

1 INHOUD		
1	Inhoud	1
2	Samenvatting	2
3	Inleiding	2
4	Thematische afbakening	4
5	Vraagarticulatie	5
6	Het netwerk	6
7	Projectplan	6
8	Datamanagementparagraaf	8
9	Prestatie-indicatoren	9

2 SAMENVATTING

Zonder expliciet bezwaar wordt de samenvatting in geval van honorering van uw projectvoorstel op www.nwo.nl en op www.sia-projecten.nl gepubliceerd.

Aantal woorden	288
----------------	-----

Titel: Doorgrond Discriminatie Digitale Differentiatie

De Nederlandse financiële sector is één van de vier grote sectoren waar AI wordt toegepast [CBS2021]. Zo maken de verzekeraars betrokken bij onze praktijkpartners MavenBlue en het Verbond van Verzekeraars gebruik van *digitale differentiatie (DD)*: het gebruik van zelflerende algoritmes om het financieel risico per klant te bepalen; op dit risico baseren verzekeraars de premiehoogte en of ze de klant aannemen of weigeren. Verzekeraars willen zo het rendement van hun producten bewaken zodat ze de *financiële duurzaamheid* van hun bedrijf kunnen waarborgen. Tegelijk, willen ze dat het onderscheid dat ze maken *eerlijk* is (m.a.w. niet discriminerend) zodat ze wetgeving volgen en boetes en imagoschade voorkomen. Bovendien, vanuit de wetgever, toezicht en de samenleving neemt de druk toe om de eerlijkheid van digitale differentiatie te kunnen verantwoorden. Dit roept de vraag op:

hoe kunnen verzekeraars digitale premie-differentiatie verantwoorden vanuit zowel eerlijkheid als rendement?

Voor onze praktijkpartners is het een uitdaging op deze vraag een antwoord te vinden doordat het een integratie vraagt van wetgeving, ethiek en techniek in huidige werkprocessen. Daarom volgt ons onderzoek een 'design science'-aanpak: we ontwerpen een methode om DD te verantwoorden en gebruiken deze methode om de behoeftes scherp te krijgen vanuit verzekeraars en software. Dit doen we door middel van interviews, een literatuurstudie, ontwerp en vier co-design workshops waar de praktijkpartners de methode evalueren en hun behoeftes aangeven. Het resultaat van dit onderzoek is tweeledig: een methode om DD te verantwoorden en een evaluatie van deze methode in termen van de behoeftes van de praktijk. Dit is relevant voor een bredere validatie van de methode in een RAAK en direct relevant voor de praktijkpartners zodat ze boetes en imagoschade voorkomen en voor de maatschappij om discriminatie door zelflerende algoritmes te doorgronden en te voorkomen.

3 INLEIDING



Hoofdstuk 3 t/m 7 wordt beoordeeld door de beoordelingscommissie. Het maximaal aantal voor deze hoofdstukken is 2.500 woorden. Teksten in tabellen, voetnoten en figuren tellen ook mee voor het woordenaantal. **Let op:** de schrijfinstructies geven een handvat bij het schrijven van het projectvoorstel en zijn richtinggevend bedoeld.

Totaal aantal woorden hoofdstuk 3 t/m 7 (max. 2.500)	634 + 58 + 646 + 423 + 739 = 2500
---	--------------------------------------

In 2021 gebruikte 25% van de financiële instellingen AI [1]. Zo maken de verzekeraars betrokken bij onze praktijkpartners MavenBlue en het Verbond van Verzekeraars (VvV) gebruik van digitale differentiatie (DD): het gebruik van zelflerende algoritmes om het financieel risico per klant te bepalen; op dit risico baseren verzekeraars de premiehoogte en of ze de klant aannemen of weigeren. Vanuit politiek, wetenschap, samenleving, wetgeving, toezichthouders en de verzekeraars zelf is er een roep voor het borgen van eerlijkheid van deze zelflerende algoritmes [2], [3]. Tegelijk vinden verzekeraars het belangrijk dat AI bijdraagt bij het rendement verhogen van hun producten zodat ze de financiële duurzaamheid van hun bedrijf kunnen bewaken. Dit roept de praktijkvraag op:

Hoe kunnen verzekeraars en software-makers digitale premie-differentiatie verantwoorden vanuit zowel eerlijkheid als rendement?

