

In dialoog: samen richting maatschappelijk verantwoorde AI

December 2022



www.amai.vlaanderen

Auteurs: Pieter Duysburgh, Carina Veeckman, Michiel Vaes, Annelies Duerinckx, Karen Verstraelen, Jef Van Laer en Alexander Torfs

Projectmanagement: Pieter Duysburgh, Annelies Duerinckx

Advies workshop methodiek: 'De Betrokken Partij'

Voor vragen of bijkomende informatie mail je naar info@amai.vlaanderen.

Deze studie werd mogelijk gemaakt door steun van het departement Economie, Wetenschap en Innovatie van de Vlaamse Overheid, in het kader van het Vlaams Actieplan AI.

Inhoudstafel

1	Inleiding.....	6
2	Wat is begeleidingsethiek?	7
3	Aanpak methodiek	9
3.1	In dialoog met burgers.....	9
3.2	Randvoorwaarden	9
3.3	Vier cases.....	12
3.4	Doem- en boemscenario's.....	14
3.5	Handelingsopties	14
3.6	De ministerspeech	15
4	Opzet van de workshops	17
4.1	Planning	17
4.2	Communicatie en promo.....	18
4.3	Verloop	18
4.4	Enkele sfeerbeelden	20
5	Resultaten	21
5.1	Overzicht aantal deelnemers.....	21
5.2	Overzicht aantal stemmingen voor randvoorwaarden.....	22
5.3	Doem- en boemscenario's.....	23
5.4	Handelingsopties	24
5.5	Ministerspeech	25
6	Aanbevelingen aanpak begeleidingsethiek	26
7	Bijlages	28
	Bijlage 1 – Risico's en voorwaarden per domein.....	28
	Bijlage 2 – Fiches randvoorwaarden	30

Figuren en tabellen

Figuur 1: Aanpak begeleidingsethiek van Professor Verbeek (Verbeek & Tijink, 2019, p.32).....	7
Figuur 2: Overzicht randvoorwaarden.	11
Figuur 3: Voorbeeld van drie fiches ‘aansprakelijkheid’, ‘autonomie’ en ‘betrouwbaarheid’. Zie bijlage 2 voor alle fiches.	12
Figuur 4: Template doem- en boemscenario’s met handelingsopties.	15
Figuur 5: Ministerspeech.....	16
Figuur 6: Promotie voor de workshops via sociale media.....	18
Figuur 7- 8: Sfeerbeelden demostand.....	20
Figuur 9-10: Sfeerbeelden workshops.	20
Tabel 1: Risico’s en voorwaarden – domein gezondheid.	11
Tabel 2: Omschrijving cases per domein.....	14
Tabel 3: Planning van workshops en lezingen.	17
Tabel 4: Agenda van de workshop	19
Tabel 5: Aantal deelnemers per locatie.....	21
Tabel 6: Aantal stemmen per randvoorwaarde – prioritering naar mate van belangrijkheid.....	22
Tabel 7: Handelingsopties – gezondheidscase.....	24

Samenvatting

Dit rapport beschrijft in detail de aanpak van begeleidingsethiek die toegepast werd binnen het amai!-project. De tweede editie van amai! wilde immers niet enkel ideeën van burgers voor AI verzamelen, maar eveneens randvoorwaarden en bezorgdheden van burgers voor AI capteren (taken T3.1 en T3.2).

Met dat doeleinde werd de aanpak van begeleidingsethiek vertaald naar een methode van dialoog met burgers over ethische dilemma's over artificiële intelligentie (AI). Deze methodiek werd opgezet in 2022 aan de hand van workshops "AI en randvoorwaarden" en lezingen. Deze activiteiten richtten zich naar het brede publiek, en voornamelijk naar het profiel van de 'maatschappelijke toepassers'.

In totaal werden er vier workshops en drie lezingen georganiseerd met 118 deelnemers tijdens de herfstperiode van 2022. Met deze methodiek beogen we positief en negatief geladen discussies op te zetten rond de technologie, randvoorwaarden te definiëren, en een inclusieve aanpak te hanteren over de rol van AI in onze maatschappij.

De (aangepaste) aanpak van begeleidingsethiek wordt uitgebreid toegelicht in dit rapport in hoofdstuk 2 en 3. De logistieke organisatie en promotie van de workshops en lezingen is omschreven in hoofdstuk 4. De resultaten met aanbevelingen over de methodiek zijn terug te vinden in hoofdstuk 5. Op basis van de resultaten is er ook een set van fiches ontworpen dat randvoorwaarden beschrijft. Dit zijn ethische richtlijnen, geformuleerd door burgers, waaraan een ontwerp wenst te voldoen. Deze randvoorwaarden zorgen ervoor dat (nieuwe) AI-toepassingen ethisch en juridisch correct zijn.

1 Inleiding

Ontwikkelingen op vlak van AI zorgen regelmatig voor onrust bij het brede publiek. Afhankelijk van de specifieke toepassing, gaat over zaken met betrekking als privacy, tewerkstelling, digitale inclusie, jobzekerheid etc. Het project amai! wil het brede publiek inzicht geven in AI en inspraak geven over de richting waarin AI-toepassingen evolueren. Daarbij wilde de tweede editie van amai! in 2022 ook aandacht hebben voor de ethische bezorgdheden van burgers bij deze AI-ontwikkelingen.

In de plannen voor amai! in 2022 stond initieel beschreven dat we samen met burgers wilden komen tot een 'adviesdocument' voor verantwoorde AI. De bedoeling was te komen tot beleidsinput die bestuurders, onderzoekers en bedrijven zouden moeten helpen om te komen tot maatschappelijk gedreven AI-innovatie. Dit laatste is altijd de focus gebleven, maar de methode om daartoe te komen is evenwel in de loop van het project bijgestuurd op basis van voortschrijdend inzicht in het projectteam van amai! en bij het Kenniscentrum Data & Maatschappij, één van de trekkers van het project.

Het idee was initieel immers dat een overzicht van randvoorwaarden zou worden gemaakt die burgers hebben bij AI-ontwikkelingen. Dit pad werd verlaten omwille van 2 redenen. 1) In voorbije jaren zijn internationaal meerdere initiatieven geweest die een overzicht hebben gemaakt van algemene randvoorwaarden voor betrouwbare AI-innovatie (e.g. de [high level expert group voor AI](#), [Unesco](#), etc.). Verschillende van deze initiatieven gingen ook actief ten rade bij burgers (e.g. [the Montreal declaration for responsible AI](#)). Het was niet duidelijk hoe het initiatief van amai! op dit vlak nog een verschil kon maken. 2) Het Kenniscentrum Data & Maatschappij is in de loop van 2022 steeds sterker vertrokken van een kader dat een contextspecifieke benadering hanteert wanneer het gaat over ethiek en (AI-)technologie, met name de 'Aanpak Begeleidingsethiek'. Dit kader bekijkt een technologie in z'n context en onderzoekt de handelingsopties die mogelijk zijn om een antwoord te bieden op ethische bezorgdheden. Dit kader vertrekt dus niet van algemene normen maar bekijkt de specifieke vragen die leven rond een specifieke technologische toepassing en neemt die als startpunt om na te denken over de manieren waarop de technologie kan worden ingezet.

Deze inzichten leidden ertoe dat we ook de aanpak die beoogd was in amai! werd geactualiseerd en aangepast. Er is voor gekozen om de 'Aanpak Begeleidingsethiek' verder te verkennen en specifiek te bekijken of en hoe burgers een stem kunnen hebben in dit proces. Hoe kunnen we burgers betrekken bij AI-innovatieprocessen en ook hen een stem geven over de ethische vragen die een toepassing oproept en de handelingsopties die hierbij mogelijk zijn. Hoe kunnen we deze aanpak inzetten om met burgers te komen maatschappelijk verantwoorde AI?

2 Wat is begeleidingsethiek?

De ‘aanpak begeleidingsethiek’ is een gedachtegoed van techniekfilosoof Professor Verbeek. Centraal in dit gedachtegoed staat de verwevenheid tussen technologie en samenleving. Technologie en de maatschappij staan daarbij in wisselwerking: technologie kan de maatschappij beïnvloeden, maar de maatschappij kan ook technologie beïnvloeden.

