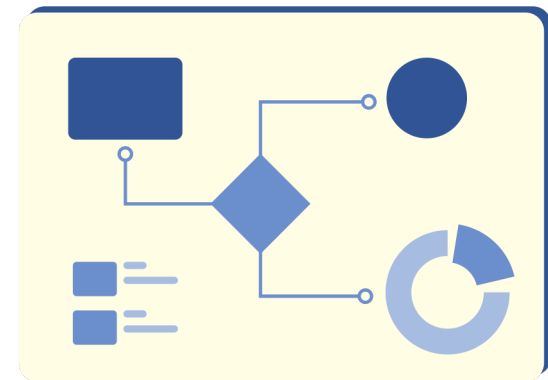


# Algoritmevalidatie

Kwalitatieve en kwantitatieve ervaringen uit de praktijk



## Overzicht seminar

1. Introductie Algorithm Audit door *Jurriaan Parie*
2. Kwantitatief – Hypothesetoets door *Floris Holstege* + Q&A
3. Kwalitatief – Algoprudentie door *Anne Meuwese* + Q&A



## Activiteiten Stichting Algorithm Audit



### Normatieve adviescommissie

Adviseren over ethische kwesties in concrete algoritmische toepassingen door samenbrengen van deliberatieve, diverse adviescommissies, met algotrudentie als resultaat



### Technische tools

Implementeren en testen van technische methoden voor biasdetectie en -mitigatie, zoals bias scan tool en synthetische data



### Kennisplatform

Samenbrengen van kennis en experts voor collectief leerproces over verantwoorde inzet van algoritmes, bijv. AI beleidsobservatorium en white papers



### Projectwerk

Ondersteuning bij specifieke vragen uit publieke en private sector over de verantwoorde inzet van algoritmes

Financieel ondersteund door

**SIDN**fonds

European  
Artificial Intelligence  
& Society Fund



Ministerie van Binnenlandse Zaken en  
Koninkrijksrelaties

Opbouwen van *publieke kennis*  
over verantwoorde AI *zonder winstoogmerk*

## Recente publicaties bieden concrete aanknopingspunten om algoritmes inhoudelijk te valideren

### Noodzaak voor algoritmevalidaties

Profilering heet hangijzer

- > Hoe te onderbouwen?
- > Omstreden, maar toch een politieke wens
- > Advies landsadvocaat en Schufa-arrest
- > Motie Bamenga-Stultiens 'algoritme APK'
- > Als grip op profilering lastig is, hoe echte AI te beteugelen?



### Vooringenomenheid voorkomen

- > Kwalitatief en kwantitatief analyse met bevindingen en aanbevelingen voor profilering door overheidsorganisaties
- > Verstuurd naar de Tweede Kamer op 01-03-2024
- > Hypothesetoets als *ex ante* validatiemethode

### Huidige soft law kaders bieden onvoldoende richting

- > Toetsingskader algoritmes, Algemene Rekenkamer (2021)
- > Impact Assessment Mensenrechten en Algoritmes (2021)
- > Algoritmeregister (2022)
- > Algoritmekader BZK (2024)



### Hoe 'algotrudentie' kan bijdragen aan een verantwoorde inzet van machine learning-algoritmes

- > Dringende behoefte om normatieve aspecten van datamodelleren openlijk te delen, bediscussiëren en te beslechten
- > Algotrudentie: concrete op casus gebaseerde en gedecentraliseerde oordeelsvorming over de verantwoorde inzet van algoritmes

1. Introductie
2. Hypothesetoets door *Floris Holstege*
3. Algoprudentie door *Anne Meuwese*



*Floris Holstege*  
Algoritmevalidator Algorithm Audit  
PhD-kandidaat Machine Learning, Universiteit van Amsterdam



The thumbnail shows a presentation slide with the Algorithm Audit logo and the University of Amsterdam logo. The slide title is "Vooringenomenheid voorkomen" (Prevention of bias). The content includes "Aanbevelingen voor risicoprofilering in het Controle Uitwonendenbeurs-proces: een kwantitatieve en kwalitatieve analyse" (Recommendations for risk profiling in the Control of Outgoing Students Scholarship process: a quantitative and qualitative analysis). The date "Februari 2024" is visible at the bottom.

