kennen, terwijl een ruime meerderheid (76% en 75%) het (heel) belangrijk vindt dat dit beschikbaar is of verder ontwikkeld wordt in de toekomst. De percentages mensen die aangeven het (helemaal) niet belangrijk te vinden zijn in beide gevallen klein. Het beveiligen van computersystemen of het detecteren van schade aan wegen of bruggen is niet iets waar veel mensen tegen zullen zijn. Toch is het opvallend dat er kennelijk wel het vertrouwen is in de AI-technologie dat zij deze taken op zich kan nemen. Dat zwakt de aanname dat men over het algemeen AI-technologie wantrouwt in dit geval een beetje af, omdat het in beide gevallen gaat over de gedeelde verantwoordelijkheid over (online) veiligheid die gegeven wordt aan AI-technologie.

Opvallend zijn de resultaten rondom de toepassing van militaire AI-technologie. In de open vragen kwamen redelijk wat vraagtekens voorbij over het gebruik van AI voor militaire doeleinden. Respondenten spraken veelvuldig hun zorgen uit over potentieel misbruik van AI of benoemden de angst dat de technologie in verkeerde handen terechtkomt. Echter, het merendeel van de respondenten (65%) geeft vervolgens bij vraag vijf aan het (heel) belangrijk te vinden dat deze technologie beschikbaar is of verder ontwikkeld wordt in de toekomst. Het verschil met neutraal (15%) en (helemaal) niet belangrijk (12%) is vrij groot. Kennelijk gaan de zorgen rondom misbruik van AI niet over toepassingen op militair gebied, wat je in eerste instantie misschien wel zou verwachten. De verschillen zijn groot: ondanks dat er veel zorgen over zijn vindt twee derde van de respondenten dat de technologie wel beschikbaar moet zijn. Een mogelijke verklaring zou kunnen zijn dat dit gaat over bescherming: als we ons in Nederland hier niet verder op door ontwikkelen, kan dit tot een nadeel leiden op het moment dat we aangevallen worden. De angst voor militaire technologie maakt dat er in de technologie geïnnoveerd wordt omdat er anders achterstanden optreden.

Gekeken naar de toepassingen waarvan de meerderheid aangeeft het (heel) onbelangrijk te vinden dat het beschikbaar is of verder ontwikkeld wordt in de toekomst, roept *deepfake technologie* de meeste weerstand op. 68% van de respondenten geeft aan het (heel) onbelangrijk te vinden. Slechts 8% vindt het (heel) belangrijk (iets meer mannen dan vrouwen) en 16% is neutraal. Ook aan seksbots hechten respondenten minder belang, of aan AI voor aanbevelingen (bijvoorbeeld in algoritmes op Netflix of Facebook) en voor de selectie van sollicitanten. Voor seksbots geldt dat mannen (15%) de toepassing vaker (heel) belangrijk vinden dan vrouwen (4%). Dit verschil zien we ook terug voor etniciteit, waar westerse en niet-westerse migranten de seksbot vaker (heel) belangrijk vinden dan mensen zonder migratieachtergrond. Jongeren tussen de 18 en 24 jaar geven vaker aan het (helemaal) niet belangrijk te vinden ten opzichte van de overige leeftijdsgroepen. Wat betreft AI voor aanbevelingen vinden de leeftijdscategorieën boven de 35 de toepassing vaker (helemaal) niet belangrijk dan jongeren onder de 24. Hoogopgeleiden (16%) lijken het dan weer vaker belangrijk te vinden dan laagopgeleiden

(7%) en niet-westerse migranten (22%) vaker dan mensen zonder migratieachtergrond (11%) en westerse migranten (8%). Voor AI voor de selectie van sollicitanten vinden middelbaar- (58%) en hoogopgeleiden (61%) het vaker (helemaal) niet belangrijk dat dit beschikbaar is vergeleken met laagopgeleiden (45%). Niet-westerse migranten (30%) kiezen juist vaker voor (heel) belangrijk dan autochtonen (10%) en westerse migranten (75). Ondanks de significante verschillen tussen de subgroepen in de reacties op deze voorbeelden is de spreiding wel vergelijkbaar met de algemene tendens, met het zwaartepunt van de antwoorden bij (helemaal) niet belangrijk.

Er zijn ook voorbeelden waarover de meningen meer verdeeld zijn. Sporthorloges zoals de Fitbit, gezichtsherkenning, ChatGPT, zelfrijdende auto's, chatbots voor digitale klantenservice en de digitale assistent roepen verschillende reacties op onder de respondenten. In sommige gevallen zijn er categorieën te vinden waartussen de antwoorden met name verschillen. Voor ChatGPT geldt bijvoorbeeld dat mannen (40%) het vaker (heel) belangrijk vinden dan vrouwen (23%) en vrouwen (42%) vinden het dan ook vaker (helemaal) niet belangrijk dan mannen (28%). Respondenten onder de 50 zijn ook positiever dan respondenten boven de 50 en hoogopgeleiden vinden het vaker (heel) belangrijk dan laag- en middelbaar opgeleiden. Ook vinden niet-westerse migranten het vaker (heel) belangrijk dan autochtonen en westerse migranten. Ook voor de zelfrijdende auto geldt dat mannen, hoogopgeleiden en respondenten uit de drie grootste steden het vaker (heel) belangrijk vinden dan respectievelijk vrouwen, laag- en middelbaar opgeleiden en respondenten uit de overige regio's in Nederland.

Thema's voor de onderzoeksagenda

Ten slotte werd de respondenten gevraagd welke thema's zij graag op de onderzoeksagenda rondom AI zouden willen plaatsen. Als we kijken welke thema's het meest gekozen werden, is de prioritering als volgt:



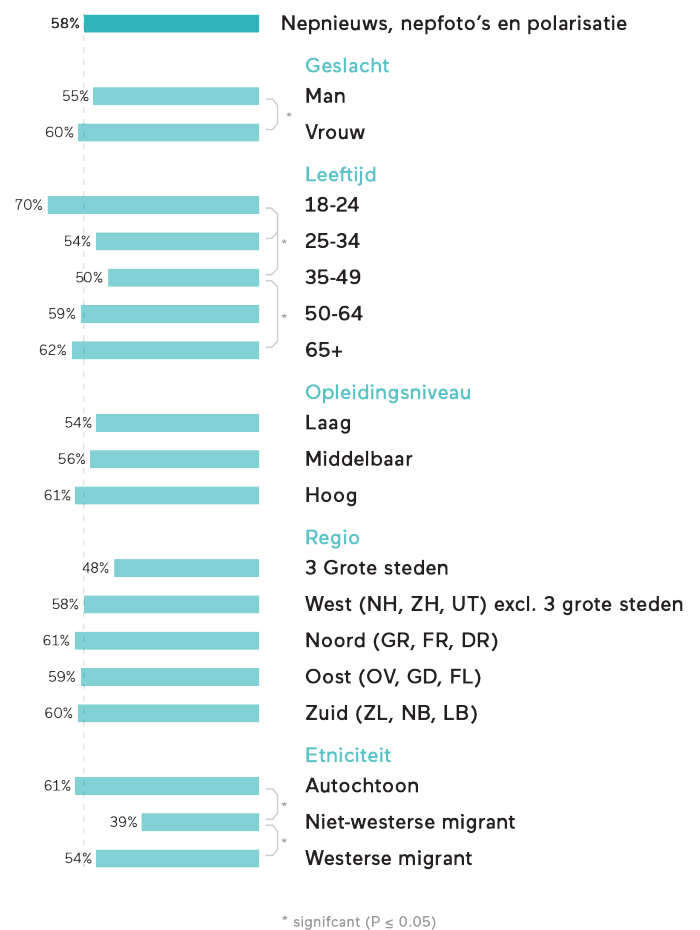
Figuur 8. De twaalf thema's en de prioritering. Kunt u uit deze 12 thema's er drie kiezen die waarvan u vindt dat wetenschappers daar rekening mee moeten houden als ze in de toekomst onderzoek doen naar AI? (% gekozen in de top-3, n=1.327)