Bij de introductie of aanpassing van een technologie kunnen er wel eens kritische stemmen klinken. Is het een goede of eerder een nefaste ontwikkeling? Vinden we de technologie ethisch correct? Welke waarden vinden we belangrijk? Binnen begeleidingsethiek wordt (nieuwe) technologie niet zomaar met een ‘ja – dit willen we’ of ‘neen – dit willen we niet’ bestempeld. De aanpak focust zich eerder op de **hoe-vraag**: hoe kan technologie op een waarden-volle wijze ontworpen, geïmplementeerd en gebruikt worden? De hoe-vraag opent ruimte voor discussie en laat toe om condities te onderzoeken waarop een technologie op een verantwoorde manier kan functioneren. Dit wil echter niet zeggen dat elke technologie een plaats krijgt in onze samenleving. Als blijkt dat de technologie niet past bij onze waarden, dan is de uitkomst dat deze beter niet gebruikt kan worden.

Naast de wisselwerking tussen technologie en maatschappij, hecht dit ethisch kader ook belang aan een **positief en negatief dialoog**. Kritische stemmen focussen enkel vaak op wat er mis is, wat er niet kan, of op de onzekerheid van de toekomst. In deze aanpak wordt er echter een balans gezocht naar zowel negatieve als positieve aspecten. Een technologie kan zowel iets positief als iets negatief zijn. Bovendien is iets wat een negatieve emotie teweegbrengt vaak een indicator voor een waarde dat we belangrijk vinden.



Figuur 1: Aanpak begeleidingsethiek van Professor Verbeek (Verbeek & Tijink, 2019, p.32).¹

De aanpak begeleidingsethiek wordt in de praktijk opgezet door drie verschillende fases (zie figuur 1). Fase 1 beschrijft daarbij de **technologie en de context** waarin die technologie functioneert in begrijpbare taal. In de tweede fase wordt er een **dialoog** opgestart met actoren. Actoren die actief betrokken zijn bij het gebruik van een technologie worden uitgenodigd, zoals burgers, ontwerpers, beleidsmakers, etc. maar dit kan ook breder opengetrokken worden naar niet-gebruikers. Tijdens een workshop sessie met actoren worden effecten en waarden verzameld rond het gebruik van een technologie. In de derde fase vindt er een oefening plaats om handelingsopties te definiëren vanuit de technologie, context en gebruiker. Deze **handelingsopties** verwijzen naar ethische waarden, richtlijnen of randvoorwaarden om een technologie goed te kunnen gebruiken in een bepaalde aangepaste omgeving, met een geïnformeerde gebruiker.

Deze aanpak van begeleidingsethiek hebben we opgezet binnen het amai!-project. AI is nog altijd een nieuwe technologie waarvan we nog niet weten wat de impact op de samenleving zal zijn. De aanpak begeleidingsethiek biedt een gestructureerde methode om over deze impact (en bijhorende handelingsopties) na te denken, en focust in eerste instantie op gesprekken met direct betrokken ‘stakeholders’.

Omdat discussies over AI ook in de brede samenleving leven, bekeken we in amai! hoe we de aanpak van begeleidingsethiek konden inzetten om met een breder publiek van burgers na te denken over AI-technologie en menselijke waarden. Tijdens een workshops gingen we met burgers in dialoog, en

¹ <https://begeleidingsethiek.nl/wp-content/uploads/2021/05/Aanpak-begeleidingsethiek-gecomprimeerd.pdf>

bespraken we ethische dilemma's over de technologie. We bespreken mogelijke positieve en negatieve effecten, prioriteerden randvoorwaarden en zochten naar handelingsopties voor de technologie, omgeving en gebruiker. De opzet van de methodiek van deze workshops kan je terugvinden in hoofdstuk 3, en de resultaten in hoofdstuk 5.

3 Aanpak methodiek

Het amai!-project koos er voor om (een licht aangepaste vorm van) de aanpak van begeleidingsethiek aan te bieden via een workshop format genaamd “**AI en randvoorwaarden**” en via lezingen.

De workshops hadden een interactieve opbouw, en de lezing was een voordracht over de basisbeginselen van AI. Tijdens de lezingen werden ook de randvoorwaarden gepresenteerd en kregen deelnemers kort de gelegenheid om hun mening te delen. In dit hoofdstuk gaan we dieper in op het interactieve format van de workshops. Voor meer informatie over de lezingen verwijzen we naar het [activiteitsaanbod](#) van amai!.

3.1 In dialoog met burgers

Het amai!-project koos ervoor om de workshops en lezingen te richten naar het **brede publiek**, en voornamelijk naar het profiel van de ‘maatschappelijke toepassers’.

Uit de [voorstudie van amai!](#) van 2021 weten we dat 26,6% van de Vlamingen nog nooit heeft gehoord van AI, en 40,3% er al van gehoord heeft maar niet weet wat het betekent. Verder gaf 48% aan niet te weten of ze al eens gebruik hadden gemaakt van AI. Gezien deze cijfers opteerde amai! er voor om zich specifiek te richten naar het profiel van de maatschappelijke toepassers. Dit profiel heeft geen of weinig kennis over AI, maar is wel heel gemotiveerd om een bijdrage te leveren als de technologie gerelateerd is aan één van de vier maatschappelijke domeinen van amai!, namelijk gezondheid, mobiliteit, werk en klimaat.

In dit opzicht is de samenstelling van de groep actoren anders dan de aanpak begeleidingsethiek van Professor Verbeek. Daar wordt aangeraden om een generieke groep aan actoren uit te nodigen, met bijvoorbeeld burgers, professionele gebruikers, beleidsmedewerkers/bestuur, academici, ontwerpers, experts, etc. Deze actoren zijn veelal gebruikers van de technologie met een rechtstreekse inbreng in de dialoog. Gezien veel burgers al gebruik maken van AI zonder dit te beseffen zijn ze geen bewuste gebruikers. Met deze workshops beoogden we bewustzijn en een inclusieve aanpak te creëren over de rol van AI in onze samenleving. Daarbij ondersteunden workshops een proces van kennismaking met de technologie en tegelijkertijd een ethisch dialoog.

3.2 Randvoorwaarden

De eerste oefening tijdens de workshops peilde naar randvoorwaarden. Randvoorwaarden zijn geformuleerde eisen en richtlijnen waaraan een ontwerp moet voldoen, en zorgen ervoor dat AI-toepassingen ethisch en juridisch correct zijn. In 2021 werd er een voortraject opgezet om randvoorwaarden te verzamelen. Het proces van verzameling van randvoorwaarden wordt uitgelegd in volgende paragrafen, alsook hoe deze verder gevalideerd werden tijdens de workshops.

- **Voorstudie naar voorwaarden en risico's**

In 2021 peilde amai! naar mogelijke risico's en voorwaarden van AI bij burgers via twee activiteiten.

Allereerst werd er bij het indienen van een idee voor een maatschappelijke toepassing met AI een optionele vraag gesteld of de indiener ethische bemerkingen had. 53 indieners vulden deze vraag in. Verder werd er tijdens de kampvuursessies in de zomer van 2021 gevraagd om risico's en randvoorwaarden voor het concept neer te schrijven. Deze workshops hadden als doel om ideeën

verder te verfijnen en uit te werken tot een concept. Meer informatie over de kampvuursessies kan je [hier](#) raadplegen.

78 bemerkingen werden verzameld tijdens de kampvuursessies. In totaal werden zo 131 inputs aan risico's en voorwaarden over AI verzameld.

De meeste input werd verzameld voor het domein gezondheid (51), gevolgd door mobiliteit (43), werk (25) en klimaat/milieu (12). Deze inputs werden samengevoegd in een Excel-lijst en thematisch gecodeerd per type risico en voorwaarde, en geclusterd per domein. Deze analyse gebeurde op een generiek niveau, waarbij er werd gekeken naar terugkerende risico's en voorwaarden los van het specifieke idee. In de meeste gevallen konden risico's worden gelinkt met voorwaarden, bijvoorbeeld het risico op 'schending van de privacy van een gebruiker' gelinkt kan worden met de voorwaarde 'conformiteit met GDPR-richtlijnen'.