## Aannames in risicoprofilerings-algoritmes moeten gevalideerd worden

Elk algoritme bestaat uit aannames. **Wanneer deze aannames niet zorgvuldig onderbouwd zijn, kan dit leiden tot onbedoelde discriminatie**

Deze onderbouwing kan voortkomen uit:

- > Statistische onderzoek (via o.a. hypothesetoets)
- > Gedragwetenschappelijk onderzoek
- > Politiek-bestuurlijke afweging

Noodzaak voor het valideren van algoritmes conform normen anno 2024

## Case study: risicoprofilering in controle uitwonendenbeurs (CUB)-proces bij DUO

### Chronologie

- > 2012: Start gebruik risicoprofiel. Onderscheid op basis van onderwijsvorm, leeftijd en afstand tot ouder(s). Wettelijke grondslag in Wet studiefinanciering 2000.
- > 2014: Willekeurige steekproef: onrechtmatig gebruik 3,6%  
Controle aan de hand van risicoprofiel: onrechtmatig gebruik 38,9%
- > 2015: Invoering leenstelsel. Alleen mbo-studenten nog recht op uitwonendenbeurs
- > 2018: Herevaluatie naar aanleiding van KOT-affaire. Geen aanleiding om profiel + onderbouwing aan te passen
- > 2023: Kamervragen gesteld over het CUB-proces naar aanleiding van media-aandacht. Gebruik risicoprofiel stilgelegd
- > 2023: Herintroductie leenstelsel voor mbo-, hbo- en wo-studenten
- > 2024: Excuses minister voor indirecte discriminatie

Uitwonendenbeurs –  
Financiële ondersteuning van  
studenten die op kamers  
wonen (€163,40 per maand)

Woensdag 21 juni 2023, 06:14

**Studenten met migratieachtergrond opvallend vaak beschuldigd van fraude, minister wil systeem grondig nagaan**

Bron: NOS

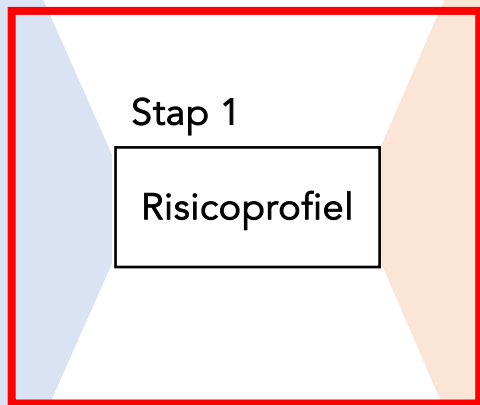
Vrijdag 1 maart, 14:11 • Aangepast vrijdag 1 maart, 15:37

**Dijkgraaf: onbedoeld discriminatie bij controle studenten**

Bron: NOS



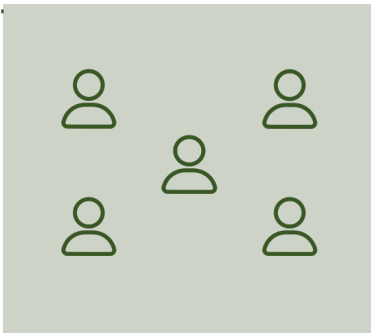
### Stap 0 – Brondata



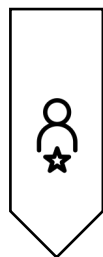
risicoscores



### Stap 2 – Preselectie huisbezoek

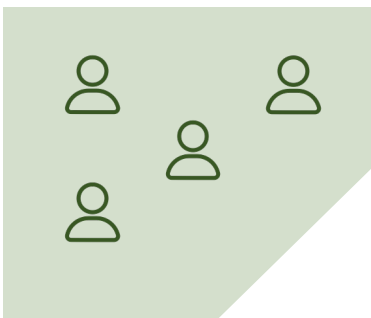


Preselectie-populatie

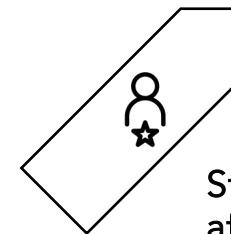


Stap 3 – Handmatige check DUO

### Stap 4 – Extern: Huisbezoek

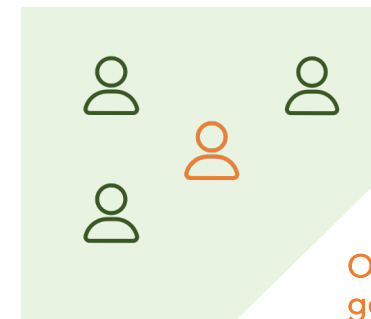


Huisbezoek-populatie



Stap 5 – Terugkoppeling, afhandeling en vervolgstappen

### Stap 6 – Bezwaar en beroep



Onrechtmatig-gebruik-populatie

## Werking risicoprofiel

Onderscheid op basis van:

- > Opleidingsvorm
- > Leeftijd
- > Afstand tot ouder(s)

Resultaat:

- > Risicoscore tussen 0-144
- > Opdeling in risicocategorie 1 (zeer hoog) tot en met 5 (zeer laag), 6 (onbekend)

## Aannames

De kans op misbruik uitwonendenbeurs is groter bij:

- > Jongere studenten
- > Mbo- of hbo-studenten in tegenstelling tot wo-studenten
- > Studenten die dichterbij ouder(s) wonen
- > Categorie 1 vs. 2 vs. 3 vs. 4 vs. 5 vs. 6.