Nepnieuws, nepfoto's en polarisatie

Dit is het meest geselecteerde thema. 58% van de respondenten vindt dat wetenschappers hiermee rekening moeten houden als ze in de toekomst onderzoek doen naar AI. Opvallend is dat als er gekeken wordt naar de verschillende subgroepen, er nauwelijks significante verschillen zijn. Dat betekent dat de aandacht voor dit thema door iedereen belangrijk wordt gevonden. Gekeken naar gender is er geen significant verschil te vinden: 55% van de mannen heeft het thema geselecteerd en 60% van de vrouwen. Onder jongeren tussen de 18 en 24 jaar wordt het thema vaker geselecteerd dan door jongvolwassenen tussen 25 en 34 en volwassenen tussen 35 en 49 jaar. Ook ouderen boven de 65 selecteren het thema iets vaker dan in de groep 35 tot 49-jarigen. De verschillen zijn echter klein. Bovendien geldt dat voor elke leeftijdscategorie dit thema het vaakst geselecteerd wordt vergeleken met de overige thema's.

Ook qua opleidingsniveau en regio bestaan er geen significante verschillen tussen de subgroepen. Het enige wat opvalt is dat onder mensen met een niet-westerse migratieachtergrond het thema minder vaak wordt geselecteerd dan onder westerse migranten of mensen zonder migratieachtergrond. Deze verschillen zijn wel significant. Gekeken naar de respons van mensen met een niet-westerse migratieachtergrond is te zien dat het verschil tussen het aantal mensen dat dit thema koos en het aantal dat koos voor een ander thema, namelijk "verantwoordelijkheid en aansprakelijkheid", een stuk kleiner is dan voor alle andere subgroepen. Ondanks dat het thema nepnieuws nog steeds het meest geselecteerd wordt, is het voor deze subgroep een minder duidelijke prioriteit.

De keuze voor dit thema past bij de antwoorden die respondenten gaven op de eerdere vragen. Ruim twee derde gaf aan het (helemaal) niet belangrijk te vinden dat deep fake technologie beschikbaar is of verder ontwikkeld wordt en respondenten uitten veelvuldig hun zorgen rondom nepnieuws en nepbeelden in hun antwoorden op de open vragen. Hoewel mensen het interessant vinden dat ze veel informatie kunnen vinden, en ook de creatieve mogelijkheden die AI biedt voor beeldredactie waarderen, vrezen ze ook dat de technologie verkeerd ingezet wordt en daarmee nadelige effecten kan hebben op



Figuur 9. Het thema 'Nepnieuws, nepfoto's en polarisatie' en de prioritering uitgesplitst voor geslacht, leeftijd, opleidingsniveau, regio en etniciteit.

individuele of op de samenleving in het algemeen. Er wordt vaak een link gelegd met sociale media: filmpjes van politici of andere bekenden waarvan het lastig te bepalen is of het echt is of niet. Daarbij ontstaat bij veel mensen de zorg zichzelf of anderen niet in staat zijn om echt van nep te onderscheiden. Ook zijn zij bezorgd dat mensen alles zomaar voor waar aannemen, zonder überhaupt de vraag te stellen of het de waarheid is.

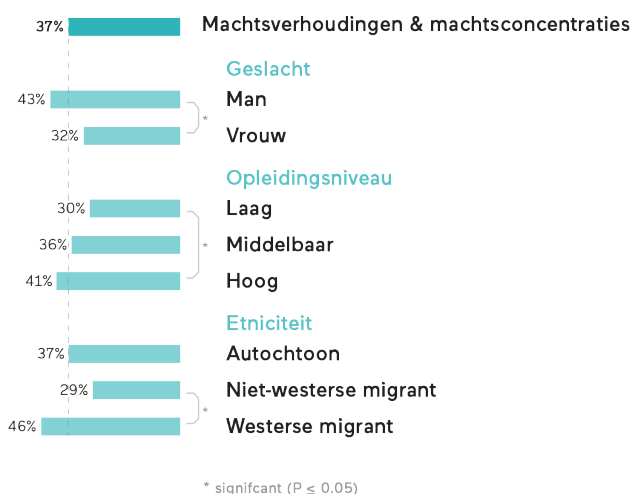
Machtsverhoudingen en machtsconcentraties

Het daaropvolgende thema wordt door 37% van de respondenten geselecteerd.

Hier zien we een significant verschil terug in gender: mannen selecteren dit thema vaker (43%) dan vrouwen (32%).

Voor leeftijdsverschil selecteren respondenten boven de 49 dit thema vaker dan jongeren. Voor 50 tot 64-jarigen is het verschil met 25 tot 34-jarigen significant: 41% van de 50 tot 64-jarigen heeft het thema geselecteerd en 30% van de 25 tot 34-jarigen. Wat betreft opleiding wordt het thema onder hoogopgeleiden vaker gekozen (41%) dan onder laagopgeleiden (30%).

Gekeken naar regio is er geen verschil terug te zien: mensen door heel Nederland kiezen dit thema even vaak. Onder Westerse migranten wordt het thema wel vaker geselecteerd dan onder niet-westerse migranten.

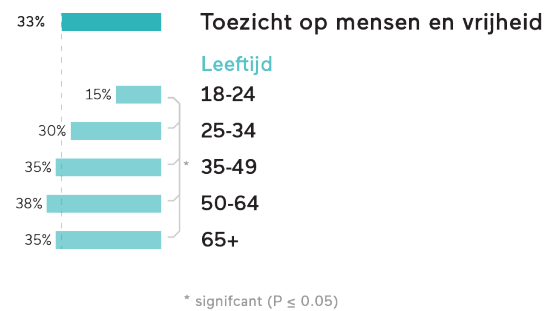


Figuur 10. Machtsverhoudingen en machtsconcentraties; significante verschillen in gender, opleiding en etniciteit

Uit de antwoorden van de respondenten komt naar voren dat mensen zich zorgen maken over gebrek aan transparantie en controleerbaarheid van AI. AI-experts (ontwikkelaars, wetenschappers en beleid) bepalen de richting van de ontwikkelingen van AI. Dit maakt de samenleving kwetsbaar, want zij heeft daarop zelf weinig tot geen invloed en is in die zin afhankelijk van experts. Mensen zien dat maatschappelijk belang in veel gevallen niet bovenaan de prioriteitenlijst staat en zijn bang dat er misbruik van ze kan worden gemaakt of dat ze worden misleid. Overigens wordt er ook met ontzag naar de ontwikkelaars van de technologie gekeken. Respondenten spreken in hun antwoorden op de open vragen hun bewondering uit voor de makers van AI en de mogelijkheden die AI biedt. Maar dat ontzag voor de makers gaat in veel gevallen gepaard met een zorg: het feit dat AI overal is, ook als je het als burger niet weet, en er weinig transparantie is over hoe het functioneert, baart mensen zorgen. Ook hebben ze het gevoel er zelf geen controle over te hebben.

Toezicht op mensen en vrijheid

Het thema “toezicht op mensen en vrijheid” is door 33% van de respondenten geselecteerd. Er is geen significant verschil als we kijken naar categorieën als gender, opleidingsniveau, regio of etniciteit. Opvallend is wel dat onder jongeren tussen de 28 en 24 jaar dit thema minder vaak wordt geselecteerd dan door alle overige leeftijdscategorieën. Jongeren gebruiken over het algemeen meer sociale media dan ouderen en zijn zich misschien minder bewust van de data die ze daarvoor “inleveren”. Ze overzien de gevolgen niet of denken mogelijk minder na over wat anderen met hun data kunnen doen en wat de risico’s daarvan kunnen zijn.