Verzekeraars vinden deze vraag relevant om imagoschade en boetes te voorkomen en op een duurzame manier huidige en aankomende wetgeving te integreren.

Voor onze praktijkpartners is het een uitdaging op deze vraag een antwoord te vinden doordat het een integratie vraagt van wetgeving, ethiek en techniek in huidige werkprocessen:

- **Wetgeving en Ethiek** - Onze partners kennen en volgen specifieke wetgeving en rechterlijke uitspraken die verbieden om gevoelige variabelen direct te gebruiken (bijv. geslacht of nationaliteit) [4]. Echter, het is onduidelijk wat de juridische normen zijn voor indirect onderscheid op deze variabele (bijv. het via postcode indirect onderscheid maken op nationaliteit); zowel bij onze directe partners, overige betrokkenen als bij de wetgever zelf [5]. De wetenschappelijke literatuur over ethiek biedt genuanceerder begrip van wat eerlijk is (en discriminatie) en kan stabiele antwoorden bieden die vooruitlopen op huidige wetgeving [6], [7], maar zal moeten worden vertaald naar huidige technische en menselijke werkprocessen.
- **Techniek** - Onze partners integreren zelflerende algoritmes in hun werkproces en testen hoe accuraat deze zelflerende algoritmes een risico kunnen voorspellen; echter, een integratie in het werkproces van een controle op eerlijkheid ontbreekt. Deze integratie is uitdagend want er is niet één manier om eerlijkheid te meten, er zijn meerdere manieren om eerlijkheidsscores te verbeteren en de techniek zelf biedt geen uitsluitel over hoe je eerlijkheid weegt tegenover rendement.

Dus, op dit moment weten de praktijkpartners niet hoe ze hun keuzes moeten verantwoorden ten opzichte van ethische principes over eerlijkheid doordat zulke principes niet geïntegreerd zijn in hun werkprocessen, terwijl de druk vanuit wetgever [8] en toezichthouders [9] wel toeneemt.

Wetenschap biedt een deel van het antwoord omdat het laat zien hoe men ethische principes integreert met techniek. Bijvoorbeeld, hoe men de eerdergenoemde eerlijkheidsscores berekent, wat men kan doen om ze te verbeteren en hoe de scores zich ethisch met elkaar verhouden [10]. Daarnaast biedt ethische literatuur helderheid over wat bepaalde variabelen gevoelig maakt (bijv. hoeveel invloed je hebt op deze variabele en hoe je ethische keuzes maakt in proportie tot andere belangen (bijv. rendement of accuraatheid ([11], [12])). De literatuur biedt geen uitsluitel over de perfecte keuze, maar wel de kennis die de praktijk kan helpen in het verantwoorden van hun keuzes. Om deze wetenschappelijke kennis over AI-ethiek te integreren in de praktijk is het nodig om de behoeftes vanuit de praktijk te begrijpen en de uitdagingen te vinden bij het toepassen van wetenschappelijke methodes.

Onze onderzoeksvraag is dan ook:

Welke behoeftes hebben verzekeraars en software-makers bij het stapsgewijs verantwoorden van digitale premiedifferentiatie in termen van eerlijkheid en rendement?

De behoeftes van de praktijk worden verzameld in termen van *user stories* en *menselijke waarden* (zie Sectie 7). We onderzoeken deze vraag met een *design science* aanpak: we ontwerpen een methode voor DD en gebruiken deze methode om de behoeftes scherp te krijgen (zie Sectie 7). De methode zelf zal worden gedeeld met de sector (via het VvV) en de evaluatie zal de focus zijn van een wetenschappelijke publicatie. Samen bieden ze een basis voor een RAAK-onderzoek waar de methode kan worden gevalideerd voor de bredere financiële sector.