RF		Risico Factor		RC				Risico Categorie			
Onderwijs	Factor			Codering	Omschrijving	Ondergrens	Bovengrens				
Bol 1-2	1,2			1	Zeer hoog risico	80	144				
Bol 3-4	1,1			2	Hoog risico	60	79				
HBO	1,0			3	Gemiddeld risico	40	59				
WO	0,8			4	Laag risico	20	39				
*Bepaling risico (-punten)				5	Zeer laag risico	1	19				
				6	Onbekend risico	0	0				

Aantal meters tot ouders	Leeftijd	Risico	Bol 1-2	Bol 3-4	HBO	WO
0	17-18	120	144	132	120	96
0	19-20	110	132	121	110	88
0	21-22	105	126	116	105	84
0	23-24	100	120	110	100	80
0	25-40	80	96	88	80	64
1-1000	17-18	100	120	110	100	80
1-1000	19-20	95	114	105	95	76
1-1000	21-22	85	102	94	85	68
1-1000	23-24	75	90	83	75	60
1-1000	25-40	65	78	72	65	52
1001-2000	17-18	95	114	105	95	76
1001-2000	19-20	85	102	94	85	68
1001-2000	21-22	75	90	83	75	60
1001-2000	23-24	65	78	72	65	52
1001-2000	25-40	60	72	66	60	48
2001-5000	17-18	85	102	94	85	68
2001-5000	19-20	75	90	83	75	60
2001-5000	21-22	65	78	72	65	52
2001-5000	23-24	55	66	61	55	44
2001-5000	25-40	45	54	50	45	36
5001-10000	17-18	75	90	83	75	60
5001-10000	19-20	65	78	72	65	52
5001-10000	21-22	55	66	61	55	44
5001-10000	23-24	45	54	50	45	36
5001-10000	25-40	35	42	39	35	28
10001-20000	17-18	50	60	55	50	40
10001-20000	19-20	40	48	44	40	32
10001-20000	21-22	30	36	33	30	24
10001-20000	23-24	25	30	28	25	20
10001-20000	25-40	20	24	22	20	16
20001-50000	17-18	35	42	39	35	28
20001-50000	19-20	25	30	28	25	20
20001-50000	21-22	20	24	22	20	16
20001-50000	23-24	15	18	17	15	12
20001-50000	25-40	10	12	11	10	8
50001-250000	17-18	20	24	22	20	16
50001-250000	19-20	20	24	22	20	16
50001-250000	21-22	15	18	17	15	12
50001-250000	23-24	10	12	11	10	8
50001-250000	25-40	5	6	6	5	4
Onbekend	17-18	0	0	0	0	0
Onbekend	19-20	0	0	0	0	0
Onbekend	21-22	0	0	0	0	0
Onbekend	23-24	0	0	0	0	0
Onbekend	25-40	0	0	0	0	0

## Aannames kunnen getoetst worden met behulp van een statistische hypothesetoets

$H_0$ : kans onrechtmatig gebruik bij mbo-studenten (niveau 3-4) =  
kans op onrechtmatig gebruik bij hbo-studenten

$H_A$ : kans onrechtmatig gebruik bij mbo-studenten (niveau 3-4) >  
kans op onrechtmatig gebruik bij hbo-studenten

- > Hypotheses moeten idealiter worden getoetst op basis van aselechte steekproef
- > Statistische significantie bepaald met Z-toets en aselechte steekproef uit 2014 (n=387)