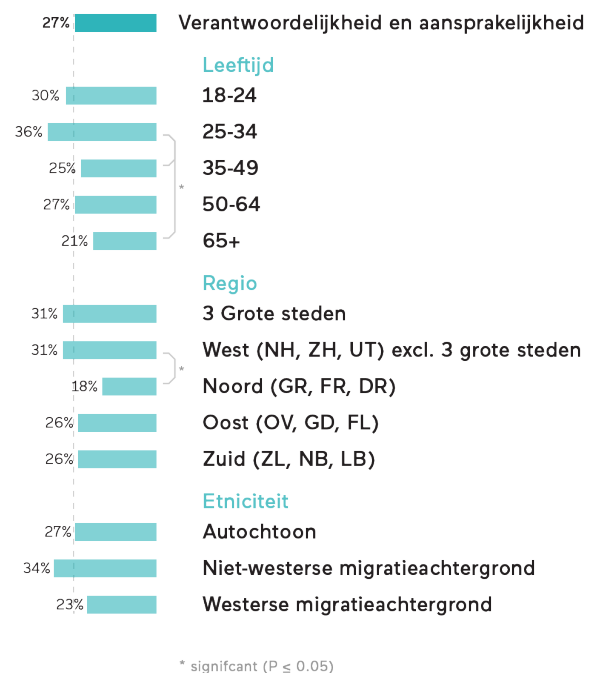


Figuur 11. Toezicht op mensen en vrijheid. Significante verschillen tussen antwoorden van mensen van verschillende leeftijden

In de antwoorden op de open vragen komen over het algemeen veel reflecties over surveillance en privacy terug. Mensen maken zich zorgen over hun data: ze weten niet wie hier toegang tot heeft en waar het voor gebruikt wordt. Het idee dat er continu (online) toezicht plaatsvindt, door overheden of bedrijven, maakt dat mensen zich beperkt voelen in hun vrijheid. Zelfs in de eigen woonkamer, bijvoorbeeld via slimme apparaten. Mensen uiten hun zorgen over het gevoel geen keuzevrijheid meer te hebben, omdat ze niet weten in hoeverre keuzes gestuurd worden door AI-technologie. Ditzelfde geldt voor bewegingsvrijheid of vrijheid van meningsuiting. Dat een derde van de Nederlandse bevolking dit graag op de onderzoeksagenda wil plaatsen komt daarom niet uit de lucht vallen.

Verantwoordelijkheid en aansprakelijkheid

Ook voor het thema “verantwoordelijkheid en aansprakelijkheid” is er geen significant verschil te vinden wat betreft gender en opleidingsniveau. Respondenten tussen de 25 en 34 jaar selecteren het thema vaker dan mensen tussen de 25 en 49 of boven 65. Gekeken naar de regio waarvan respondenten afkomstig zijn, zien we dat het thema in het Westen (exclusief de drie grote steden) vaker wordt gekozen dan in Groningen, Friesland of Drenthe. Ook wordt het thema door mensen met een westerse migratieachtergrond vaker geselecteerd dan door mensen met een niet-westerse migratieachtergrond.

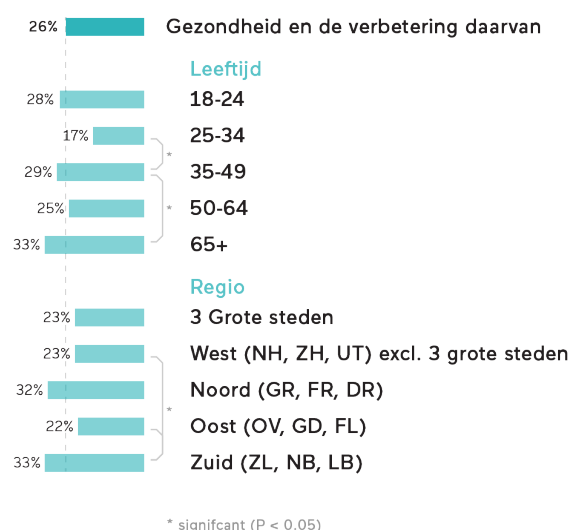


Figuur 12. Verantwoordelijkheid en aansprakelijkheid. Significante verschillen tussen mensen van verschillende leeftijd, regio en etniciteit

Verantwoordelijkheid en aansprakelijkheid gaat over de vraag wie iets te verwijten valt als er iets misgaat, of als er schade is. Bij AI is dat soms lastig, want hoewel AI door mensen is gemaakt, is het de bedoeling dat AI min of meer onafhankelijk functioneert, zonder dat mensen er de hele tijd naast zitten om te kijken of het goed gaat. Dit levert vragen op over wie er dan verantwoordelijk kan worden gehouden (en aansprakelijk). Uit de open vragen blijkt dat respondenten denken aan tal van mogelijkheden waarbij schade kan optreden (waaronder diefstal, fraude, plagiaat etc.) waarbij het nodig is te kunnen vaststellen wie daarvoor verantwoordelijk is. Het speelt ook een rol bij toepassingen waarbij schade mogelijk fataal kan zijn zoals bij medische toepassingen bijvoorbeeld. Respondenten leggen die link overigens niet expliciet, want meer dan drie kwart vindt het (heel) belangrijk dat medische AI-technologie beschikbaar is en verder ontwikkeld wordt en niemand noemt verantwoordelijkheid en aansprakelijkheid in relatie tot medische toepassingen in de open vragen. Een verklaring hiervoor zou kunnen zijn dat het vertrouwen in de medische sector zelf al vrij hoog is en daarom ook hoog is als het gaat om nieuwe technologie dat daarvoor wordt ingezet.

Gezondheid en de verbetering daarvan

26% van de respondenten vindt dat wetenschappers rekening moeten houden met “gezondheid en de verbetering daarvan”. Er is geen significant verschil voor gender, of opleidingsniveau. Qua leeftijdscategorie zien we dat het thema vaker wordt geselecteerd onder 35 tot 49-jarigen en 65+’ers dan onder 25 tot 34-jarigen. Respondenten afkomstig uit het Zuid-Nederland selecteren het thema vaker dan respondenten uit West of Noord. Opvallend is dat bij het voorleggen van voorbeelden van AI, 81% van de respondenten aangaf het (heel) belangrijk te vinden dat medische toepassingen van AI beschikbaar zijn of verder ontwikkeld worden. Dat percentage was voor geen van de andere toepassingen zo hoog. Maar als vervolgens de respondenten wordt gevraagd welke thema’s prioriteit zouden moeten krijgen als het gaat om toekomstig onderzoek naar AI, dan komt het thema gezondheid niet bovenaan te staan, maar op plek vijf. Dat is opvallend. Blijkbaar overheersen de zorgen over AI bij de keuze van respondenten voor de 12 thema’s en niet de mogelijkheden die zij in AI-ontwikkelingen zien.



Figuur 13. Gezondheid en de verbetering daarvan. Significante verschillen tussen mensen van verschillende leeftijden en woonplaats

Onderwijs, ontwikkeling en vaardigheden

21% van de respondenten heeft het thema “onderwijs, ontwikkeling en vaardigheden” geselecteerd. Onder vrouwen wordt het thema significant vaker geselecteerd (25%) dan onder mannen (16%). Een verklaring hiervoor zou kunnen zijn dat vrouwen vaker in het onderwijs werkzaam zijn dan mannen en zich er daarom meer mee bezig houden. Er is geen significant verschil tussen verschillende leeftijdsgroepen terug te vinden. Datzelfde geldt voor regio en etniciteit. Middelbaar opgeleiden selecteren het thema vaker dan hoogopgeleiden.