Onderstaande tabel vat de risico's en voorwaarden samen voor het domein gezondheid. De resultaten van de andere domeinen zijn terug te vinden in bijlage 1.

Domein: Gezondheid	
Risico's	Voorwaarden
<p>1) Een gezondheidsrisico als gevolg van gebruik:</p> <ul style="list-style-type: none"> • Het systeem werkt niet meer • Er wordt een foute diagnose of therapie voorgesteld • Wat met hiaten of ontbrekend onderzoek? • Wat met specifieke gevoeligheden, zoals allergieën? <p>2) Een juridisch risico: wie is er aansprakelijk bij fouten</p> <p>3) Risico op bias of filter bubble</p> <p>4) Discriminatie en sociale sortering bij delen van data</p> <p>5) Risico dat de technologie niet inclusief is:</p> <ul style="list-style-type: none"> • Digitale kloof • Kost van het gebruik: sociale uitsluiting <p>6) Privacy van de gebruiker wordt niet gerespecteerd. Gezondheidsdata is gevoelige data.</p> <p>7) Risico op fraude en misbruik met de data</p> <p>8) Risico op averechts effect. Bijvoorbeeld gebrek aan capaciteiten als gevolg van veelvuldig gebruik</p>	<p>1) Correctheid van een diagnose (gelinkt aan risico 1 en 3):</p> <ul style="list-style-type: none"> • Aansluitende diagnose • Waarschijnlijkheid van de diagnose opnemen <p>2) Datakwaliteit (gelinkt aan risico 1 en 3)</p> <ul style="list-style-type: none"> • Goede training algoritme • Up-to-date data • Herkennen van duplicaten • Herkennen van synoniemen <p>3) 'Human in the loop' – controle van zorgverlener (risico 2, 8, 9 en 10)</p> <p>4) Ondersteunen van digitale vaardigheden (gelinkt aan risico 5)</p> <p>5) Conformiteit met GDPR-richtlijnen volgen (gelinkt aan risico 4, 6 en 7)</p> <p>6) Transparantie in beslissingen (gelinkt aan risico 3, 6, 7 en 8)</p> <ul style="list-style-type: none"> • Informeren van gebruiker over doelstellingen app • Informeren over welke data verzameld wordt

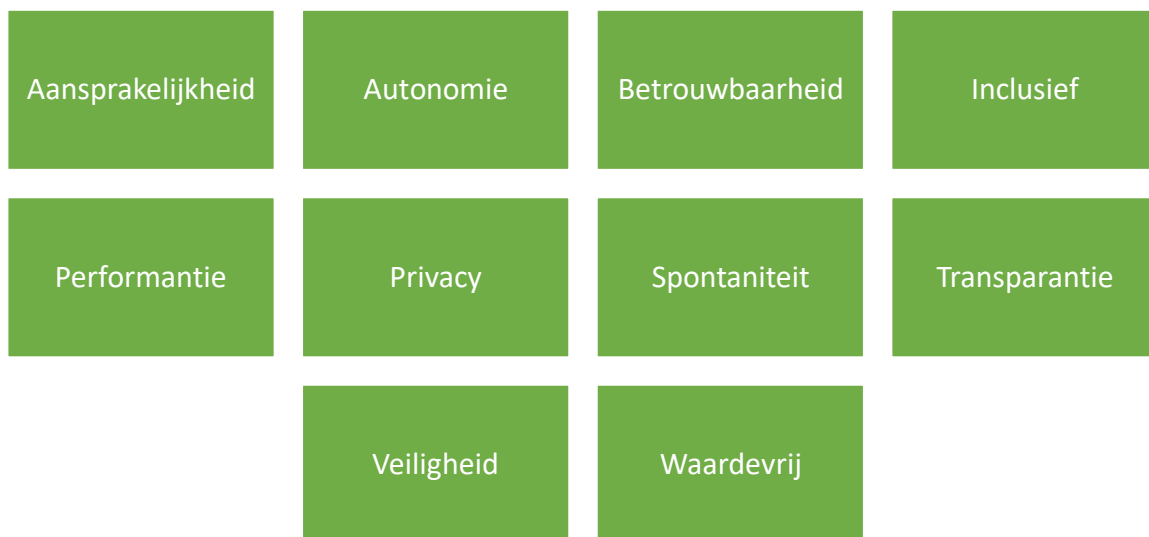
9) Risico dat de gebruiker te veel vertrouwen heeft in de technologie, en blindelings advies opvolgt	
10) Risico op verlies autonomie van de zorgverlener of arts	

Tabel 1: Risico's en voorwaarden – domein gezondheid.

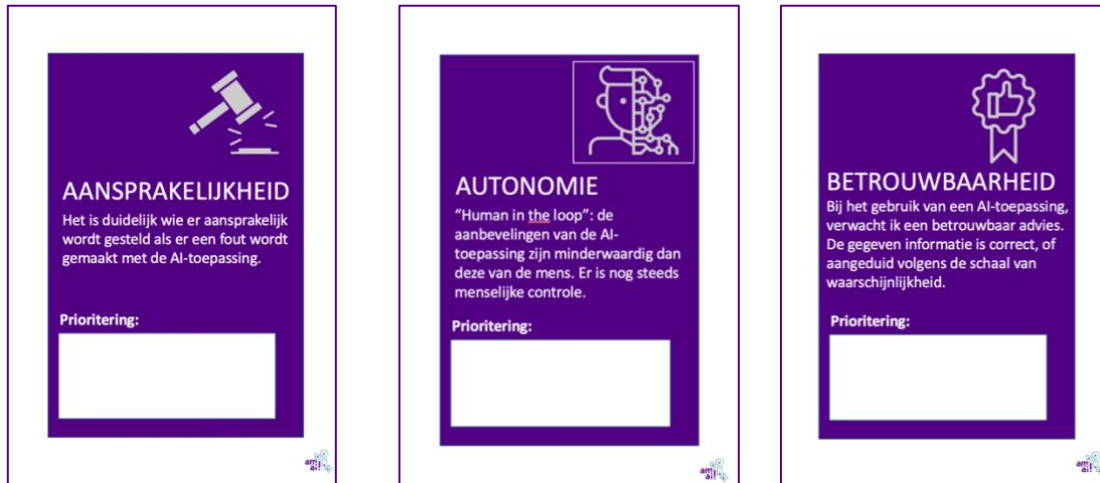
o Clustering van randvoorwaarden

Op basis van de thematische contentanalyse in de Excel-lijst werden er tien clusters van mogelijke risico's en voorwaarden geïdentificeerd. Deze werden gelabeld als **randvoorwaarden**. Randvoorwaarden zijn geformuleerde eisen en richtlijnen waaraan een ontwerp moet voldoen, en zorgen ervoor dat AI-toepassingen ethisch en juridisch correct zijn.

De volgende tien clusters werden geïdentificeerd binnen het amai!-project, waarvoor elk een fiche werd opgemaakt met korte beschrijving en een voorbeeld. De fiches kan je terugvinden in bijlage 2.



Figuur 2: Overzicht randvoorwaarden.



Figuur 3: Voorbeeld van drie fiches ‘aansprakelijkheid’, ‘autonomie’ en ‘betrouwbaarheid’. Zie bijlage 2 voor alle fiches.

○ Identificeren en prioriteren van randvoorwaarden

Tijdens de workshops “AI en randvoorwaarden” gingen we verder aan de slag met de fiches. Tijdens een eerste oefening werden de randvoorwaarden één voor één voorgesteld aan de deelnemers, met daarbij een korte uitleg. Ze kregen ook informatie over hoe de randvoorwaarden tot stand waren gekomen. Na het presenteren van de tien randvoorwaarden kregen de deelnemers elk drie stickers. De deelnemers werden gevraagd om deze te klevan bij hun drie belangrijkste randvoorwaarden. De stickers konden ze klevan in het daarvoor bestemde witte vak op de fiche.