## Beperkt statistisch bewijs voor verband tussen onderwijsvorm en onrechtmatig gebruik uitwonendenbeurs

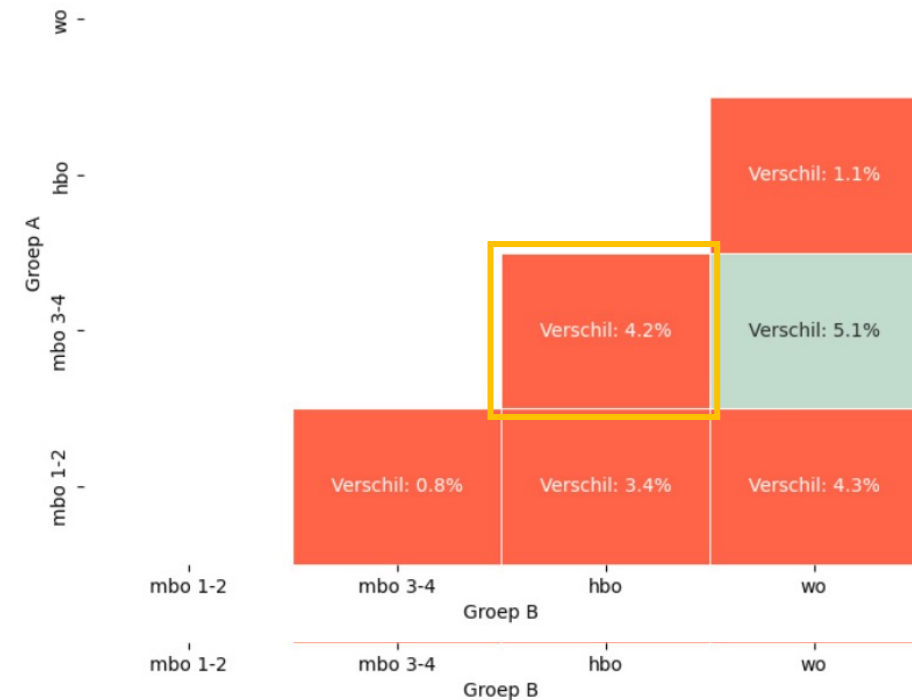
2014	Grootte groep	#onrechtmatig	Percentage
mbo 1-2	15	1	6,7%
mbo 3-4	53	4	7,5%
hbo	150	5	3,3%
wo	164	4	2,4%
<b>Totaal</b>	<b>387</b>	<b>14</b>	<b>3,6%</b>

### Aannames Z-toets

Onderhevig aan gebruikelijke aannames rondom statistische toetsen, zoals:

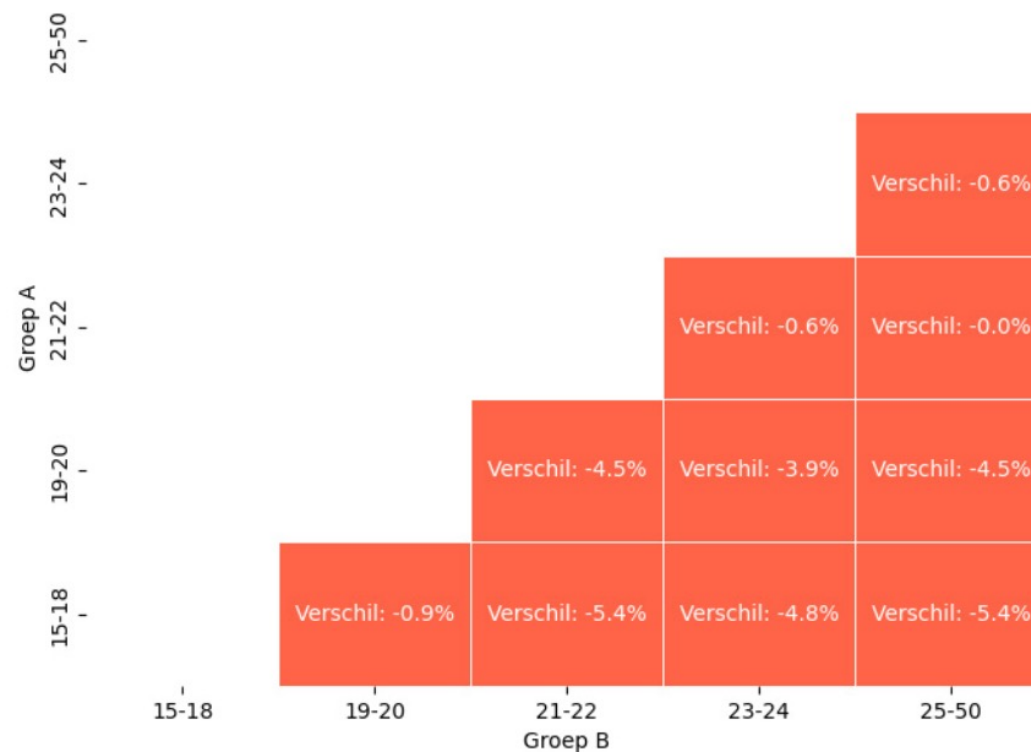
- > Steekproef moet groot genoeg zijn (voor beide groepen)
- > Onafhankelijk en identiek verdeelde observaties