4 THEMATISCHE AFBAKENING

Dit voorstel sluit aan bij de sleuteltechnologie Artificial intelligence in de categorie Digital and Information Technologie. Dit voorstel bestudeert DD: het onderscheid maken tussen klanten met behulp van zelflerende algoritmes. Dit voorstel test hoe we zelflerende algoritmes zowel eerlijk en winstgevend kunnen inzetten in de praktijk en draagt hiermee bij aan de duurzame inzetbaarheid en ontwikkeling van AI.

5 VRAAGARTICULATIE

Uit rondgang onder financiële instellingen kwam de volgende vraag uit de praktijk: wanneer is differentiatie door een zelflerend algoritme te verantwoorden? Dit speelde ook in onze huidige FIN-X RAAK-mkb consortium bestaande uit verschillende partijen in de financiële sector. Waar FIN-X zich richt op een algemene uitleg van AI, focust deze KIEM zich op het specifiek verantwoorden van Digitale Differentiatie (DD) in termen van eerlijkheid. Dat discriminatie door DD meer aandacht verdient werd bevestigd door verschillende partijen binnen en buiten het consortium: Tjerry Smit van Nationale Nederlanden, Hylke Niermeijer van Deloitte, Marijn van Aerle CTO van Floryn en Uri Schimron van het AP. Daarnaast, hebben we hebben met drie partijen verdiepende gesprekken gevoerd om tot een verscherping van de praktijkvraag te komen; deze doen mee als praktijkpartner: het Verbond van Verzekeraars, MavenBlue en Info Support.

Jos Schaffers (van het Verbond van Verzekeraars) bevestigt dat verzekeraars het lastig vinden DD te verantwoorden omdat ze een balans willen vinden tussen financiële duurzaamheid van het bedrijf en de eerlijkheid van AI-modellen. Specifiek noemt Jos dat een eerdere poging om een Canadees bedrijf te koppelen aan verzekeraars geen weerklank vond omdat dit bedrijf niet goed aansloot bij de praktische en juridische context van Nederlandse verzekeraars. Jos zet zich in voor dit onderzoek zodat het wel aansluit bij Nederlandse verzekeraars, en ze helpt om te gaan met nationale wetgeving, eerlijkheidsscores, proxyvariabelen en de balans tussen techniek en ethiek.

Joost van Bruggen (consultant bij MavenBlue) bevestigt dat MavenBlue graag functionaliteit toevoegt aan haar software waardoor ze DD kunnen verantwoorden vanuit modelperformance (rendement) en eerlijkheid. MavenBlue levert software waarmee medewerkers van verzekeraars de premiehoogte van schadeverzekeringen kunnen bepalen. Hij herkent de complexiteit van de balans tussen rendement en eerlijkheid en benoemt als voorbeeld dat er binnen sommige verzekeraars discussies worden gevoerd tussen sales en productmanagement die blijven hangen in deze tegenstelling zonder een oplossing te zien.

MavenBlue ziet dat hun klanten zich aan de concrete wetgeving houden (zoals de AVG, WBFT en WFT) en niet direct gevoelige variabelen meenemen die benoemd worden in deze wetgeving (bv. geslacht). Echter, in zijn algemeenheid is er geen concrete aanpak voor proxy-variabelen of variabelen die niet expliciet worden genoemd maar wel ethisch twijfelachtig kunnen zijn (bv. postcode). Dit zorgt nu al voor spanning, echter met de EU AI-verordening moet hier ook juridisch meer transparantie over komen. MavenBlue wil graag het overzicht hebben over huidige methodes over eerlijke differentiatie; en geadviseerd worden bij het implementeren van deze methodes in hun software.