Na de prioritering werd aan de deelnemers ook gevraagd of ze aan andere randvoorwaarden dachten, dan deze reeds omschreven. Nieuwe randvoorwaarden werden toegevoegd door middel van een jokerkaart.

Na het klevan van de stickers en aanvullen van nieuwe randvoorwaarden was er een korte discussie met de groep. De deelnemers konden hun meningen delen over de randvoorwaarden die de meeste en minste stemmen hadden gekregen. De moderator vroeg de deelnemers om eventuele voorbeelden aan te halen van situaties waarbij ze reeds dergelijke ethische dilemma’s hadden ervaren. Verdere bedenkingen en bezorgdheden konden gedeeld worden met de groep.

Na deze sessie bleven de fiches met randvoorwaarden centraal liggen op de tafel om deze verder te kunnen gebruiken tijdens de workshop.

In de reguliere ‘aanpak begeleidingsethiek’ komt deze stap met randvoorwaarden niet voor. Er is voor gekozen om dit in deze amai!-workshops wel te doen. Enerzijds was dat om de eerdere input van burgers niet verloren te laten gaan, anderzijds wilden we burgers ook voor de start van de oefening sensibiliseren rond mogelijke randvoorwaarden. Doordat zij geen direct betrokken stakeholders zijn, kon het namelijk zijn dat het voor hen moeilijker is om mogelijke maatschappelijke / ethische implicaties in te schatten. Door hen eerst inzicht te geven in randvoorwaarden die eerder werden geopperd door burgers bij andere AI-oplossingen, wilden we hun reflectieproces ondersteunen bij de cases die het onderwerp waren van deze workshops.

3.3 Vier cases

Tijdens de tweede oefening van de workshop kregen de deelnemers de keuze uit vier verschillende cases. Elke case was gelinkt aan een maatschappelijk domein van amai! en omschreef een scenario waarbij een nieuwe AI-toepassing wordt geïntroduceerd²:

Gezondheid	Mobiliteit
<p>Een internationaal team van onderzoekers ontwikkelde een AI-systeem dat louter op basis van oogscans kan inschatten wie een verhoogd risico loopt om binnen het jaar een hartaanval te krijgen. Wanneer artsen en patiënten van deze AI-oplossing gebruik willen maken, moeten ze - net zoals bij andere nieuwe technologieën – met veel rekening houden. Wanneer, waarvoor en hoe willen we de technologie gebruiken?</p> <p>Wat vind jij belangrijke voorwaarden om een AI-toepassing, zoals deze oogscan, te kunnen vertrouwen?</p>	<p>Tegen het einde van 2023 komen er 250 ‘slimme’ verkeerslichten in Vlaanderen. Op basis van o.a. smartphone-signalen zal kunnen bepaald worden welke vervoersmiddelen op dat moment voorrang krijgen. Zo kunnen bijvoorbeeld bussen met tientallen inzittenden of fietsers bij regenweer langer groen licht krijgen. De bedoeling is dat zoveel mogelijk weggebruikers via hun smartphone verbinding maken met deze verkeerslichten om op deze manier de nauwkeurigheid van het systeem te verbeteren. Een goede oplossing om efficiënt verkeersstromen aan te pakken of een probleem voor de privacy? Onder welke voorwaarden zou jij je smartphone koppelen aan deze slimme verkeerslichten?</p> <p>Wat vind jij belangrijke voorwaarden om een AI-toepassing, zoals deze slimme verkeerslichten, te kunnen vertrouwen?</p>
Werk	Klimaat & milieu
<p>Iedereen laat wel eens een pakje thuisbezorgen. Voor de transportmaatschappijen is de planning van welke chauffeurs welk pakje moet leveren op welk moment vaak een enorme uitdaging. Nieuwe AI-systemen helpen die maatschappijen om die planning zo vlot mogelijk te laten verlopen en de chauffeurs optimaal in te zetten. Deze algoritmes helpen planners om de pakjes sneller en efficiënter te leveren en dus het beperkte personeel zo veel mogelijk te laten leveren in de beschikbare tijd.</p> <p>Wat vinden we belangrijk als we deze algoritmes inzetten? Aan welke voorwaarden moet de technologie voldoen om het belang van alle betrokkenen zo goed mogelijk te bedienen?</p>	<p>Vervuiling van rivieren is een toenemend probleem. Slimme camera’s uitgerust met AI kunnen drijvend afval detecteren. Een autonoom varende werkboot kan vervolgens het afval onderscheppen en duwt het naar een collectiepontoon. In de haven hangt een signalisatiebord dat voorbijgangers wijst op de aanwezigheid van de slimme camera’s. Op deze manier vermindert het afval in het rivierwater en de zee.</p> <p>Maar aan welke voorwaarden moet deze technologie voldoen om echt een bijdrage te leveren aan het klimaat, milieu én mens?</p>

² De cases mobiliteit en gezondheid worden ook behandeld in het [AI-kompas](#).

Tabel 2: Omschrijving cases per domein.

Volgens de voorkeur van de deelnemers koos men voor één of meerdere cases en ging men in groep of in kleinere groepjes te werk. De deelnemers kregen 10 minuten de tijd om een specifieke case te kiezen en door te nemen.

3.4 Doem- en boemscenario's

De deelnemers gingen vervolgens aan de slag om doem- en boemscenario's te bedenken per case. **Doemscenario's** zijn scenario's die mogelijke negatieve effecten, problemen of fouten omschrijven van de technologie. Omgekeerd zijn **boemscenario's** manieren waarop alles succesvol of efficiënt kan verlopen.

Per scenario werd gevraagd om rekening te houden met een aantal actoren of stakeholders: welke negatieve ('doem') en positieve ('boem') effecten kunnen we voor elke actor bedenken? De actoren waren afhankelijk per gekozen domein. Bijvoorbeeld voor het domein gezondheid waren dat de patiënten, de arts en de mutualiteit.

Ter ondersteuning van de brainstorm werd ook gebruik gemaakt van de fiches met de beschrijving van de randvoorwaarden. De verschillende negatieve en positieve effecten konden op deze manier gelinkt worden met bestaande of nieuwe geïdentificeerde randvoorwaarden.

3.5 Handelingsopties

Naast het bedenken van doem- en boemscenario's werden de deelnemers ook uitgenodigd om na te denken over handelingsopties. **Handelingsopties** zijn mogelijke oplossingen met concrete acties voor de technologie (het ontwerp), omgeving (fysiek, sociaal of juridisch) en de gebruiker (gedrag, kennis of vaardigheden). Op deze manier bedacht men oplossingen om doemscenario's te vermijden of boemscenario's te realiseren.

Onderstaande afbeelding toont de template die de deelnemers invulden voor het identificeren van doem- en boemscenario's met bijhorende handelingsopties:

DOEM-scenario

"Op welke manieren kan het mis lopen met AI?"

"Welke negatieve effecten kan het hebben voor de actoren?"

"Welke waarden zit achter welk effect verborgen?"

OPLOSSINGEN & ACTIES

"Hoe zorg ik ervoor dat doemscenario's vermeden worden?"
Specificeer acties voor de technologie, omgeving en gebruiker.

BOOM-scenario

"Op welke manieren kan AI leiden tot succes en groei?"

"Welke positieve effecten kan het hebben voor de actoren?"

"Welke waarden zit achter welk effect verborgen?"

OPLOSSINGEN & ACTIES

"Hoe zorg ik ervoor dat boomsenario's werkelijkheid worden?"
Specificeer acties voor de technologie, omgeving en gebruiker.

Figuur 4: Template doem- en boemscenario's met handelingsopties.

3.6 De ministerspeech

De laatste sessie tijdens de workshop was de 'ministerspeech'. De deelnemers kregen hiervoor ongeveer tien minuten de tijd om een advies te formuleren per gesprekstafel. De speech begon met "Als ik minister van innovatie zou zijn, dan vind ik het belangrijk dat AI". Per groep stelde 1 deelnemer het advies voor.