Aselecte steekproef 2014: Statistisch significant onderscheid in groen (n=387)  
(resultaten gebaseerd op eenzijdige Z-toets)



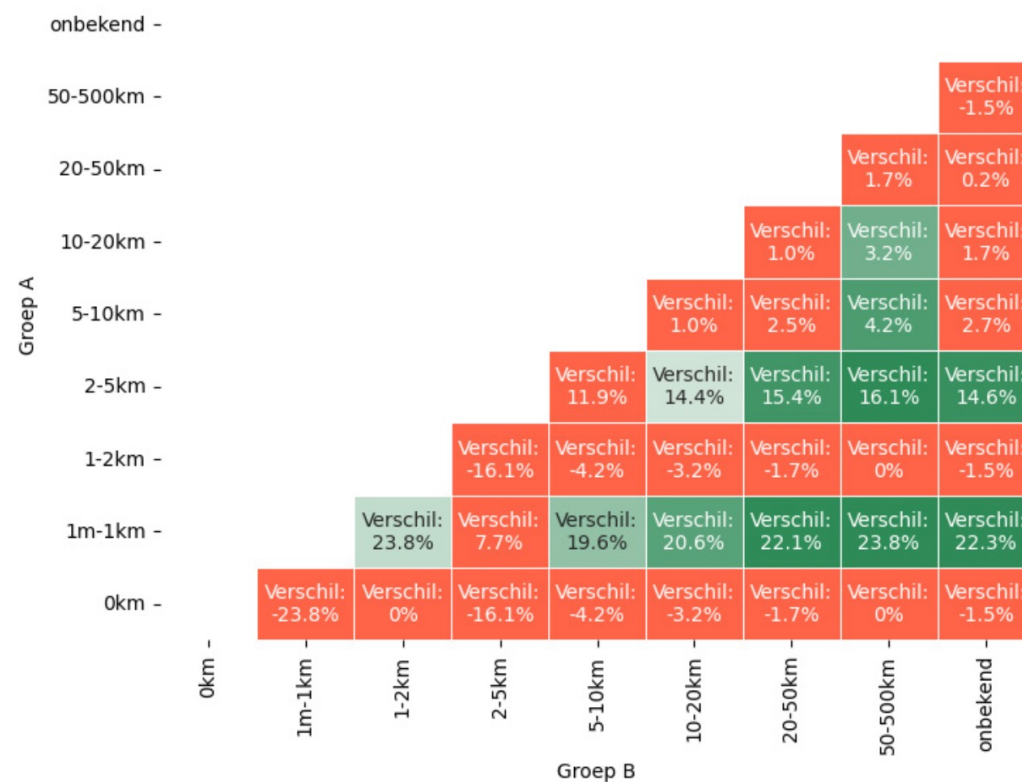
## Geen statistisch bewijs voor verband tussen leeftijd en onrechtmatig gebruik uitwonendenbeurs

Aselecte steekproef 2014: Statistisch significant onderscheid in groen (n=387)  
(resultaten gebaseerd op eenzijdige Z-toets)



## Statistisch bewijs voor verband tussen afstand tot ouder(s) en onrechtmatig gebruik uitwonendenbeurs

Aselecte steekproef 2014: Statistisch significant onderscheid in groen (n=387)  
(resultaten gebaseerd op eenzijdige Z-toets)



## Geen statistisch bewijs voor 6 risicocategorieën, wel voor 2 risicocategorieën

Groep	% Misbruik	Aantal Misbruik	Totaal Aantal
1 - Zeer hoog	10,0%	2	20
2 - Hoog	17,2%	5	29
3 - Midden	13,2%	5	38
4 - Laag	0,0%	0	106
5 - Zeer laag	0,8%	1	125
6 - Onbekend	1,4%	1	69
Zeer hoog (1) + Hoog (2)	14,3%	7	49
Midden (3) + Laag (4) + Zeer laag (5)	2,2%	6	269
Midden (3) + Laag (4) + Zeer laag (5) + Onbekend (6)	2,1%	7	338
Vershil Groep 1 + 2 & Groep 3 + 4 + 5	12,1%	-	-
Vershil Groep 1 + 2 & Groep 3 + 4 + 5 + 6	12,2%	-	-
Vershil Zeer hoog (1) & hoog (2)	-7,2%	-	-
Vershil Hoog (2) & Midden (3)	4,1%	-	-
Vershil Midden (3) & Laag (4)	13,2%	-	-
Vershil Laag (4) & Zeer laag (5)	-0,8%	-	-

## Bij herontwerp CUB-risicoprofiel moet gemaakt statistisch gevalideerd worden

**Bevinding 2** – Uit de aselechte steekproeven uit 2014 en 2017 volgt onvoldoende statistisch verband tussen de selectiecriteria onderwijsvorm en leeftijd en onrechtmatig gebruik van de uitwonendenbeurs. Voor specifieke categorieën binnen het selectie criterium afstand is wel voldoende verband met onrechtmatigheid gevonden. Voor de opdeling in zes risicocategorieën is onvoldoende statistische onderbouwing gevonden ten opzichte van een binaire risicoclassificatie.