Joop Snijder (Hoofd van het onderzoekscentrum AI van Info Support) bevestigt dat ze een inschatting willen kunnen maken wanneer DD te verantwoorden is. Info Support is een software- en AI-bedrijf dat ICT-oplossingen biedt voor klanten o.a. in de financiële sector. Op dit moment wijst Info Support klanten af wanneer ze niet weten of een project ethisch wel verantwoord is: hoe kunnen ze verantwoorden dat ze DD toepassen op de dataset die wordt aangeboden door de klant? Info Support zou geholpen zijn met een concrete aanpak waarmee ze deze vraag kunnen beantwoorden en op deze manier kunnen bepalen of ze een project kunnen aannemen.

Samenvattend willen alle praktijkpartners weten wanneer DD te verantwoorden is en dat de balans tussen rendement en eerlijkheid hierin centraal staat. Hun huidige instrumenten kunnen deze vraag niet beantwoorden en de state-of-the-art is niet concreet genoeg om gelijk te gebruiken in bedrijfscontext. Uit gesprekken met de eerdergenoemde overige betrokkenen blijkt dat deze vraag breed leeft in de sector. De combinatie van de verkennende en de verdiepende gesprekken met de partners uit de praktijk komen we dan ook tot de praktijkvraag: **Hoe kunnen verzekeraars en software-makers digitale premie-differentiatie verantwoorden vanuit zowel eerlijkheid als rendement?** Om dit te bereiken moeten we weten hoe wetenschappelijke methodes voor dit probleem passen bij de praktijk. De onderzoeksvraag die hierbij past is: **welke behoeftes hebben verzekeraars en software-makers bij het stapsgewijs verantwoorden van digitale premiedifferentiatie in termen van eerlijkheid en rendement?**

6 HET NETWERK

Dit project bouwt voort op het FIN-X RAAK-mkb-consortium. We versterken de band tussen een deel van de consortiumpartners door een nauwe samenwerking tussen:

- De Hogeschool Utrecht, een kennisinstelling die werkt aan Coöperatieve AI binnen het Lectoraat AI met verschillende Responsible AI onderzoeken. Bijvoorbeeld, het eerdergenoemde RAAK FIN-X onderzoek en een ZonW-project over het ontwerpen van ethische AI in de zorg. Verder sluit dit onderzoeksvorstel aan bij de Human-centered AI Master waar afstudeerders en gastcolleges zullen worden gegeven.
- Het Verbond van Verzekeraars, een koepelvereniging, en daarmee publieke instelling, die de belangen van zijn leden, de verzekeraars, behartigt en verbindt aan de belangen van de samenleving. Het verbond informeert ons over de belangen en uitdagingen van verzekeraars: in de beginfase bij de Datamiddag 2024 en gesprekken, tijdens de stakeholder-workshops door de belangen van verzekeraars te vertegenwoordigen en achteraf door te helpen met de disseminatie via nieuwsbrieven en de Datamiddag 2025.
- MavenBlue, een MKB-organisatie die een software-oplossing maakt waarmee verzekeraars de prijs van verzekeringen kunnen bepalen. MavenBlue kent zowel de technische uitdaging als de uitdaging in consultancy en biedt dit project een casus om de methode op uit te proberen. MavenBlue is betrokken in het opstellen van de eisen en de workshops waar we iteratief de methode testen.
- Info Support, een grootbedrijf dat AI-software maakt voor verschillende domeinen, waaronder verzekeraars. Ze zijn betrokken in alle fases van het project en bieden ons inzicht of de behoeftes die we uitvragen breder worden gedragen, vooral of ze technisch realistisch zijn. Verder helpen ze bij de disseminatie door de resultaten te verspreiden onder hun klantennetwerk.