Deze oefening had als doelstelling om adviezen te verzamelen over ethische richtlijnen voor AI in Vlaanderen. De deelnemers konden daarbij prioriteren wat ze belangrijk vinden, wat ze huidig willen veranderen of een duidelijke wens uitbrengen voor de technologie naar de toekomst toe. Met deze afsluitende oefening willen we de stem van gebruikers van AI laten horen en hen ondersteunen in een ethisch denkproces.



Advies over de toekomst van AI in Vlaanderen – op een maatschappelijk en ethische verantwoorde manier vooruit!

Groepje:
Locatie:

“Als ik minister van innovatie zou zijn, dan vind ik belangrijk dat AI ...”

Figuur 5: Ministerspeech.

4 Opzet van de workshops

In dit hoofdstuk kan je de praktische opzet van de workshops raadplegen: de planning, de communicatie met medeorganisatoren en de agenda.

Er werd ervoor gekozen om de aanpak begeleidingsethiek in twee vormen aan te bieden:

- Een **workshop** rond randvoorwaarden voor AI, zoals beschreven in hoofdstuk 3: een dialoog met burgers rond de prioritering van randvoorwaarden, boem- en doemscenario's en handelingsopties voor AI, de omgeving en gebruiker.
- Een **lezing** met een kort intermezzo over de geïdentificeerde randvoorwaarden

De workshops en lezingen werden georganiseerd in samenwerking met bibliotheken in Vlaanderen, met sociaal-culturele verenigingen of educatieve organisaties. Bij het organiseren van een workshop werd er ook een **AI-demostand** voorzien. Deze demostand liet deelnemers spelenderwijs kennis maken met AI. Het was een opwarmer voor de workshops waarbij deelnemers de basisbeginselen van AI konden ontdekken. Op de AI-demostand kom men aan volgende activiteiten deelnemen: AI of ni quiz, het AI-kompas, try-out pictionary, edubox AI, wordcloud AI-ideeën, etc. Men kon er ook kennis maken met de vier goedgekeurde amai! projecten van 2021.

4.1 Planning

Onderstaande tabel geeft de planning weer van de workshop sessies en lezingen die tijdens de herfstperiode van 2022 werden georganiseerd. In kolom 3 wordt het aantal activiteiten per locatie vermeld:

Datum	Locatie	Activiteit	AI-stand
7 september 2022	ARhus bibliotheek, Roeselare	Workshops (2)	Ja
5 november 2022	Bib Beringen	Workshop (1)	Ja
10 november 2022	Avansa Oostende – Astropolis	Lezing (1)	Neen
22 november 2022	Stadsbibliotheek Leuven	Lezingen (2)	Ja
3 december 2022	Samana vorming dag, Erasmushogeschool Brussel	Workshop (1)	Ja

Tabel 3: Planning van workshops en lezingen.

De workshops vonden zowel overdag als tijdens de avond plaats. Het begin- en eindstip werd steeds met de medeorganisator vastgelegd. Een workshop duurde twee uren, de duurtijd van een lezing was variabel (drie kwartier tot anderhalf uur).

Het aantal deelnemers per workshop werd vastgelegd op minimum 10 tot maximaal 20 personen. Na het bezoeken van de AI-demostand werden de deelnemers uitgenodigd om plaats te nemen aan een gesprekstafel. Afhankelijk van het aantal deelnemers waren er twee of drie gesprekstafels per workshop, met zes tot acht personen per tafel.

4.2 Communicatie en promo

De workshops werden gepromoot onder de noemer “**AI en randvoorwaarden**” voor het brede publiek. De promotekst had een duidelijke ethische klemtoon:

*Om 11u start een workshop over AI en randvoorwaarden. Ben jij soms ook tegelijk verbaasd en bezorgd over de razendsnelle technologische veranderingen met behulp van AI?
Zou je soms meer inzicht en controle willen in de beslissingen die op dat vlak worden genomen?
Dat kan in deze workshop! Je leert bij over de nieuwe technologie en je kan tegelijk je bezorgdheden en aandachtspunten meegeven. Jouw mening telt, want met de uitkomsten van deze workshop gaan onderzoekers aan de slag.*

Deelnemers konden zich inschrijven via een registratiemodule van de medeorganisator. Hiervoor werd een promotekst en afbeelding voorzien door het amai!-team. De link naar de registratiemodule, samen met de nodige info, werden verspreid via een blogbericht op de amai!-website. Dit bericht werd vervolgens gedeeld op onze sociale media: Twitter, Facebook en Instagram. De partnerorganisaties deelden deze ook op hun website en hun sociale media kanalen.

De workshops en lezingen werden ook aangekondigd in de nieuwsbrief van amai! Tot slot werden er een aantal workshops ook toegevoegd aan de online kalender [Uit in Vlaanderen](#). Een aantal workshops uitmaakten deel uit van een groter event. In Roeselare maakte de workshop “AI en randvoorwaarden” bijvoorbeeld deel uit van het event ‘Tournée Digitale’, de dag van de digitale geletterdheid.



Figuur 6: Promotie voor de workshops via sociale media.

4.3 Verloop

Onderstaande tabel geeft de agenda weer van een workshopssessie. Elke workshop had dezelfde agenda.

De opstelling van de ruimte voor de workshop samen met de AI-demostand gebeurde minimaal een uur voor de starttijd. Deelnemers konden de stand ontdekken en bij aanvangsuur werden ze uitgenodigd om plaats te nemen aan een gesprekstafel. Een medewerker van amai! bleef de stand permanent bemannen.

Een onthaalverantwoordelijke heette de deelnemers welkom en begeleidde hen naar een gesprekstafel. De persoonsgegevens werden geregistreerd als bewijs van effectieve deelname.

Een workshopmoderator heette op zijn beurt iedereen welkom, legde het doel van de workshop uit en organiseerde een informele introductieronde. De moderator stelde zichzelf voor en vroeg aan de deelnemers om naam en eventueel beroep te vermelden. Ook nodigde de moderator de deelnemers uit om spontaan de volgende twee zinnen aan te vullen:

- “Mijn grootste hoop voor AI is ”
- “Mijn grootste bezorgdheid over AI is ...”

Vervolgens startte de eerste sessie waarbij de workshop moderator uitleg gaf over de randvoorwaarden. Deze randvoorwaarden werden kort omschreven met een voorbeeld. Vervolgens werden de deelnemers gevraagd om een sticker te kleven op de drie belangrijkste randvoorwaarden, of nieuwe te identificeren.

Tijdens de tweede sessie werden er vier cases voorgesteld aan de deelnemers. Volgens de voorkeur van de deelnemers koos men één of meerdere cases en ging men in groep of in kleinere groepjes te werk. Voor elke case werden er boom- en doemscenario's omschreven, alsook handelingsopties. Na deze sessie kon men even bij de burens kijken en verder aanvullen bij een andere case.

De laatste sessie nodigde de deelnemers uit voor een ministerspeech. Daarbij werd hen gevraagd om een advies te formuleren als minister van innovatie over de toekomst van AI in Vlaanderen.

Tussen sessie 2 en 3 konden de deelnemers even pauzeren.

0	Vorbereiding locatie: AI demo stand en workshop ruimte	'60
0	Ontdekrunde: kennismaking met AI via de AI demo stand	'30
1	Onthaal en informele intro	'10
2	Kennismaking	'10
3	Sessie 1: "Prioritering en identificatie randvoorwaarden"	'25
4	Sessie 2: "Boom en doemscenario's"	'40
5	Sessie 3: "Vul aan bij de andere tafel"	'20
6	Sessie 4 - plenair: Speech als de minister van innovatie	'15
7	Afronden en outro	'5
8	Debriefing/opruimen

Tabel 4: Agenda van de workshop

4.4 Enkele sfeerbeelden



Figuur 7-8: Sfeerbeelden demostand.



Figuur 9-10: Sfeerbeelden workshops.

5 Resultaten

In dit hoofdstuk worden de resultaten samengevat van de workshops. De aanpak van de methodiek wordt geëvalueerd en een aantal resultaten van de cases worden uitgelicht.