Elk algoritme bestaat uit aannames. Wanneer deze aannames niet zorgvuldig statistisch gecontroleerd worden, kan dit mogelijk leiden tot onbedoelde discriminatie



# Q&A



1. Introductie
2. Hypothesetoets door *Floris Holstege*
3. **Algoprudentie door *Anne Meuwese***



*Anne Meuwese*  
Raad van Advies Algorithm Audit  
Hoogleraar Staats- en bestuursrecht Universiteit Leiden



## Voortdurende onzekerheid over gebruik van algoritmes door overheidsorganisaties

- > Veel normen, weinig houvast
- > De trias hapert
- > Onzekerheid wegnemen door 'contextualisering' van wettelijke kaders

### Definitie algotrudentie

Concrete op casus gebaseerde en gedecentraliseerde oordeelsvorming over de verantwoorde inzet van algoritmes

### Soft law kaders

- > Toetsingskader algoritmes, Algemene Rekenkamer (2021)
- > Impact Assessment Mensenrechten en Algoritmes (2021)
- > Algotrudentregister (2022)
- > Algotrudentkader BZK (2024)

### Wettelijke kaders

- > AVG
- > AI-Verordening
- > Algemene wet bestuursrecht
- > Bijzondere wetten (e.g. Wsf 2000)

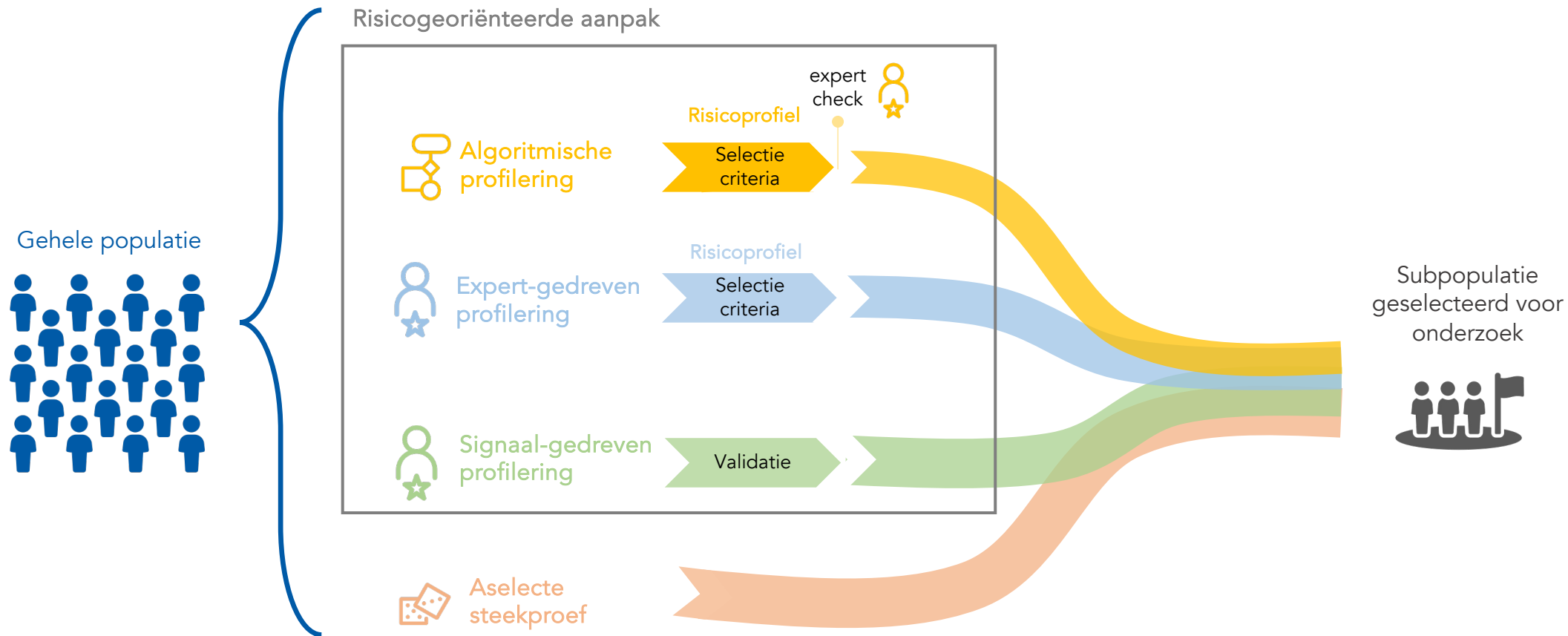
### Institutionele impasse

- > Rechter
- > Parlement
- > Toezichthouder

### Waarom algotrudentie?

- > Creëert draagvlak bij verschillende belanghebbenden
- > Deliberatieve, inclusieve en transparante oordeelsvorming over normatieve kwesties
- > Publieke kennisopbouw over interpretatie van open juridische normen

# Keuzes maken rond selectiemethoden onvermijdelijk



## Selectiemethoden zijn o.a. onderhevig aan de algemene beginselen van bestuur

### Motiveringsbeginsel

#### Openstaande vragen:

- > Geen concrete uitlegbaarheidsvereisten voor ML
- > Juridische roep om uitleg in 'natuurlijke taal': hoe en in hoeverre?
- > Hoe uitlegbaar zijn ML ensemble methoden? Zoals xgboost, explainable boosting → black box

### Zorgvuldigheidsbeginsel

#### Openstaande vragen:

- > Is ML een geschikte methode? Zo ja, welk type ML?
- > Zijn alle relevante data bekend?
- > Welke inputvariabelen zijn (on)geschikt om door het algoritme als profileringscriteria te worden overwogen?
- > Moet het voorspellende vermogen van een variabele een rol spelen bij selectie?

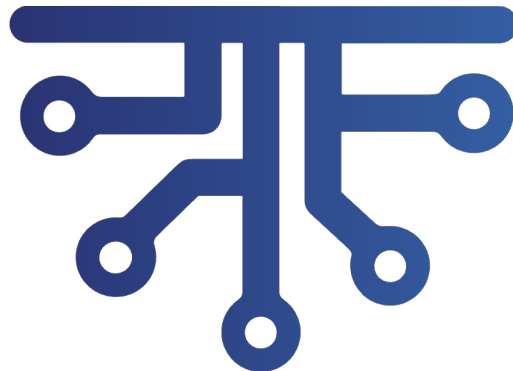
### > Beginsel van *fair play*

#### Openstaande vragen:

- > Verplichting bias detectie en -mitigatie?
- > Hoe en in hoeverre (proxy)discriminatie te voorkomen?



*algoprudentie*



## Voorbeeld uit Rotterdam: risicoprofilering heronderzoek bijstandsuitkering



Inwoners onderworpen aan het algoritme



Afgevaardigden van betroffen groepen



Gemeente Rotterdam



Onderzoeksjournalisten



Gemeentelijke instanties (Rotterdamse Rekenkamer, Rotterdamse ombudsman)



Juridische experts en wetenschappers



### Normatieve adviescommissie



Maarten van Asten, wethouder Financiën, Digitalisering en Evenementen, gemeente Tilburg



Munish Ramlal  
Ombudsman metropoolregio Amsterdam



Abderrahman Al Aazani,  
Onderzoeker Rotterdamse Ombudsman



Francien Dechesne, Associate Professor Law and Digital Technologies, Universiteit Leiden



Oskar Gstrein, Assistant Professor Governance and Innovation, Rijksuniversiteit Groningen

# De normatieve dimensie van expert- en algoritme-gedreven profilering lijken op elkaar

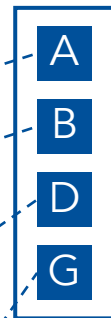
## Stap 1

Beschikbare variabelen in database



## Stap 2

Variabeleselectie



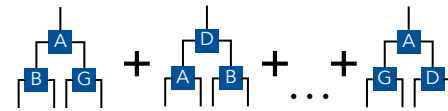
✓  
Geschikt onder bepaalde voorwaarden

✗  
Ongeschikt

## Stap 3

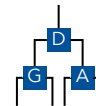
Profielsamenstelling

Algoritme-gedreven profilering → profile



Zoals explainable boosting en extreme gradient boosting

Expert-gedreven profilering → profiel



## Stap 4

Expert check op uiteindelijke risicoprofiel

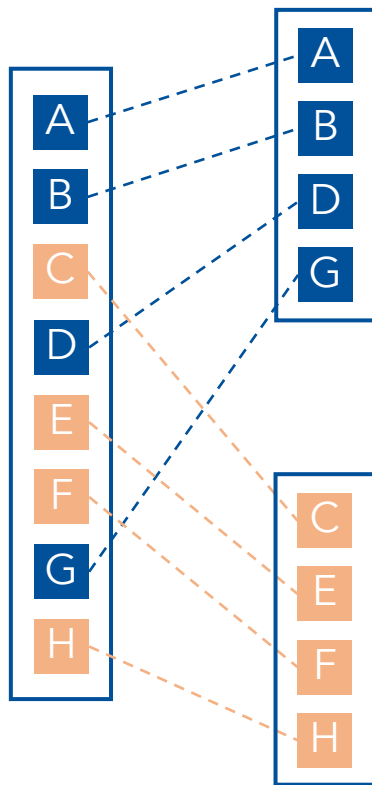




# Algorudentie voorziet in flexibele en concrete normen

## Stap 1

Beschikbare variabelen in database



## Stap 2

Variabeleselectie

✓  
Geschikt onder bepaalde voorwaarden

✗  
Ongeschikt

### Niet te gebruiken variabelen

Postcode, wijk	⚡
Geslacht, gender	⊖
Reden voor afspraak met gemeente (jaarlijks gesprek, intake)	?
Type contact (mailen, bellen, appen, post)	👉
Laaggeletterdheid	⚡
ADHD	⊖
Bezoekt psycholoog	⊖
Aantal kinderen	👉
Sector (werk)ervaring (horeca, bouw, logistiek)	🔄
Assertiviteit	🔗 🗒️
Professioneel voorkomen	🗒️

### Wel te gebruiken variabelen

Leeftijd	🔗	👤
No show bij afspraak met gemeente	🔗	👤
Reminders voor aanleveren van informatie	🔗	👤
Participatie in traject naar werk (training, werkplaats, maatschappelijke taak)	🔗	👤 ?
Type woonvorm (met huisgenoot, met partner)	🔗	👤
Kostendelersnorm	🔗	👤

### Legenda

⊖	Wettelijk verboden	🗒️	Subjectief
🔗	Inhoudelijke verband met doel	🔄	Veranderlijk
👉	Geen inhoudelijk verband met doel	👤	Beheersbare risico's
?	Onduidelijke variabele	⚡	Proxy-discriminatie

## Rotterdamoordeel: context-afhankelijk én generaliseerbaar

- > Algoritme-gedreven besluitvorming kan onder strikte voorwaarden worden ingezet
- > Heronderzoek moet meer op dienstverlening en minder op wantrouwen zijn gebaseerd
- > Met het oog op uitlegbaarheidsvereisten is een xgb-algoritme ongeschikt
- > Combineer algoritme-gedreven profilering met expert-gedreven profilering en aselechte steekproeven om tunnelvisie en feedbackloops te vermijden
- > Lijst van (on)geschikte selectiecriteria

## Samenvattend: algoprudentie

### Leren & harmonizeren

- > Jaagt collectieve leerproces aan over de verantwoorde inzet van algoritmes
- > Harmoniseert interpretatie van wet- en regelgeving in de context van algoritme-gedreven besluitvorming

### Bevragen & bekritisieren

- > Beslecht normatieve keuzes in democratisch zicht
- > *Ex ante* mechanisme voor AI validatie
- > Definieert een *state-of-the-art*, ie. normstelling

### Inclusie & participatie

- > Betreft verschillende belanghebbenden bij het ontwerpen van AI
- > Nederlands antwoord hoe AI op een democratische manier ontwikkeld en ingezet kan worden

# Q&A



## Wat u kunt doen om algoritrudentie en Stichting Algorithm Audit te helpen ontwikkelen

1. Reageer op een normatief oordeel en deel uitkomsten met andere experts

<https://algorithmaudit.eu/nl/algorithmprudence>

2. Samenwerking middels projectwerk of dien een case in voor review

<https://algorithmaudit.eu/algorithmprudence/submit-a-case/>

3. Word onderdeel van een normatieve adviescommissie

[info@algorithmaudit.eu](mailto:info@algorithmaudit.eu)



[www.algorithmaudit.eu](http://www.algorithmaudit.eu)



[info@algorithmaudit.eu](mailto:info@algorithmaudit.eu)



<https://www.linkedin.com/company/algorithm-audit/>



<https://github.com/NGO-Algorithm-Audit>



Stichting Algorithm Audit is bij de Kamer van Koophandel  
geregistreerd onder nummer 83979212