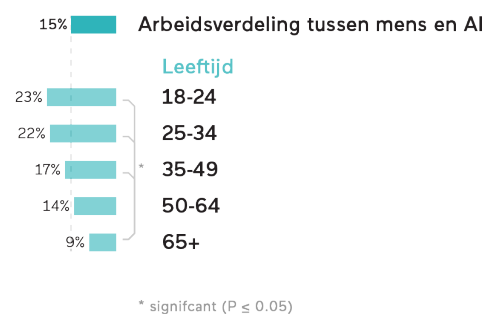


Figuur 14. Onderwijs, ontwikkeling en vaardigheden. Significante verschillen tussen gender en opleidingsniveau

Deze thematiek relateert aan een bekend debat: de invloed van AI en dan met name ChatGPT op het onderwijs. Respondenten uiten in hun reflecties hun zorgen over het verlies van vaardigheden en denkcapaciteit, als mensen (leerlingen) te veel taken delegeren aan AI, zoals verslagen en werkstukken schrijven. Ook zijn er zorgen over het verlies van de vaardigheid om echt van nep te onderscheiden: mensen zijn bang dat de vraag of iets echt waar is überhaupt nog wel gesteld blijft worden. Opvallend is wel dat 62% van de respondenten aangeeft het (heel) belangrijk te vinden dat interactieve leerplatforms die gebruik maken van AI verder worden ontwikkeld. Kennelijk zien mensen ook kansen waarbij de toepassing van AI het onderwijs zou kunnen ondersteunen. In de reflecties die mensen delen in hun antwoorden op de open vragen, zeggen ze echter weinig positiefs over de rol die zij voor AI zien in het onderwijs. Ze verwoorden daar eerder hun zorgen over de negatieve gevolgen van het gebruik van generatieve AI voor de ontwikkeling en vaardigheden van de mens. Dat 22% van de mensen dit thema op de onderzoeksagenda wil zetten is daarom lastig te duiden: enerzijds vinden mensen het een goed idee als AI het onderwijs ondersteunt, anderzijds vinden zij het geen goed idee als AI taken weghaalt bij leerlingen waardoor ze daarvan niet meer leren.

Arbeidsverdeling tussen mens en AI

Het thema “arbeidsverdeling tussen mens en AI” wordt door 15% van de respondenten geselecteerd. Er is geen significant verschil in gender, regio, opleidingsniveau en etniciteit. Opvallend is wel dat het thema vaker wordt geprioriteerd door respondenten tussen de 18 en 24, tussen de 25 en 49 en tussen de 35 en 49 vergeleken met respondenten boven de 65 jaar oud. Een verklaring hiervoor is dat mensen boven de 65 vaak reeds met pensioen zijn en zich



Figuur 15. Arbeidsverdeling tussen mens en AI. Significante verschillen tussen mensen van verschillende leeftijden

hier mogelijk minder om bekommeren omdat ze er zelf niet meer mee te maken krijgen. Werkenden krijgen dat wel en vooral voor jongeren is de kans groot dat zij in hun toekomstige loopbaan met AI zullen worden geconfronteerd.

Sturing van gedrag en beïnvloeding van keuzes

14% van de respondenten heeft het thema “sturing van gedrag en beïnvloeding van keuzes” geselecteerd. Er is geen significant verschil terug te vinden voor gender, opleidingsniveau, regio of etniciteit.

Gekeken naar de verschillende leeftijdscategorieën valt op dat het thema onder 25 tot 34 en 35 tot 49-jarigen vaker geselecteerd wordt dan onder jongeren tussen de 18 en 24 jaar. Sowieso is het thema onder die laatste leeftijdscategorie

heel weinig geselecteerd, slechts 4%. Mogelijk is dat deze subgroep zich nog niet goed bewust is van de sturende werking die AI kan hebben. Andere verklaringen zouden kunnen zijn dat ze zich er wel bewust van zijn maar het ze simpelweg niet interesseert, of dat ze het zijn gewend en het accepteren als een gegeven.

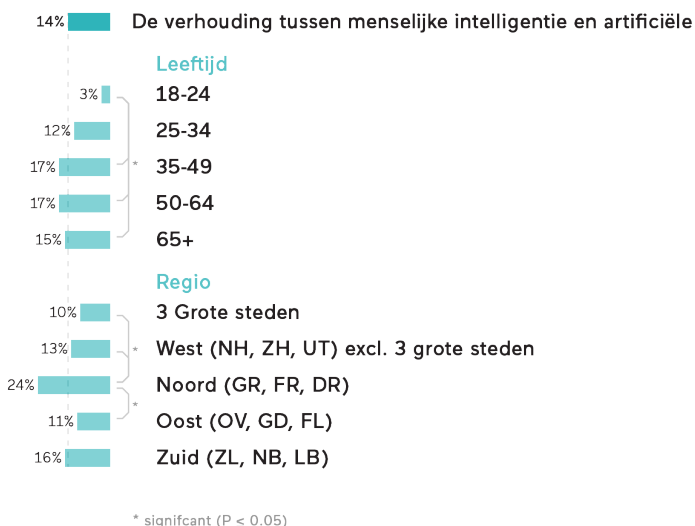


Figuur 16. Sturing van gedrag en beïnvloeding van keuzes. Significante verschillen tussen mensen van verschillende leeftijden

De verhouding tussen menselijke intelligentie en artificiële

Hoewel veel mensen in hun antwoorden op de open vragen omschreven wat AI was door termen te gebruiken die verwijzen naar menselijke capaciteiten, blijkt de verhouding tussen menselijke en artificiële intelligentie niet iets te zijn dat veel mensen onderzocht willen hebben. Dit thema wordt door 14% van de respondenten geselecteerd. Er is geen significant verschil terug te vinden in gender, opleidingsniveau en etniciteit. Opvallend is dat dit thema significant vaker wordt geselecteerd onder respondenten tussen de 35 en 49, 50 en 64 en boven de 64

vergeleken met jongeren onder de 25. Onder die groep wordt het thema sowieso heel weinig geselecteerd, namelijk slechts door 3%. Gekeken naar de regio waar respondenten vandaan komen, zien we dat dit thema vaker geselecteerd wordt door respondenten uit Groningen, Friesland of Drenthe dan onder mensen uit de drie grote steden, West en Oost.



Figuur 17. De verhouding tussen menselijke en artificiële intelligentie. Significante verschillen tussen mensen van verschillende leeftijden en woonplaatsen

Relaties en psychologische ontwikkeling

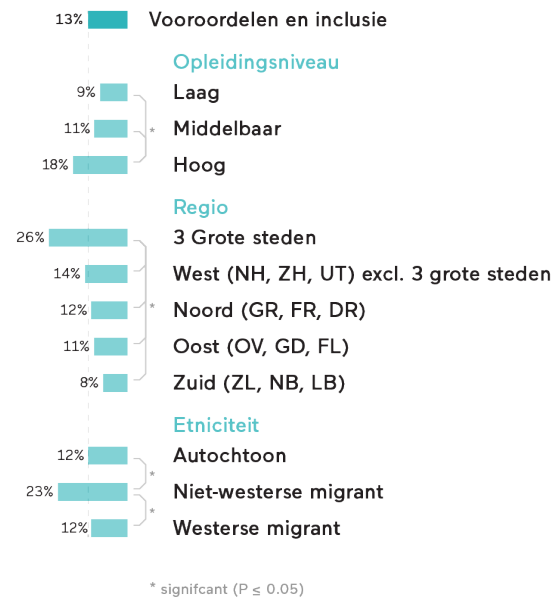
Het thema “relaties en psychologische ontwikkeling” wordt door 14% van de respondenten geselecteerd. Er zijn geen significante verschillen terug te vinden tussen subgroepen. Dat betekent dat dit thema ongeacht subcategorie niet heel vaak wordt geselecteerd. Dit is opvallend omdat dit thema wel veelvuldig ter sprake kwam tijdens de burgerpanels aan de hand waarvan we de themalijst hebben samengesteld (zie hierboven het hoofdstuk over Achtergrond). Mensen reflecteerden op voorbeelden als de AI-vriend of een relatie met een chatbot. Het ging erover wat het betekent voor een mens als hij/zij een relatie heeft met een ander die niet echt bestaat. Er waren zorgen over wat dat doet met de psychologische ontwikkeling van de mens. In antwoorden van de respondenten op de open vragen in de huidige vragenlijst kwamen we nauwelijks gedachten tegen over wat AI doet met de menselijke ontwikkeling, behalve dan over wat het doet met de ontwikkeling van menselijke intelligentie als steeds meer denkwerk wordt overgelaten aan AI. Kennelijk hebben mensen vooralsnog weinig affectie voor AI en dichten zij dit geen grote rol toe in de ontwikkeling van mensen, behalve dan als het gaat om cognitieve ontwikkeling.