5.1 Overzicht aantal deelnemers

In totaal werden er vier workshops “AI en randvoorwaarden” georganiseerd en drie lezingen. Het totaal aantal deelnemers was 118. Een overzicht van het aantal deelnemers kan je terugvinden in onderstaande tabel:

Datum	Locatie	Activiteit	Aantal deelnemers
7 september 2022	ARhus bibliotheek, Roeselare	Workshops (2)	Workshop 1: 12 Workshop 2: 13
5 november 2022	Bib Beringen	Workshop (1)	15
10 november 2022	Avansa Oostende – Astropolis	Lezing (1)	22
22 november 2022	Stadsbibliotheek Leuven	Lezingen (2)	Lezing 1 (volwassenen): 15 Lezing 2 (scholen): 28
3 december 2022	Samana vorming dag, Erasmushogeschool Brussel	Workshop (1)	13
		Totaal	118 deelnemers

Tabel 5: Aantal deelnemers per locatie.

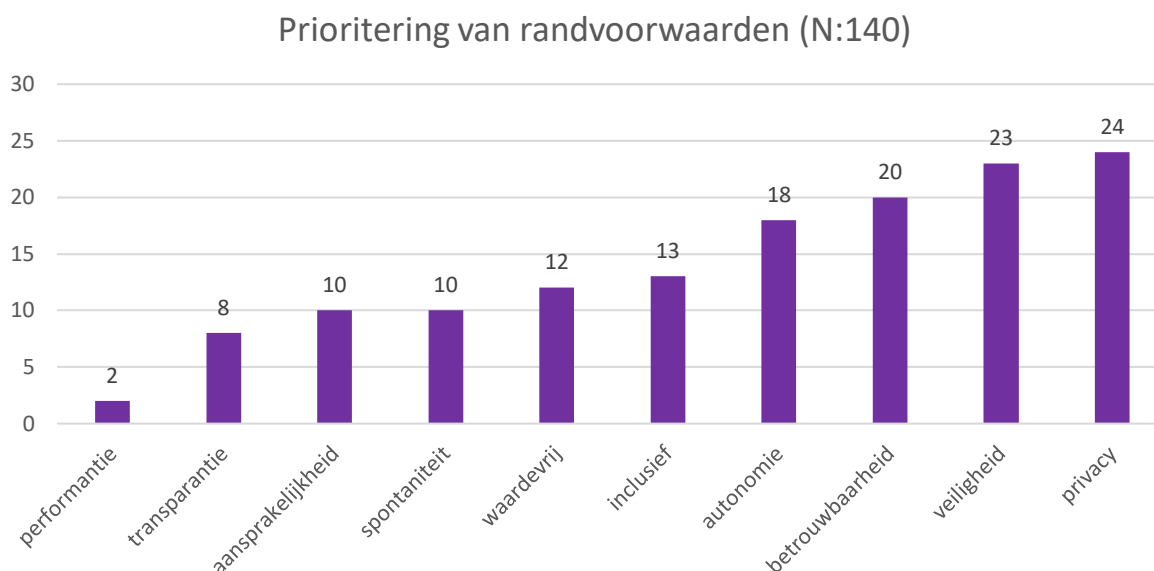
Het profiel van de maatschappelijke toepassers was duidelijk aanwezig tijdens de workshops en lezingen. Er waren geen AI-experten of professionals, en er was een goede mix tussen oud en jong. Het merendeel van de deelnemers had nog niet gehoord van AI, of wist niet goed wat het juist betekent. De voorkennis over AI was iets hoger tijdens de lezingen dan tijdens de workshops. Het spel “AI of ni” werd indien nodig als extra opwarmer gebruikt tijdens de introductieronde van de workshops. Ook werd er gevraagd in de groep of iemand in één zin AI kon omschrijven. Op deze manier startte iedereen met een basiskennis over de technologie.

Een aantal deelnemers schreven zich in via hun organisatie voor een workshop of lezing, het merendeel van de deelnemers schreef zich in uit interesse en om bij te leren over de technologie.

We merkten echter op dat de AI-demostand minder bezoekers lokte dan verwacht. Ondanks dat hier promotie werd voor gemaakt leek het erop dat deelnemers aan de workshops hier minder interesse voor hadden. Een mogelijke reden hiervoor is dat deelnemers niet de tijd hadden om extra vroeger te komen of even na te blijven voor of na te workshop. Ook kan de opstelling een mogelijke verklaring zijn. De demostand werd steeds geïnstalleerd nabij een workshop ruimte, maar niet in de workshop ruimte. Het exploreren van de demostand op eigen houtje kon een mogelijke drempel zijn.

5.2 Overzicht aantal stemmen voor randvoorwaarden

In de eerste sessie van de workshops gingen deelnemers aan de slag met de randvoorwaarden. De randvoorwaarden werden kort voorgesteld en vervolgens prioriteerden deelnemers deze met behulp van drie stickers. In totaal werden er tijdens de vier workshops 140 stemmen uitgebracht. Maximaal bracht een gesprekstafel stemmen uit verspreid over negen randvoorwaarden en minimaal over vijf. In onderstaande grafiek kan je het aantal stemmen per randvoorwaarde aflezen:



Tabel 6: Aantal stemmen per randvoorwaarde – prioritering naar mate van belangrijkheid.

De randvoorwaarden ‘**privacy**’, ‘**veiligheid**’ en ‘**betrouwbaarheid**’ behaalden de meeste stemmen. Bij de discussie rond privacy verwees men vaak naar niet naleving of bescherming van persoonlijke gegevens, voornamelijk binnen het domein gezondheid. Ondanks dat er een duidelijk wetgevend kader hieromtrent bestaat, bleek de bezorgdheid toch hoog. Ook veiligheid behaalde een hoge score. Hierbij werd opgemerkt door een deelnemer dat systemen zich nog zo goed kunnen beschermen, maar als iemand kwade bedoelingen heeft dat er steeds misbruik kan worden gemaakt. Een aantal bezorgde deelnemers deelden ook dat veiligheid boven alles staat, voornamelijk omdat AI-systemen zelflerend zijn en men hier de bezorgdheid heeft dat men controle erover kan verliezen.

Het minste aantal stemmen werd gegeven aan ‘performantie’ en ‘transparantie’. Er werden acht jokerkaarten ingediend, deze bestempelden de randvoorwaarden:

- Authenticiteit
- Transparante communicatie (zowel positief als negatief)
- Macht (voor patiënten)
- Datavolledigheid (wanneer is een AI-systeem volleerd?)
- Moraliteit
- Inspraak
- Controle (voor gebruikers: bepalen wanneer je AI kan aan-of afzetten)
- Censuur

Vier van deze acht randvoorwaarden verwijzen naar de verhouding tussen technologie en maatschappij, waarbij men vanuit de maatschappij meer zeggenschap wenst, namelijk door de nieuwe randvoorwaarden ‘macht’, ‘inspraak’, ‘controle’ en ‘censuur’.

De jokerkaart 'datavolledigheid' linkt met de randvoorwaarde van betrouwbaarheid. Tijdens de voorfase werd er al meermaals gehamerd dat adviezen van AI-systemen volledig, correct en betrouwbaar moeten zijn.

De jokerkaarten 'authenticiteit' en 'moraliteit' zijn nieuwe waarden.

5.3 Doem- en boemscenario's

Tijdens de tweede sessie van de workshops gingen de deelnemers aan de slag met de vier cases. Zij konden zelf kiezen welke case ze graag wouden behandelen. De deelnemers hadden een grote voorkeur voor de gezondheidscase en is in totaal vijf keer behandeld over de vier workshops. Tijdens de voorfase van het amai-project in 2021 werd voor dit domein ook de meeste input verzameld. De andere cases over werk, klimaat & milieu werden elk tweemaal besproken over de vier workshops. Gezien de **gezondheidscase** het meeste werd behandeld, gaan we in de volgende paragrafen hier verder op in.