Naast de consortiumpartners betrekken we andere organisaties om de relevantie van de resultaten van dit project voor een bredere groep te testen, bestaande methodefragmenten te inventariseren en om de informatie te verspreiden. Dit doen we met NN (verzekeraar), Deloitte (consultancy partij), AP (toezichthouder), Dong Nguyen (Onderzoeker Universiteit Utrecht) en Floryn (software-maker)

Dit projectvoorstel sluit aan bij de expertise en doel van het van het SPRONG-programma Responsible Applied AI (RAAIT, 2022-2029) van de Hogeschool van Amsterdam, Hogeschool Rotterdam en HU: AI met oog voor ethiek, in dit geval, eerlijkheid en discriminatie. We halen kennis op en verspreiden kennis binnen RAAIT via vier wegen: bijeenkomsten van de learning community zakelijke dienstverlening, een bijdrage aan de jaarlijkse RAAIT workshop, contact met Sophie Horsman (Onderzoeker HVA naar DD), afstudeerders en gastlezers bij AI-bachelors en masters uit SPRONG.

Met dit KIEM project zullen we behoeftes uit de praktijk ophalen en deze gebruiken voor een RAAK-voorstel met een uitgebreider consortium inclusief de eerdergenoemde overige betrokkenen (zie Projectplan).

7 PROJECTPLAN

Voortbouwing op bestaande kennis

Op dit moment zijn er juridisch-ethische wetten en beschouwingen over Digitale Differentiatie (DD) ([4], [8], [11]) en technische instrumenten voor DD ([12]), echter er ontbreekt een stapsgewijze methode die deze instrumenten integreert in de praktijk. Handleidingen uit de praktijk zoals de handleiding van de Gemeente Amsterdam [13] zijn concreter en dienen als inspiratie maar moeten worden vertaald naar een stappenplan voor Fintech. Onze focus ligt echter voornamelijk op wetenschappelijk literatuur over AI en Ethiek (e.g., [10], [14], [15]) omdat deze probeert een combinatie te vinden tussen gefundeerde ethische (duurzame) antwoorden en technische aspecten.

Zoals eerder gezegd, gebruikt dit onderzoek *user stories* om de behoefte uit de praktijk in kaart te brengen. User stories, uit Agile softwareontwikkeling, worden gebruikt om stakeholders, functionaliteiten en abstracte

doelen te koppelen ([16], [17]). Voor de integratie van ethiek en techniek nemen we een ruim begrip van ‘functionaliteit’ en verzamelen we zowel technische als ethische methode-fragmenten uit de literatuur die helpen bij DD. De abstracte doelen conceptualiseren we in termen van menselijke waarden zoals het eerdergenoemde rendement en eerlijkheid. Het koppelen van ontwerpkeuzes aan gedeelde menselijke waarden zorgt voor wederzijds begrip [18]. Vanuit dit wederzijds begrip is het makkelijker om conflicten tussen ontwerpkeuzes op te lossen of aftewegen [19]. Samengevat, door de behoefte uit de praktijk te vatten in *user stories* en *menselijke waarden* maken we een koppeling tussen concrete ontwerpkeuzes in DD terwijl we oog houden voor menselijke waarden.

Onderzoeksmethode

Het project volgt een *design science* aanpak waarbij op basis van behoeften uit de praktijk en beschikbare kennis een artefact wordt ontworpen [20]. Het artefact in deze studie is de methode om DD te verantwoorden. We gaan door de volgende fases:

1. Ten eerste, wordt de *praktische relevantie* van de methode verzekerd door contact met de praktijk via de Datamiddag 2024 (T1) van de VvV, semi-gestructureerde interviews (T2) en workshops. Hierbij gebruiken we *user stories* en *menselijke waarden* om de behoefte in kaart te brengen en de interviews vorm te geven.
2. Ten tweede, wordt de *gronding* van de methode verzekerd doormiddel van een literatuurstudie en het gebruik van literatuur bij het ontwerp.
3. Ten derde, wordt *het ontwerp en de evaluatie* van de methode verzekerd door zelfstandig onderzoek en co-design workshops. We ontwerpen de methode rond een use case van MavenBlue, namelijk een synthetische dataset die zij intern gebruiken om DD te testen. We herhalen hierbij de volgende stappen:
 - a. Door zelfstandig conceptueel ontwerp wordt een proof-of-concept methode gemaakt door methode-fragmenten direct, aangepast of gecombineerd te integreren. Hierbij wordt rekening gehouden met de eerder geëliciteerde *user stories* en *menselijke waarden* en cruciale conflicten tussen methode-fragmenten expliciet gemaakt.
 - b. Door zogenaamde co-design workshops evalueren we de methode. Hierbij wordt uitgevraagd (1) welke behoeftes worden nog niet vervuld (in termen van *user stories* en *menselijke waarden*) (2) hoe zouden jullie prioriteiten stellen tussen ontwerpkeuzes?

Op deze manier levert dit onderzoek een antwoord op de praktijkvraag: een methode voor DD; en een antwoord op de onderzoeksvraag: een evaluatie van dit artefact in hoeverre het de behoefte van de praktijk behartigt. Door verschillende kanalen worden deze resultaten verspreid en wordt er feedback gevraagd vanuit de bredere sector: de Datamiddag 2025, Nieuwsbrief Bond van Verzekeraars, Publicatie in een vaktijdschrift, Wetenschappelijke Publicatie en optredens in de Podcast AI Voor Moraalridders. Dit KIEM-onderzoek biedt een concrete methode, en een evaluatie waaruit de behoeftes van de praktijk duidelijk zijn. Dit biedt de basis voor een vervolg RAAK-onderzoek waarin we deze methode kunnen valideren voor de bredere financiële sector: met meer praktijkpartners en met een reproduceerbaar onafhankelijk validatieproces.

Activiteitenplan

Werk-pakket	Taak	Omschrijving taak	Beoogd (deel)resultaat	Verantwoordelijk	Wanneer?
1	1	Kick-off & Datamiddag 2024	Exploratieve lijst van methode-fragmenten en waarden	HU + VvV	Sep 2024
1	2	Literatuurstudie naar methode-fragmenten en waarden rond DD	Lijst van methode-fragmenten en waarden	HU	Sep-Nov 2024

1	3	Eliciteren user stories door middel van semi-gestructureerde interviews	Lijst van user stories	HU + Consortiumpartners	Okt-Dec 2024
2	4	Ontwerpen van concept-methode	Proof-of-concept van NDDD-methode	HU	Dec 2024
2	5	Evalueren van methode door workshops	Evaluatie van methode in VSD-termen	HU + Consortiumpartners	Jan 2025
2	6	Driemaal herhalen van taak 4 en 5	(zie hierboven)	(zie hierboven)	Feb, Maart, April 2025
3	7	Verspreiden en publiceren resultaten	Presentaties, Nieuwsbrief, Publicaties, Optredens LC RAAIT NLAIC	HU + VvV	Mei-Jun 2025
3	8	Opstellen RAAK-aanvraag, netwerkvorming	RAAK-aanvraag	HU	Jun-Jul 2025
3	9	Datamiddag 2025		HU +VvV	Sep 2025
4	10	Projectmanagement		HU	Sep 2024- Sep 2025

8 DATAMANAGEMENTPARAGRAAF

1. Worden er data verzameld of gegenereerd die voor hergebruik geschikt zijn?

Verslagen van interviews en workshops zijn vertrouwelijk en niet voor hergebruik beschikbaar. De synthetische dataset van MavenBlue blijft eigendom van MavenBlue en is niet voor hergebruik beschikbaar. De methode, de resultaten van de methode op de dataset, en de eisen uit de praktijk worden wetenschappelijk gepubliceerd.

2. Waar worden de data gedurende het onderzoek opgeslagen?

Voor het project wordt een projectruimte ingericht in HU Research Drive, de opslagvoorziening voor onderzoeksdata van Hogeschool Utrecht. HU Research Drive wordt gehost door SURF en de opslag vindt plaats in Nederland. De omgeving voldoet aan ISO27001 (de standaard voor informatiebeveiliging) en het "Juridisch Normenkader Cloud Services Hoger Onderwijs" en is AVG-compliant. Toegang en rechten worden bepaald door de projectleider en beheerd door de aan het project verbonden datasteward.

3. Hoe worden de data na afloop van het project voor de lange termijn opgeslagen en voor hergebruik beschikbaar gesteld voor derden? Voor wie zijn de data toegankelijk?

Na afloop van het project blijven de geanonimiseerde en gepseudonimiseerde onderzoeksdata gedurende de minimale bewaartermijn van tien jaar opgeslagen op de HU Research Drive en zijn alleen toegankelijk voor de aan dit project verbonden onderzoekers van de HU.

4. Welke voorzieningen (ICT*, (beveiligd) archief, koelkasten of juridische expertise) zijn naar verwachting nodig voor de opslag van data gedurende het onderzoek en na het onderzoek? Zijn deze beschikbaar?

Er zijn geen specifieke voorzieningen voor dit project nodig. Zoals bij de vragen 2 en 3 aangegeven wordt gebruik gemaakt van reeds beschikbare voorzieningen.

9 PRESTATIE-INDICATOREN

Regieorgaan SIA informeert zich graag over de mate waarin dit project netwerkvorming bewerkstelligt en, daarnaast, in hoeverre het aansluit bij het onderwijs. Daartoe vragen wij onderstaande prestatie-indicatoren uit.

Afhankelijk van de subsidieregeling vragen we deze prestatie-indicatoren nogmaals uit in de voortgangs- en eindrapportages. In geaggregeerde vorm geven de prestatie-indicatoren inzicht in de betrokkenheid van praktijk en onderwijs bij het onderzoek dat Regieorgaan SIA met verschillende subsidieregelingen stimuleert. Regieorgaan SIA gebruikt die informatie om regelingen beter af te stemmen op de doelstellingen.

Netwerkvorming			
Type organisatie	Aantal	Waarvan consortiumpartner	Waarvan overige betrokken partijen
Hogeschool	3	1	2
Kennisinstelling (o.a. universiteit, TO2-instituut)	1		1
Publieke instelling	1		1
Mkb-onderneming	2	1	1
Grootbedrijf	3	1	2
Koepel- of brancheorganisatie	1	1	
Beroepsvereniging			
Overig			

Onderwijs		
Onderwijssector*	Aantal docenten betrokken bij onderwijssector	Aantal studenten betrokken bij onderwijssector
Betatechniek	1-2	2-3

* Kies een (of meerdere) van toepassing zijnde onderwijssectoren:

Agro en Food Gezondheidszorg Sociale Studies Onderwijs
Bètatechniek Kunst Economie

10 REFERENTIES

- [1] CBS, "ICT, kennis en economie," Den Haag/Heerlen/Bonaire, 2021. [Online]. Available: <https://longreads.cbs.nl/ict-kennis-en-economie-2021/internetgebruik-van-huishoudens-en-personen/#:~:text=In 2015 nam 56 procent,ook vaker om te bellen>
- [2] Verbond van Verzekeraars, "Gedragscode Verzekeraars," 2011. [Online]. Available: https://www.verzekeraars.nl/overhetverbond/zelfregulering/Documents/Gedragscodes/Gedragscode_Verzekeraars_juni2011.pdf
- [3] Wetenschappelijke Raad voor Regeringsbeleid, *Opgave ai*. 2021. [Online]. Available: <https://www.rijksoverheid.nl/documenten/rapporten/2021/12/08/bijlage-bij-kamerbrief-toezending-wrr-rapport-opgave-ai-de-nieuwe-systeemtechnologie>
- [4] "Algemene Wet Gelijke Behandeling." Accessed: Apr. 26, 2024. [Online]. Available: <https://wetten.overheid.nl/jci1.3:c:BWBR0006502&z=2020-01-01&g=2020-01-01>
- [5] F. J. Zuiderveen Borges, "Digitale discriminatie en differentiatie: het recht is er nog niet klaar voor." Nijmegen, 2022.
- [6] S. Barocas, M. Hardt, and A. Narayanan, "Fairness and Machine Learning," 2019, pp. 9–35, 2020, [Online]. Available: <https://fairmlbook.org/%0Ahttps://fairmlbook.org>
- [7] B. Hutchinson and M. Mitchell, "50 Years of Test (Un)fairness: Lessons for machine learning," *FAT* 2019 - Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency*, pp. 49–58, 2019, doi: 10.1145/3287560.3287600.
- [8] C. and T. European Commission, Directorate-General for Communications Networks, "EU AI ACT: Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS." Accessed: Apr. 26, 2024. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>
- [9] Autoriteit Persoonsgegevens | Directie Coördinatie Algoritmes (DCA), "Werkagenda coördinerend algoritmetoezicht in 2024." 2024.
- [10] D. Pessach and E. Shmueli, "A Review on Fairness in Machine Learning," *ACM Comput Surv*, vol. 55, no. 3, 2022, doi: 10.1145/3494672.
- [11] College voor de Rechten van de Mens, "Advies aan Dazure B.V. over premiedifferentiatie op basis van postcode bij de Finvita overlijdensrisicoverzekering | Mensenrechten," 2014, [Online]. Available: <https://mensenrechten.nl/nl/publicatie/19173>
- [12] B. Johnson, J. Bartola, R. Angell, S. Witty, S. Giguere, and Y. Brun, "Fairkit, fairkit, on the wall, who's the fairest of them all? Supporting fairness-related decision-making," *EURO Journal on Decision Processes*, vol. 11, no. May 2022, p. 100031, 2023, doi: 10.1016/j.ejdp.2023.100031.
- [13] Gemeente Amsterdam, "The Fairness Handbook," no. May, 2022, [Online]. Available: <https://drive.google.com/file/d/18FTdKvvmnmtpAwagJlg6YwAa6KEU-Km/view>
- [14] B. Ustun, A. Spangher, and Y. Liu, "Actionable recourse in linear classification," *FAT* 2019 - Proceedings of the 2019 Conference on Fairness, Accountability, and Transparency*, pp. 10–19, 2019, doi: 10.1145/3287560.3287566.
- [15] S. Kagan, "What's Wrong with Speciesism? (Society of Applied Philosophy Annual Lecture 2015)," *J Appl Philos*, vol. 33, no. 1, pp. 1–21, 2016, doi: 10.1111/japp.12164.
- [16] K. Beck *et al.*, "The agile manifesto." 2001.

- [17] S. Leijnen, H. Aldewereld, R. van Belkom, R. Bijvank, and R. Ossewaarde, "An agile framework for trustworthy AI," *CEUR Workshop Proc*, vol. 2659, no. June, pp. 75–78, 2020.
- [18] B. Friedman, "Value-sensitive design," *Interactions*, vol. 3, no. 6, pp. 16–23, 1996, doi: 10.1145/242485.242493.
- [19] B. Friedman, D. G. Hendry, and A. Borning, "A survey of value sensitive design methods," *Foundations and Trends in Human-Computer Interaction*, vol. 11, no. 23. Now Publishers Inc, pp. 63–125, 2017. doi: 10.1561/1100000015.
- [20] A. R. Hevner, "A Three Cycle View of Design Science Research," *Scandinavian Journal of Information Systems*, vol. 19, no. 2, pp. 87–92, 2007, doi: <http://aisel.aisnet.org/sjis/vol19/iss2/4>.