Vervuiling en milieu

Het thema “vervuiling en milieu” wordt door 13% van de respondenten geselecteerd. Dat is weinig, vergeleken met de overige thema’s. Opnieuw is het opvallend, net als het hiervoor beschreven thema “relaties en psychologische ontwikkeling” dat er geen significante verschillen zijn onder de antwoorden van verschillende categorieën respondenten. Dat betekent dat onder de gehele Nederlandse bevolking dit thema niet vaak als belangrijk wordt beschouwd vergeleken met de andere thema’s. Dat is opvallend omdat milieu een belangrijk thema is in de media, en ook vaak in verband wordt gebracht met AI. Enerzijds wordt gezegd dat AI kan helpen om natuurrampen te voorspellen, waardoor mensen beter zijn voorbereid, anderzijds wordt in de media aandacht gegeven aan de ecologische voetafdruk van AI zelf. Het opslaan, verwerken en delen van data kost veel energie op de manier waarop het nu gebeurt. Dit geldt ook voor AI: de enorme hoeveelheden data en het trainen van algoritmen vraagt veel opslagruimte en rekenvermogen, en dus energie. Hoewel dit thema genoemd wordt in de media, lijkt dit onder de Nederlandse bevolking (nog) niet zozeer te spelen als het gaat om de ontwikkelingen rondom AI.

Vooroordelen en inclusie

Van de respondenten heeft 13% het thema “vooroordelen en inclusie” geselecteerd. Opnieuw is dat relatief weinig, vergeleken met de keuze voor de overige thema’s. Wat betreft gender en leeftijd is er geen significant verschil tussen de verschillende subgroepen in hoeverre ze dit thema selecteren. Gekeken naar opleidingsniveau valt op dat onder hoogopgeleiden het thema significant vaker gekozen wordt vergeleken met laag- en middelbaar opgeleiden. Ook wordt het thema in de drie grote steden vaker geselecteerd vergeleken met de rest van Nederland. Gekeken naar etniciteit valt op dat niet-westerse migranten het thema vaker aandragen ten opzichte van mensen zonder migratieachtergrond. Een mogelijke verklaring hiervoor is dat dit een groep is die zelf vaker te maken krijgt met de nadelige gevolgen van vooroordelen. Denk bijvoorbeeld aan racisme dat in historische data voorkomt en ervoor zorgt dat beslissingen van een AI aan datzelfde patroon voldoen.



Figuur 18. Vooroordelen en inclusie.
Significante verschillen voor opleiding,
woonplaats en etniciteit

Conclusie en reflectie

Op basis van de bevindingen kunnen we meerdere conclusies trekken. Allereerst kunnen we concluderen dat **de Nederlandse bevolking over het algemeen van mening is op de hoogte te zijn van ontwikkelingen op het gebied van AI**. Van 16 van de 20 toepassingen van AI geeft het overgrote deel aan er bekend mee te zijn en mensen hebben een betekenisvolle associatie met AI als ze gevraagd wordt naar hun gedachten bij het woord. Slechts een kleine groep geeft aan geen idee te hebben wat AI is. Dit verschilt van de resultaten van eerder onderzoek door Motivaction uit 2021 naar de beleving van digitalisering onder Nederlandse burgers. Daar gaf één op de vier mensen aan niet te weten wat AI is. Dit verschil is mogelijk te verklaren door de opkomst van ChatGPT, de aandacht die generatieve AI krijgt in de media en het feit dat er steeds meer toepassingen van AI in het dagelijks leven gebruikt worden.

Als het gaat om thema's die de Nederlandse bevolking op de agenda wil zetten bij wetenschappers en ontwikkelaars van AI, dan valt op dat **zorgen over AI leidend zijn bij hun keuzes**. Mensen geven veelvuldig uiting aan hun zorgen en zelfs wanneer zij een positieve verwachting of mogelijkheid van AI uitspreken, wordt dit vaak direct gevolgd door opmerking over wat er mis kan gaan. Dat wil niet zeggen dat respondenten niet in staat zijn de mogelijkheden te zien. Er is bewondering voor 'wat de technologie allemaal kan' en er wordt onzorgzaam geuit 'voor de slimme mensen die dat kunnen ontwikkelen'. Daarnaast geven zij bij meerdere toepassingen van AI aan het belangrijk te vinden dat deze beschikbaar zijn of verder ontwikkeld worden. Maar ondanks dat maar liefst 81% van de inwoners van Nederland medische AI-toepassingen (heel) belangrijk vinden, waar dat percentage voor geen van de andere toepassingen zo hoog is, eindigt het thema gezondheid niet bovenaan bij de uiteindelijke prioritering van de thema's voor de onderzoeksagenda. In plaats daarvan geven de top drie geprioriteerde thema's vooral blijk van zorgen over AI. Blijkbaar vindt de Nederlandse bevolking dat deze zorgen de nodige aandacht behoeven.

Kijken naar de thema's blijkt **dat 58% van de bevolking het thema "Nepnieuws, nepfoto's en polarisatie" op de onderzoeksagenda wil plaatsen**. Dit resultaat is consistent met de antwoorden op de open vragen. Daar worden nepnieuws, nepbeelden en *deep fakes* ook al veel genoemd. Uit de antwoorden blijkt dat inwoners zich er zorgen over maken dat zij niet langer in staat zijn om echt van nep te onderscheiden en daardoor de greep verliezen op hun eigen levens. Ook wordt het ontwrichtende effect van nepnieuws en nepbeelden op de samenleving genoemd, maar komt het begrip polarisatie eigenlijk niet terug. Hieruit leiden we af dat de zorg voor polarisatie mogelijk niet de reden was om dit thema de hoogste prioriteit toe te kennen. In plaats daarvan wortelt de reden vooral een gevoel van controleverlies en een besef dat de gewone mens, de non-expert op gebied van AI, makkelijk slachtoffer kunnen worden van misleiding.

Het thema **“Machtsverhoudingen en machtsconcentraties”** wordt als tweede geprioriteerd, waarbij 37% van de Nederlandse bevolking dit op de onderzoeksagenda wilt plaatsen. Mensen spreken hun zorg uit dat experts de macht hebben over de gewone mens zonder expertise in AI omdat zij bepalen welke doelen AI dienen. Mensen verwachten dat deze doelen niet per se in het belang van de samenleving zijn opgesteld, maar in het belang van de experts. Dit kunnen wetenschappers en ontwikkelaars zijn, werkzaam bij kennisinstellingen of grote en kleine techbedrijven, maar ook overheden (in China, Rusland of zelfs Nederland) die opdracht geven tot het maken van AI-technologie voor eigen belang en daarbij het belang van de samenleving ondermijnen.

Het thema **“Toezicht op mensen en vrijheid”** wordt door 33% van de Nederlandse bevolking op de onderzoeksagenda geplaatst en eindigt daarmee op de derde plek. Mensen uiten hun zorgen over het inleveren van hun data en het verlies van controle over waar de data naartoe gaat. Ook benoemden ze de mate van controle die (ontwikkelaars van) AI-technologie op mensen kunnen uitoefenen doordat ze deze data bezitten; ze kunnen mensen misleiden of hun keuzes beïnvloeden. Ook kunnen ze mensen in de gaten houden. In sommige landen is er al sprake van stevige surveillance en mensen willen niet dat (buitenlandse) overheden of techbedrijven hun bewegingen en handelen in het dagelijks leven kunnen monitoren. Privacy en vrijheid willen inwoners van Nederland gewaarborgd hebben.

Discussie

De samenleving wordt zelden systematisch betrokken bij het tot stand brengen van onderzoeksagenda's. Voor zover de auteurs van dit rapport konden nagaan, zijn er nog geen maatschappelijke onderzoeksagenda's opgesteld rond AI door leden van de samenleving zelf. Wel zijn er gerelateerde initiatieven, die meer zijn gericht op het betrekken van burgers bij beleid rond AI. Zo is er een burgerpanel (met 60 burgerleden) actief sinds Maart 2024 waarmee inzichten worden verzameld over de richting die de EU moet inslaan rond de verdere ontwikkeling van AI.⁶ Ook is het Rathenau Instituut bezig met de ontwikkeling van burgerpanels rond AI, die vooral beleidsinvloed voor burgers tot doel hebben.

Nieuw aan dit onderzoek is dus vooral de poging om de samenleving een stem te geven als het gaat om de verdere ontwikkeling van AI. De onderwerpen zelf, waarmee leden van de samenleving komen zijn meestal al wel bekend en beschreven in de literatuur. Data, privacy en surveillance zijn veel bediscussieerd thema's. Ook is al eerder beschreven hoe AI veranderingen teweeg kan brengen in de manier waarop we werken (Kool et al 2015; Pekkeriet et al 2020⁷), het versnellen en efficiënter maken van

⁶ <https://belgian-presidency.consilium.europa.eu/nl/evenementen/burgerpanel-beeu-1/>

⁷ Kool, L., Van Est, R.: Kansen en bedreigingen: negen perspectieven op werken in de robotsamenleving. De Robot de Baas 1, 49–67 (2015); Pekkeriet, E., Splinter, G.: Arbeid in de toekomst: Inzicht in arbeid en goed werkgeverschap in de tuinbouw (No 2020–002). Wageningen Economic Research (2020)

bedrijfsprocessen (Roßmann et al., 2018⁸), het verbeteren van de gezondheidszorg (Bertolaso et al 2019⁹), vervoer en mobiliteit (Sparrow et al 2017; Abduljabbar et al., 2019¹⁰), het voorspellen en managen van milieurampen (Fotovatikhah et al., 2018¹¹), het efficiënter maken en minder milieu belastend van landbouw en veeteelt (Ryan et al 2022; Van der Burg et al 2022)¹² en voor het slimmer, efficiënter en groener maken van steden (Allam and Dhunny, 2019¹³). Ook zijn er veel vragen over het zelfstandiger en autonomer worden van robots en de relatie tot mensen, vooral over de vraag hoe verantwoordelijkheid en aansprakelijkheid geregeld moet worden (Mavroforou et al. 2010¹⁴) of AI de intelligentie van mensen evenaart en of AI ook een morele actor kan zijn (Arkin et al. 2011; Arnold et al 2016¹⁵) en over de aard en kwaliteit van relaties tussen mens-en robots (Briggs et al 2014¹⁶). In de meer recente literatuur komen ook nepnieuws en nepbeelden naar voren, soms als probleem dat samenhangt met de verspreiding van nieuws via online social networks met behulp van AI (Aïmeur et al 2023), soms ook wordt AI opgevoerd als methode om fake news juist mee te detecteren en tegen te gaan (Iqbal et al 2023; Arthira et al. 2023). Machtsverhoudingen zijn natuurlijk ook een belangrijke thematiek in de literatuur (Luitse et al 2024; Ryan 2019¹⁷).

⁸ Roßmann, B., Canzaniello, A., von der Gracht, H., & Hartmann, E. (2018). The future and social impact of Big Data Analytics in Supply Chain Management: Results from a Delphi study. *Technological Forecasting and Social Change*, 130, 135-149.

⁹ Bertolaso, M., Rocchi, M.: Specifically human: human work and care in the age of machines. *Bus. Ethics: European Rev.* (2019)

¹⁰ Sparrow, R., Howard, M.: When human beings are like drunk robots: driverless vehicles, ethics, and the future of transport. *Trans Res Part C: Emerging Technol* 80, 206–215 (2017); Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of artificial intelligence in transport: An overview. *Sustainability*, 11(1), 189.

¹¹ Fotovatikhah, F., Herrera, M., Shamshirband, S., Chau, K. W., Faizollahzadeh Ardabili, S., & Piran, M. J. (2018). Survey of computational intelligence as basis to big flood management: Challenges, research directions and future work. *Engineering Applications of Computational Fluid Mechanics*, 12(1), 411-437.

¹² Ryan, M., van der Burg, S., & Bogaardt, M. J. (2022). Identifying key ethical debates for autonomous robots in agri-food: a research agenda. *AI and Ethics*, 1-15; van der Burg, S., Giesbers, E., Bogaardt, M. J., Ouweltjes, W., & Lokhorst, K. (2022). Ethical aspects of AI robots for agri-food; a relational approach based on four case studies. *AI & SOCIETY*, 1-15.

¹³ Allam, Z., & Dhunny, Z. A. (2019). On big data, artificial intelligence and smart cities. *Cities*, 89, 80-91.

¹⁴ Mavroforou, A., Michalodimitrakis, E., Hatzitheofilou, C., Giannoukas, A.: Legal and ethical issues in robotic surgery. *Int. Angiol.* 29(1), 75 (2010)

¹⁵ Arkin, R.C., Ulam, P., Wagner, A.R.: Moral decision making in autonomous systems: enforcement, moral emotions, dignity, trust, and deception. *Proc. IEEE* 100(3), 571–589 (2011); Arnold, T., Scheutz, M.: Against the moral Turing test: accountable design and the moral reasoning of autonomous systems. *Ethics Inf. Technol.* 18(2), 103–115 (2016)

¹⁶ Briggs, G., Scheutz, M.: How robots can affect human behavior: Investigating the effects of robotic displays of protest and distress. *Int. J. Soc. Robot.* 6(3), 343–355 (2014)

¹⁷ Luitse, D., Blanke, T., & Poell, T. (2024). AI competitions as infrastructures of power in medical imaging. *Information, Communication & Society*, 1-22; Ryan, M. (2019). Ethics of using AI and big data in agriculture: The case of a large agriculture multinational. *The ORBIT Journal*, 2(2), 1-27.

Leden van de Nederlandse samenleving die meededen met de enquête komen dus niet met een totaal nieuw inhoudelijk inzicht in de maatschappelijke impact van AI, die ze op de onderzoeksagenda willen zetten. Het is vooral bijzonder dat de mensen die meededen ongeveer alle bekende thema's wel noemen, en dus gezamenlijk een behoorlijk compleet beeld verschaffen van wat er maatschappelijk gezien toe doen bij AI. Voorts is de prioritering van thema's die zij geven interessant. Het onderzoek brengt een hiërarchie aan in de onderwerpen, waardoor je kunt zien waaraan de Nederlandse samenleving anno 2024 belang hecht. De thema's die in dit onderzoek uit 2024 naar voren komen in relatie tot AI, suggereren een verschuiving in die hiërarchie ten opzichte van de thema's die in eerder onderzoek naar de beleving van digitalisering onder de Nederlandse bevolking belangrijk werd gevonden door Motivaction (uit 2021). In 2021 gaf 74% prioriteit aan veiligheidsrisico's, zoals hacking en fraude. Nepnieuws werd door 58% gekozen, maar figureert in het Motivaction onderzoek als onderdeel van een breder thema: 'spreiding van illegale praktijken, nepnieuws en schadelijke ideeën voor de samenleving'. Opvallend is dat wanneer er gevraagd wordt naar de bedreigingen van AI, nepnieuws in het onderzoek van Motivaction niet terugkwam. Echter, het verlies van controle werd toen al wel door 23% van de respondenten genoemd. Een mogelijke verklaring hiervoor is dat **de komst van generatieve AI de zorgen van de Nederlandse bevolking heeft doen toenemen rondom (de verspreiding van) nepnieuws, nepbeelden en misbruik**.

De drie geprioriteerde thema's uit dit onderzoek (nepnieuws, machtsverhoudingen en toezicht op vrijheid) geven gezamenlijk het beeld af dat de Nederlandse samenleving zich zorgen maakt over verlies aan grip op het alledaagse bestaan. AI is alleen begrijpelijk en controleerbaar voor mensen met AI-expertise. In veel gevallen is AI ontoegankelijk voor mensen zonder expertise. Zoals altijd wanneer er sprake is van een verschil in expertise (bijvoorbeeld bij de dokter, of de tandarts, of in het vliegtuig), bestaat er een risico van een disbalans van macht tussen experts en mensen die deze expertise niet hebben (leken), maar die wel afhankelijk zijn van de experts in bepaalde opzichten. Ongelijke kennis en expertise kan leiden tot misbruik van expert-kennis of misleiding. Het ligt voor de hand te denken dat dit ook rond AI een rol kan spelen, en dat het belangrijk is om die risico's te verminderen.

Bij AI is die machtsongelijkheid tussen experts, en de misstanden die daaruit voort kunnen vloeien, lastiger te bestrijden dan in andere contexten (zoals bv het medische domein). De experts die opdracht geven om AI te maken, die het ontwikkelen of implementeren, vinden het zelf waarschijnlijk fijn dat ze daarmee in hun dagelijks leven te maken krijgen. Maar het grotere publiek krijgt ook met die tools te maken, en heeft er niet om gevraagd. AI onderscheidt zich doordat het alomtegenwoordig is, en omdat het moeilijk is om een keus te maken omtrent of je het gebruikt of niet: het is er vaak al, voordat je dat besluit kunt nemen. Dat neemt controle weg van een individu: een keus kunnen nemen biedt controle, maar die is er niet als bijvoorbeeld een werkgever een

sollicitatie-AI gebruikt om kandidaten te selecteren, of als er een smart omgeving wordt ingevoerd in een verpleeghuis.

Daarbij kan AI in toenemende mate zelfstandig functioneren, zonder dat menselijke actoren er continu op toezien. De machtsongelijkheid tussen AI-experts en de bevolking zonder AI-kennis kan dan ook minder gemakkelijk gelijk worden getrokken. Methodes die in het verleden afdoende waren om de machtsongelijkheid te mitigeren, passen dan niet bij AI: het voldoet bijvoorbeeld niet om (zoals in het medische domein) experts te verplichten informatie te verschaffen aan de niet-expert en de niet-expert vervolgens de kans te geven om op basis van die informatie te kiezen om van AI-gebruik te maken of niet. Los van het feit dat AI er vaak al is voordat je kiest, staat er ook meestal geen expert naast die er tekst en uitleg bij verschaft en AI pas 'aanzet' op het moment dat mensen een geïnformeerd besluit nemen en zeggen: ja doe maar.

Om mensen desondanks grip te geven op AI, is het **een goed idee om hen een rol te geven in de creatie van AI**. Het maken van deze maatschappelijke onderzoeksagenda is daarvoor een goed begin. Deze agenda geeft aan waar de Nederlandse bevolking onderzoek naar wil hebben gedaan. Het geeft een startpunt om onderzoeksprojecten rond AI te vormen, die aansluiten bij de wensen en belangen van de samenleving in plaats van bij die van wetenschappers, overheden of bedrijven die gewoonlijk projecten bedenken of financieringsinstrumenten ontwerpen om die projecten mogelijk te maken. Maar de invloed van de samenleving moet ook verder reiken dan alleen het ontwikkelen van een onderzoeksagenda. Om de ongelijke machtsverhoudingen tussen AI-experts en leden van de samenleving zonder AI-expertise te kenteren zou het een goed idee zijn **om de samenleving tot opdrachtgever van AI-projecten te maken**. Energiegemeenschappen zouden bijvoorbeeld hun eigen AI-tool kunnen laten ontwikkelen, die de goedkope scenario's ontwikkelt voor energieverbruik met de minste CO2 uitstoot. Inwoners van steden zouden zelf data kunnen verzamelen omtrent de kwaliteit van lucht, water en bodem, en op basis daarvan bepalen wat hun buurt nodig heeft en feedback geven aan het stadsbestuur. Textielrecyclingbedrijven zouden een AI-tool kunnen laten maken die materiaal beoordeelt op kwaliteit en selecteert om te worden verzonden naar de juiste verwerkers. Voedselproducenten, verwerkers en consumenten zouden efficiënter kunnen samenwerken in de korte keten met hulp van AI-tools die de kwaliteit van voedsel meten en berichten verzenden wanneer ze kunnen worden opgehaald en verwerkt, waardoor voedsel minder snel verloren gaat en geconsumeerd wordt wanneer het rijp is. Groepen in de samenleving kunnen, kortom, opdracht geven voor de ontwikkeling van AI die een belang van de samenleving dient, in plaats van dit allemaal over te laten aan experts.

Daarnaast zijn er kaders nodig die zorgen dat AI zo transparant en uitlegbaar mogelijk is, waardoor het verschil in expertise wordt tegengegaan. Vertrouwen heeft alleen kans te ontwikkelen in een ingewikkelde technologie zoals AI, als deze basale democratische waarden respecteert en toegankelijk is en doceerbaar voor iedereen. Onderwijs, informatie, zelf experimenteren, en meedoen aan het opzetten en co-creëren van AI,

draagt bij aan een vermeerdering van vertrouwen. Daarnaast moet AI controleerbaar worden en uitlegbaar aan de hand van publieke waarden. Het moet duidelijk zijn dat publieke waarden die de democratie waarborgen gerespecteerd moeten worden in ontwerp en uitvoering van AI, in plaats van alleen de private belangen van een selecte groep AI-experts. Die regels moeten er zijn, en de overheid moet erop toezien dat ze worden nageleefd. Ook moeten mensen de kans krijgen om misstanden die zij zien of vermoeden te melden, waarna actie volgt. Alleen op die manier is er een kans dat ook deze macht controleerbaar wordt in een democratie.

Beperkingen

Dit onderzoek heeft ook enkele beperkingen. Het onderzoek is in twee fases uitgevoerd, waarbij burgerpanels en een online enquête hebben geleid tot een lijst thema's. Vervolgens hebben we een representatieve selectie van de Nederlandse bevolking gevraagd naar hun voorkeuren voor de thema's. De methode van een enquête geeft kans om een brede representatieve groep mensen uit Nederland om hun prioriteiten te vragen. Maar de resultaten van een enquête bieden ook slechts een momentopname en laten niet zien hoe inzichten en opinies van mensen (kunnen) ontwikkelen. Andere methodes kunnen dit wel bewerkstelligen, zoals een narratieve scenario-studie, een *backcasting* experiment of een burger beraad. Dit soort methodieken maken het mogelijk om deelnemers informatie te geven en met behulp van verschillende methodes kans te geven om hun denken over een toekomst met AI te verdiepen, verbreden en verrijken. Met dergelijke methodieken waarbij deelnemers zich langduriger engageren, komen ze doorgaans tot een meer diepgaande reflectie over het onderwerp. Er is dan kans dat ze in dialoog met elkaar komen tot ideeën over een meer aantrekkelijke toekomst met AI.

Aanbevelingen

De uitkomsten van dit onderzoek vragen om een reactie van iedereen die een rol heeft in de ontwikkeling van AI in Nederland. Vooral is het nodig om een passende respons te bieden op de zorgen die ten grondslag liggen aan de keuzes van inwoners van Nederland. Wij doen daarvoor de volgende aanbevelingen:

1. Onderzoek en innovatie rond AI stimuleren die de onderwerpen op de onderzoeksagenda van de samenleving adresseert; om te beginnen moet er concrete actie worden ondernomen om nepnieuws en nepbeelden tegen te gaan; door te investeren in strategische projecten én open calls waarbij ook partijen uit de samenleving zelf kunnen deelnemen.
2. De machtsverhoudingen kantelen door projecten en calls te ontwikkelen die een maatschappelijk doel dienen en waarbij de samenleving de opdrachtgever is. Niet alleen de overheid, wetenschap of bedrijven mogen bepalen wat voor AI we ontwikkelen, maar ook de samenleving.
3. Kaders (regels) opstellen samen met de samenleving die aangeven waaraan AI moet voldoen; daarin uitwerken wat uitlegbaarheid en transparantie van AI

minimaal betekent en wat het betekent om publieke waarden te waarborgen in ontwerp zoals vrijheid, democratie, inclusiviteit, privacy en veiligheid

4. Substantieel werk maken van vertrouwen in AI en in het maken van AI die vertrouwen verdient, bijvoorbeeld door:
 - a. Mensen een kans te geven (middels co-creatie) om invloed uit te oefenen op de ontwikkeling van AI en hun waarden te gebruiken als basis voor ontwerp
 - b. Kennis van en ervaring met AI vergroten bij de samenleving middels onderwijs en experimenten voor iedereen
 - c. Realiseren van een organisatie of meldpunt dat AI controleert en waar mensen naartoe kunnen om informatie te krijgen of problemen en/of misstanden te melden

Bijlage 1: Vragenlijst

1.	Waar denkt u aan bij het woord AI ofwel Artificiële intelligentie?			
2.	Wat vindt u interessant aan AI?			
3.	Zijn er ook dingen rondom AI waarover u zich zorgen maakt?			
4.	Hoe komt u aan uw kennis over AI? Antwoordopties: Het nieuws / Mijn werk / Sociale media / Via familie of vrienden / Televisieprogramma's / Anders, namelijk ... / Weet niet			
5.	Digitale assistent	Zelfrijdende auto	Robot in de voedingsindustrie	Zorgrobot
	Sexbot	Militaire technologie met AI	Deepfake technologie	Chatbot zoals ChatGPT
	AI voor aanbevelingen	AI voor herkennen van fraude en risico's	AI voor selectie van sollicitanten	Chatbot voor digitale klantenservice
	Sporthorloge	Medische technologie met AI	Slimme vertaalmachine	AI voor de beveiliging van computersystemen
	AI voor planning en optimalisering	Leerplatforms met AI	Gezichtsherkenning	AI voor schade detectie
	Had u voor dit onderzoek al eens gehoord van ...? Antwoordopties: Ja / Nee			
Hoe belangrijk vindt u het voor de samenleving dat ... beschikbaar is of verder ontwikkeld wordt in de toekomst? Antwoordopties: Helemaal niet belangrijk / Niet belangrijk / Niet belangrijk, niet onbelangrijk / Belangrijk / Heel belangrijk / Weet niet, geen mening				

6. Kunt u van het onderstaande lijstje drie thema's kiezen waarvan u vindt dat wetenschappers daar rekening mee moeten houden als ze in de toekomst onderzoek doen naar AI?

Machtsverhoudingen en machtsconcentraties. Een paar grote technische bedrijven zijn heel actief bezig met ontwikkeling van AI. Is het wel een goed idee als zij zoveel macht krijgen over AI?

Sturing van gedrag en beïnvloeding van keuzes. AI kan mensen helpen om gezonder te leven of keuzes te maken die beter zijn voor het milieu. Maar AI kan ook ongezond gedrag stimuleren, zoals een shop-verslaving of gokken. Welke keuzes willen we overlaten aan AI?

Toezicht op mensen en vrijheid. AI kan volgen wat mensen doen en daarover data verzamelen. Is dat wel altijd wenselijk?

Verantwoordelijkheid en aansprakelijkheid. Met AI zijn machines steeds beter in staat om zelfstandig te bewegen en keuzes te maken. Wie is er dan verantwoordelijk voor de daden van AI als er fouten gemaakt worden?

Nepnieuws, nepfoto's en polarisatie. AI kan foto's bewerken en tekst verspreiden. Maar hoe zie je dan het verschil tussen echt en nep? En hoe weet je of de mening juist is, die je op basis daarvan hebt gevormd?

Onderwijs, ontwikkeling en vaardigheden. AI kan leerlingen een lesprogramma bieden dat precies past bij hun niveau en behoeften. Maar AI kan ook taken overnemen; zoals moeilijke berekeningen of werkstukken schrijven. Hoe moet AI ingezet worden in het onderwijs?

Vervuiling en milieu. AI verbruikt zelf veel vervuilende energie, maar kan mensen ook helpen om beslissingen te nemen die beter zijn voor het milieu. Hoe kan AI een positieve bijdrage leveren aan verbetering van het milieu?

Vooroordelen en inclusie. AI is gebaseerd op data uit het verleden. Hoe kunnen we zorgen dat AI geen vooroordelen van vroeger herhaalt?

De verhouding tussen menselijke intelligentie en artificiële. Wat betekent het als we een machine 'intelligent' noemen? Hoe verschilt de intelligentie van mensen van die van machines?

Relaties en psychologische ontwikkeling. Robots en chatbots nemen soms de rol over van echte mensen, zowel op het werk als privé. Wat is het effect daarvan op de ontwikkeling van mensen als ze zich moeten verhouden tot deze bots?

Arbeidsverdeling tussen mens en AI. AI kan werk overnemen van mensen. Hoe kunnen mens en AI goed samenwerken?

Gezondheid en de verbetering daarvan. AI kan helpen om diagnoses te stellen of voorspellingen te doen over iemands gezondheid. Wil je wel alles weten over je (toekomstige) gezondheid?

Bijlage 2: Voorbeelden van AI en hun belang op vijfpuntsschaal

	Heel belangrijk	Belangrijk	Niet belangrijk, niet onbelangrijk	Niet belangrijk	Helemaal niet belangrijk	Weet niet/ geen mening
Medische technologie met AI	37%	45%	9%	2%	3%	4%
AI voor beveiliging van computersystemen	28%	48%	9%	1%	3%	10%
AI voor herkennen van schade aan wegen of bruggen	21%	54%	14%	3%	3%	5%
Robots in de voedingsindustrie	20%	51%	18%	5%	4%	3%
AI voor fraude en risico's door banken	23%	47%	16%	3%	4%	7%
Slimme vertaalmachines	15%	55%	21%	4%	2%	3%
AI voor planning en optimalisering	18%	49%	22%	3%	4%	4%
Militaire technologie met AI	21%	44%	15%	5%	7%	8%
Leerplatforms voor kinderen zoals sQula	18%	44%	18%	5%	6%	9%
Robots in de zorg	18%	38%	22%	6%	10%	6%
Sporthorloges zoals de Fitbit	8%	33%	30%	13%	12%	5%
Technologie voor gezichtsherkenning	8%	32%	31%	11%	15%	4%
Chatbots zoals ChatGPT	7%	25%	25%	18%	17%	9%
Zelfrijdende auto's	6%	24%	23%	21%	25%	2%
Chatbots voor digitale klantenservice	4%	23%	30%	17%	24%	3%
Digitale assistent als Google Home of Siri	4%	18%	33%	19%	20%	5%
AI voor de selectie van sollicitanten	3%	9%	22%	27%	30%	9%
AI voor aanbevelingen	3%	9%	30%	25%	29%	4%
Seksbots	3%	6%	20%	21%	43%	7%
Deepfake technologie	1%	7%	16%	22%	46%	8%

Colofon

Uitgave

Waag Futurelab 30 april 2024

Auteurs/redactie

Simone van der Burg, Jikke van den Ende

Vormgeving

Alain Otjens, Bouwe van der Molen

Fotografie/illustraties

Waag Futurelab

Contactgegevens

Waag Futurelab

Sint Antoniesbreestraat 69

1011 HB Amsterdam