Gezondheidscase

Een internationaal team van onderzoekers ontwikkelde een AI-systeem dat louter op basis van oogscans kan inschatten wie een verhoogd risico loopt om binnen het jaar een hartaanval te krijgen. Wanneer artsen en patiënten van deze AI-oplossing gebruik willen maken, moeten ze - net zoals bij andere nieuwe technologieën - met veel rekening houden. Wanneer, waarvoor en hoe willen we de technologie gebruiken? Wat vind jij belangrijke voorwaarden om een AI-toepassing, zoals deze oogscan, te kunnen vertrouwen?

Het bedenken van doem- en boemscenario's verliep vlot voor de deelnemers. In totaal werden er meer doemscenario's (19) geïdentificeerd dan boemscenario's (12).

De **doemscenario's** voor de gezondheidscase speelden zich af rond volgende waarden:

- **Inbreuk op je privacy:** Gezondheidsdata worden ongevraagd gedeeld of opgevraagd met of door werkgevers, uitzendkantoren, verzekeraars, bankinstellingen, etc. Hierdoor kan het mogelijk zijn dat je minder kans hebt op een nieuwe job, jouw job verliest, geen lening krijgt of extra moet betalen bij afsluiten van een verzekering.
- **Gevoel van wantrouwen en onveiligheid:** Wat als er een verkeerde diagnose ('vals negatief') wordt gesteld, met als gevolg een verkeerde behandeling voor de patiënt? Wat als er verborgen gebreken zijn aan het AI-systeem? Wat als het AI-systeem plots faalt? Wantrouwen in het systeem kan ontstaan.
- **Performantie in gedrang:** Wat als iedereen een dergelijke oog scan wil laten nemen uit angst? Bij een hoge vraag naar het afnemen van een oog scan kan het systeem falen, wachttijden worden lang en de kostprijs kan stijgen.
- **Inspraak wordt niet toegestaan:** Wat als je deze test niet meer vrijblijvend kan doen, maar ertoe verplicht wordt? Wil iedereen deze test doen en betalen? Er is ook de vrees dat het delen van deze gezondheidsdata een informele verplichting zou worden bij bepaalde instanties.
- **Inclusie & solidariteit in gedrang:** Hoeveel kost deze test? Kan iedereen het betalen? Gaat de mutualiteit het dekken voor iedereen? Bepaalde groepen kunnen mogelijk uitgesloten worden.

- **Gebrek aan vakmanschap:** Artsen zouden er kunnen van uit gaan dat het AI-systeem het werk voor hen doet, en dat het systeem noch de arts bijkomende scholing nodig heeft. Dit zou kunnen leiden tot verkeerde diagnoses en een paardebrileffect. Ook stelden deelnemers zich de vraag of artsen nog beroep zouden doen op specialisten buiten hun eigen vakgebied. Zou de samenwerking in het gedrag komen doordat men nu kan leunen op dit systeem?
- **Gebrek aan persoonlijk contact:** Deelnemers zagen het mogelijke risico dat de relatie arts-patiënt kon verwateren, waarbij er geen opvang meer is voor de patiënt.

De *boemscenario's* voor de gezondheidscase speelden zich af rond volgende waarden:

- **Vernieuwing van gezondheidszorg:** Het AI-systeem speelt in op preventieve zorgverlening, waardoor de curatieve zorgverlening afneemt. Men kan patiënten in een vroeg stadium behandelen, en dit kan leiden tot minder of geen sterfte. Bovendien kan dit ook leiden tot minder invasieve ingrepen in een later stadium van de ziekte.
- **Verbetering van levenskwaliteit:** Patiënten kunnen gezonder en langer leven. Er zijn minder complicaties, en dus minder onderzoeken nodig voor de patiënt waardoor het comfort stijgt.
- **Deskundigheid:** Door het AI-systeem kunnen artsen nauwkeuriger en vlotter werkt. De arts gebruikt het als een extra tool en hulpmiddel voor triage van patiënten. Het is tijdsbesparend.

5.4 Handelingsopties

Vanuit de aanpak van begeleidingsethiek worden handelingsopties bedacht om technologie en samenleving op een ethische manier te begeleiden. *Handelingsopties* zijn mogelijke oplossingen met concrete acties voor de technologie (het ontwerp), omgeving (fysiek, sociaal of juridisch) en de gebruiker (gedrag, kennis of vaardigheden). Tijdens de workshop brainstormden de deelnemers over oplossingen en acties om doemscenario's te vermijden of boemscenario's te realiseren.

Onderstaande tabel vat de handelingsopties samen voor de technologie, omgeving en gebruiker voor de gezondheidscase:

Technologie	Omgeving	Gebruiker
<ul style="list-style-type: none"> • Transparantie in het ontwerp: het AI-systeem informeert welke data wordt gedeeld. Data kan niet automatisch gedeeld worden met derden zonder toestemming. Er is een opt-out optie. • Voor eeuwig zelflerend: Een AI-systeem moet continue blijven bijleren en leren uit resultaten • Tussenkoms: Bij foute interpretatie komt een arts tussen 	<ul style="list-style-type: none"> • Wetgeving GDPR • Controle: er is een controlerende instantie dat data niet gedeeld wordt met derde partijen. Het medische dossier is geheim • Opleiding artsen: Opleidingen leren artsen kritisch om te gaan met AI-systemen, er is een blijvende scholing • Criteria vastleggen voor screening van patiënten, bijvoorbeeld erfelijkheidsfactoren 	<ul style="list-style-type: none"> • Kritische mindset van patiënten • Patiënten bouwen vertrouwen op in AI door raadplegen van informatie- en communicatiecampagnes • Patiënten - Sensibiliseringscampagnes met (bekende) Vlamingen die getuigen over hun ervaring

Tabel 7: Handelingsopties – gezondheidscase.

5.5 Ministerspeech

De laatste sessie in de workshop focuste op het formuleren van een advies vanuit de rol van de minister van innovatie in Vlaanderen. Met deze afsluitende oefening willen we de stem van gebruikers van AI laten horen en hen ondersteunen in een ethisch denkproces.

We vatten de ministerspeech samen door enkele quotes over te nemen van de deelnemers:

*“Breng het sociale aspect van AI in beeld: hoe kan AI ten goede zijn voor de mens? Breng ervaringsdeskundigen in beeld, en maak er geen top-down verhaal van”
(deelnemer, Roeselare)*

*“AI staat den dienste van de mens, het is in het belang van de mens – en niet andersom”
(deelnemer, Brussel)*

“Maak geen blinde innovatie. Informeer en breng een positief en negatief verhaal” (Deelnemer, Roeselare)

“Stimuleer een kritische bevolking: iedereen moet innovatief mee helpen denken, en mee werken aan innovatie (Deelnemer, Roeselare)”

“Stel autonomie voorop: mensen moeten kunnen kiezen wanneer en hoe ze met AI in aanraking komen” (Deelnemer, Beringen)

6 Aanbevelingen aanpak begeleidingsethiek

In dit afsluitende hoofdstuk formuleren we een aantal richtlijnen op basis van de goede praktijken van de workshops en de lezingen. De verschillen met de aanpak begeleidingsethiek van Professor Verbeek worden uitgelicht en geëvalueerd in hun opzet.

De aanpak van begeleidingsethiek hielp ons om een dialoog aan te gaan met het profiel van de maatschappelijke toepassers. Dit profiel heeft geen of een beperkte kennis over AI, en is al dan niet een bewuste gebruiker van de technologie. Door in dialoog te gaan met hen konden we bezorgdheden, ethische dilemma's alsook opportuniteiten van de technologie bespreken. Dit deden we aan de hand van concrete cases, waarvoor doem- en boemscenario's werden bedacht alsook oplossingen met concrete acties voor de omgeving, de technologie en de gebruiker.

Onze richtlijnen en aanbevelingen groeperen we volgens de drie fases van de aanpak begeleidingsethiek:

Fase 1: Technologie in context

- Binnen het amai!-project is er veel aandacht besteed om deelnemers op de hoogte te brengen over de basisbeginselen van AI. Door het organiseren van een demostand konden deelnemers AI spelenderwijs ontdekken. Door het aanreiken van diverse voorbeelden probeert het amai!-project bewustzijn te creëren en aan te tonen dat de technologie veelzijdig en alledaags is. Bij het begin van de workshop werd er ook gevraagd om AI te omschrijven in één zin.
- Bij de voorstelling van de 4 cases werd er gewerkt met een concreet scenario. Op deze manier konden de deelnemers de technologie in context scherpstellen.
- Het grote verschil met de aanpak van Professor Verbeek is dat er geen concrete case "in situ" is gekozen die effectief erna opgevolgd wordt met behulp van de geïdentificeerde handelingsopties. Voor de workshops koos het amai!-project ervoor om te werken met actuele voorbeelden en hierover een inclusieve, ethische discussie te stimuleren.

Fase 2: Dialoog

- In totaal namen 118 mensen deel aan de workshops en de lezingen. De dialoogvorm kon enkel toegepast worden binnen het workshop format.
- In tegenstelling tot de aanpak van Professor Verbeek werd ervoor gekozen om specifiek te focussen op het profiel van de maatschappelijke toepassers. In dit opzicht was de samenstelling van de deelnemers minder divers, maar wel inclusief voor deze specifieke groep.
- De groepsgrootte was goed bij alle workshops en lezingen. Er was voldoende interactie tussen de deelnemers van 1 groep, maar niet tussen de verschillende groepen. Doordat de deelnemers zelf konden kiezen welke case ze graag wilden behandelen, ofwel gezamenlijk of in groep, verschilde de timing tussen de verschillende gesprekstafels.
- Het samenstellen van de gesprekstafels kon evenwichtiger. Er waren soms sterke verbindingen tussen personen aanwezig waardoor de conversatie door hen gedomineerd werd.
- De fiches met de randvoorwaarden bleef centraal liggen op tafel en dit bevorderde de discussies. Deze toevoeging aan het oorspronkelijke format van de sessie, leek de maatschappelijke toepassers te helpen in de gesprekken.
- Het creatief nadenken werd gestimuleerd door de workshop moderator. Doem- en boemscenario's werden gemakkelijk geïdentificeerd en uitgelicht met voorbeelden uit het dagelijkse leven.

Fase 3: Handelingsopties

- Het brainstormen over de handelingsopties verliep vlot. Echter werd er geen duidelijk onderscheid gemaakt voor deze opties voor boem- of doemscenario's. In dit opzicht was het overbodig om dit tweemaal te bediscussiëren.
- De deelnemers slaagden er niet altijd in om de handelingsoptie in de juiste categorie te plaatsen (omgeving, technologie of gebruiker). Hier assisteerde de moderator.
- De doem-en boemscenario's werden besproken vanuit de verschillende stakeholderperspectieven, maar dit was minder het geval voor de handelingsopties. De handelingsopties waren eerder generiek van aard.

Deze workshops bleken succesvol in hun bedoeling om ook bij burgers volgens de principes van de 'aanpak begeleidingsethiek' in gesprek te gaan over maatschappelijke en ethische implicaties van AI-toepassingen. Het format zal daarom verder worden aangeboden in volgende edities van amai! en de methodiek zal verder worden ontsloten in train-the-trainer opleidingen, waarbij leerkrachten en trainers ook zelf met de aanpak aan de slag kunnen gaan.

7 Bijlages

Bijlage 1 – Risico's en voorwaarden per domein

Domein: Klimaat & milieu	
Risico's	Voorwaarden
1) Privacy van de gebruiker wordt niet gerespecteerd 2) Risico op averechts effect (Bijvoorbeeld gebrek aan capaciteiten als gevolg van veelvuldig gebruik)	1) Een performant systeem: <ul style="list-style-type: none"> • Voldoende rekencapaciteit 2) Datakwaliteit: <ul style="list-style-type: none"> • Goede training algoritme • Up-to-date data 3) Ethisch correcte analyse (bijvoorbeeld geen greenwashing) 4) Transparantie: <ul style="list-style-type: none"> • Informeren hoe de analyse van de data gebeurt • Informeren welke stakeholders zijn betrokken in het proces 5) Steun van de overheid bij ontwikkeling van bepaalde systemen in voor publiek belang (bijvoorbeeld communicatie ondersteuning)

Domein: Mobiliteit	
Risico's	Voorwaarden
1) Privacy van de gebruiker wordt niet gerespecteerd (met bijvoorbeeld risico op traceerbaarheid) 2) Misbruik maken van de data (bijvoorbeeld risico op terreur, hacking, manipulatie van data, valse informatie, etc.) 3) Bias in het geven van een advies of oplossing (bijvoorbeeld voorkeur in gesponsorde activiteiten) 4) Mogelijke veiligheidsrisico's: <ul style="list-style-type: none"> • Wat als het het systeem faalt? • Wat als de oplossing niet strookt met gedrag van mensen? 5) Stijgende complexiteit bij combineren van datasets	1) Datakwaliteit: <ul style="list-style-type: none"> • Betrouwbare data • Voldoende data • Evidence-based data • Goede meta data • Near-real time data 2) Conformiteit met GDPR richtlijnen <ul style="list-style-type: none"> • Gebruik van drones

Domein: Werk

Risico's	Voorwaarden
<ol style="list-style-type: none"> 1) Privacy van de gebruiker wordt niet gerespecteerd 2) Risico op bias of filter bubble 3) Verlies aan spontaniteit (de gebruiker leunt te veel op de technologie) 4) Bias in het geven van een advies of oplossing 5) Risico op averechts effect (bijvoorbeeld de kans op vervolging vergroot) 6) Onbedoelde bias/ unconscious bias (bijvoorbeeld te algemene aannames) 	<ol style="list-style-type: none"> 1) Datakwaliteit: <ul style="list-style-type: none"> • Voldoende data • Relevante data • Actuele data 2) Conformiteit met GDPR-richtlijnen 3) Transparantie 4) Human in the loop – controle inbouwen 5) Ethisch, correcte analyse (geen fake nieuws, geen politiek gekleurde oplossing)

Bijlage 2 – Fiches randvoorwaarden



AANSPRAKELIJKHEID

Het is duidelijk wie er aansprakelijk wordt gesteld als er een fout wordt gemaakt met de AI-toepassing.

Prioritering:



AUTONOMIE

"Human in the loop": de aanbevelingen van de AI-toepassing zijn minderwaardig dan deze van de mens. Er is nog steeds menselijke controle.

Prioritering:



BETROUWBAARHEID

Bij het gebruik van een AI-toepassing, verwacht ik een betrouwbaar advies. De gegeven informatie is correct, of aangeduid volgens de schaal van waarschijnlijkheid.

Prioritering:



INCLUSIEF

AI-toepassingen sluiten niemand uit en zijn afgestemd op diverse profielen. Er is nagedacht over mogelijke barrières, zoals vaardigheden of kostprijs.

Prioritering:



PERFORMANTIE

De AI-toepassing werkt naar behoren. Het systeem heeft de nodige (reken)capaciteiten om vraag met aanbod af te stemmen.

Prioritering:



PRIVACY

AI-toepassingen respecteren de privacy richtlijnen en gaan correct om met gevoelige data. Ik weet welke data er over mij verzameld worden, en met welk doel.

Prioritering:



SPONTANITEIT

Een gebruiker mag niet afhankelijk worden van een AI-toepassing. Er moet nog steeds ruimte zijn voor creativiteit en ontdekking.

Prioritering:



TRANSPARANTIE

De AI-toepassing geeft uitleg over het doel van het systeem, de verzamelde data of informatie en de type analyses.

Prioritering:



VEILIGHEID

De AI-toepassing is goed beveiligd en dekt zich in tegen mogelijke veiligheidsrisico's, zoals fraude of misbruik. De verzamelde data wordt veilig bewaard.

Prioritering:



WAARDEVRIJ

De AI-toepassing discrimineert niet. De data is representatief, en de analyses zijn ethisch correct. De gekozen oplossing is onpartijdig.

Prioritering:

