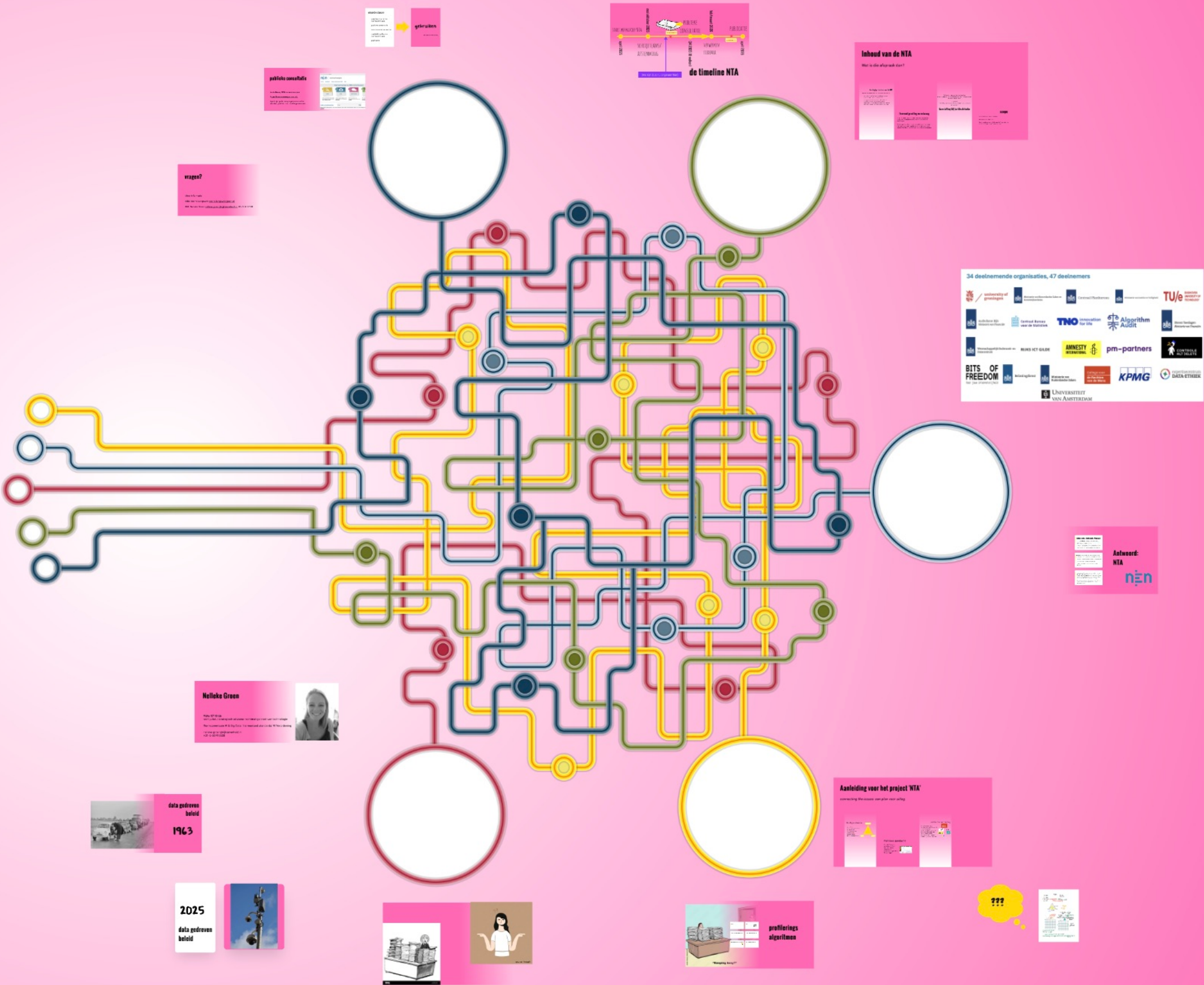


algoritmen gebruiken om te profileren en het onderscheid dat ze maken...

proces en methode voor de het ontwikkelen van
rechtmatige profileringsalgoritmes



Nelleke Groen

Rijks ICT Gilde

tech jurist / strategisch adviseur rechtmatige inzet van technologie

Normcommissie AI & Big Data: 'harmonised standards' AI Verordening

nelleke.groen@rijksoverheid.nl

+31-6-50419530



data gedreven beleid

1963



2025

data gedreven beleid





source: freepik



wens

tienduizenden individuele dossiers elk jaar
enkele personen voor het controleren
belofte van profileren met algoritmen:
'met zo min mogelijk dossierchecks, zoveel mogelijk issues vinden'

maar...

!!!

1: data is nooit onbevooroordeeld

'noot onbevooroordeeld' = 'vooringenomen'
Nederland dubieuze recordhouder...

2: algoritmen selecteren op..?

... op basis van bepaalde factoren of een combinatie ervan, maar:
- welke factoren?
- waarom zijn dit goede criteria om op te selecteren?
- waarom is de conclusie dat nu juist jouw dossier gekozen is?

Nederland: dubieuze recordhouder

meerdere heel gevoelige cases,
waarbij na beter kijken (lees:
stevig onderzoek doen) bleek:
het algoritme discrimineert.



maar ook: best practices

"What not to do?"

profilerings algoritmen

“Keeping busy?”



Prezi

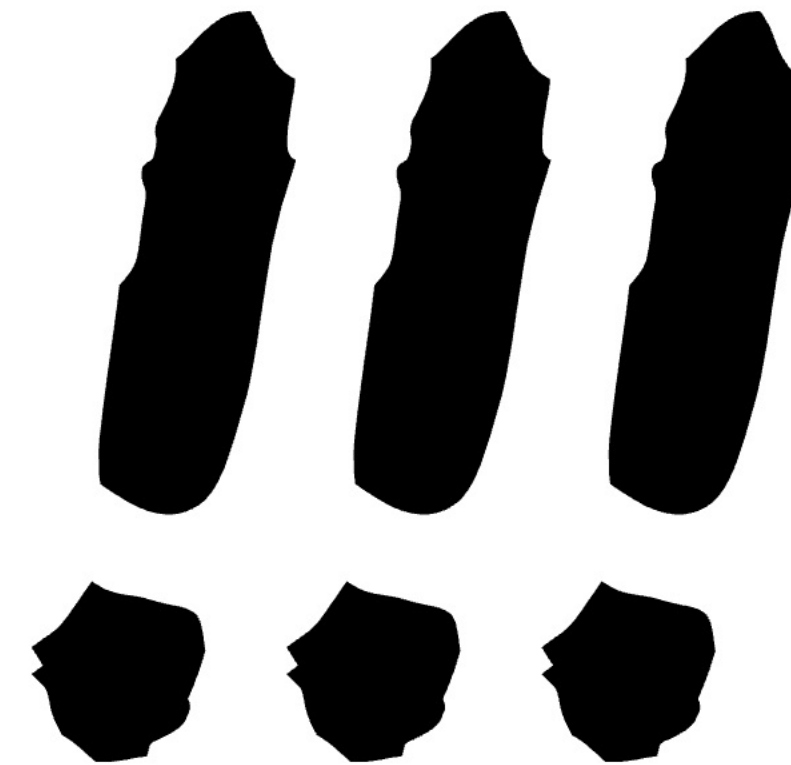
wens

tienduizenden individuele dossiers elk jaar
enkele personen voor het controleren

belofte van profileren met algoritmen:

'met zo min mogelijk dossierchecks, zoveel mogelijk issues vinden'

maar...



1: data is nooit onbevooroordeeld

'nooit onbevooroordeeld' = 'vooringenomen'?

Nederland dubieuze recordhouder...

2: algoritmen selecteren op..?

... op basis van bepaalde factoren of een combinatie ervan, maar:

- welke factoren?
- waarom zijn dit goede criteria om op te selecteren?
- waarom is de conclusie dat nu juist *jouw* dossier gekozen is?

Nederland: dubieuze recordhouder

meerdere heel gevoelige cases,
waarbij na beter kijken (*lees:
stevig onderzoek doen*) bleek:



het algoritme discrimineert.

maar ook: best practices

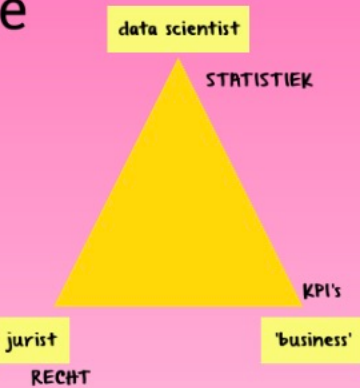
"What not to do?"

Aanleiding voor het project 'NTA'

connecting the issues: een plan voor uitleg

Huidige situatie

behoefte bij uitvoeringsorganisaties aan duidelijke richtlijnen voor de verantwoorde toepassing van profileringsalgoritmes



Politieke aandacht

Er is aanhoudende politieke aandacht voor de inzet en validatie van profileringsalgoritmes: **Kamervragen**.



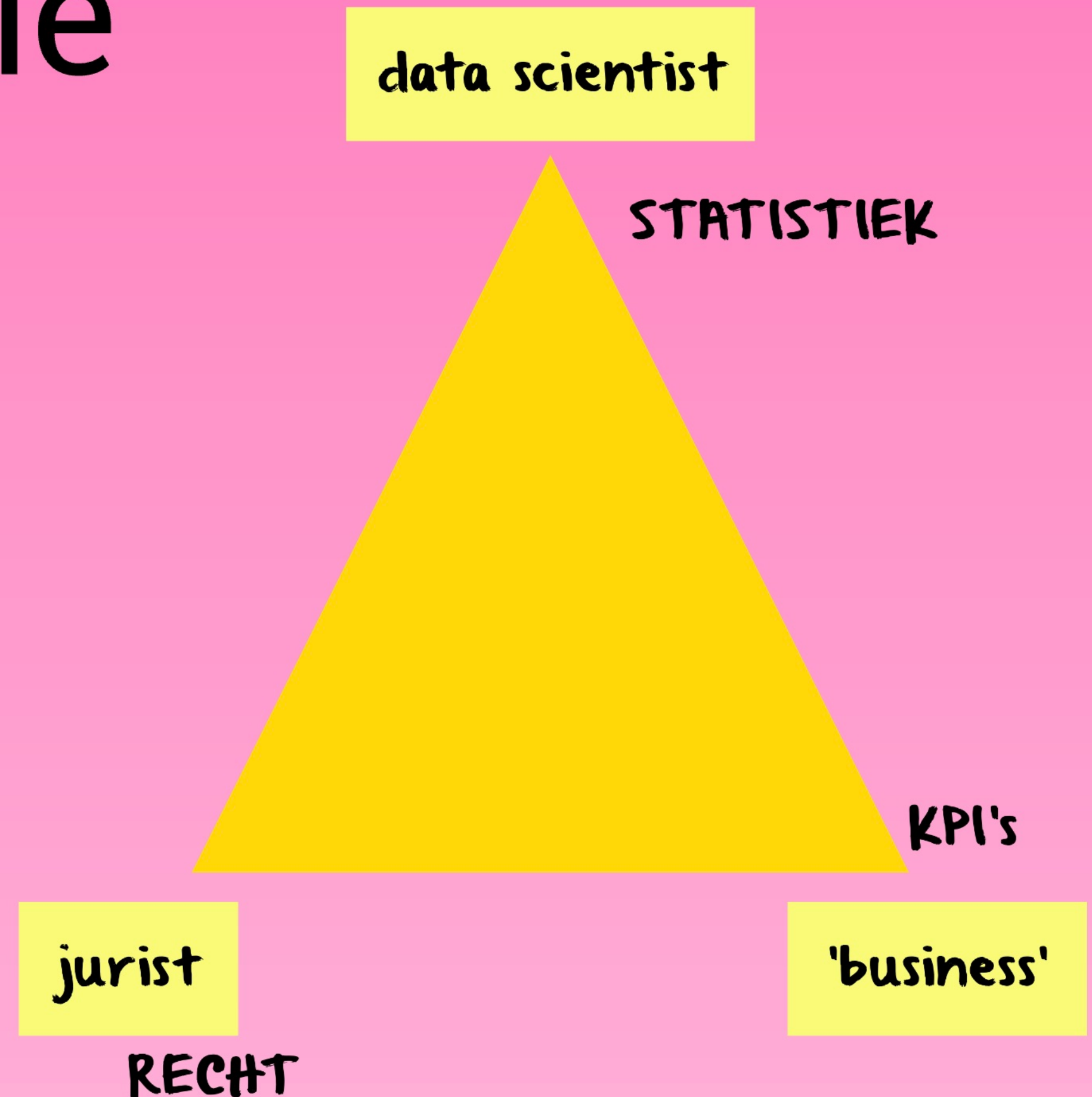
Juridische normuitleg

Er is discussie over juridische normuitleg, en de vraag of indirect onderscheid op grond van etniciteit überhaupt ooit gerechtvaardigd kan worden (in deze context: profilering door de overheid)



Huidige situatie

behoefte bij uitvoeringsorganisaties aan duidelijke richtlijnen voor de verantwoorde toepassing van profileringsalgoritmes



Politieke aandacht

Er is aanhoudende politieke aandacht voor de inzet en validatie van profileringsalgoritmes: **Kamervragen.**



Juridische normuitleg

Er is discussie over juridische normuitleg, en de vraag of indirect onderscheid op grond van etniciteit überhaupt ooit gerechtvaardigd kan worden (in deze context: profilering door de overheid)

College voor
de Rechten
van de Mens



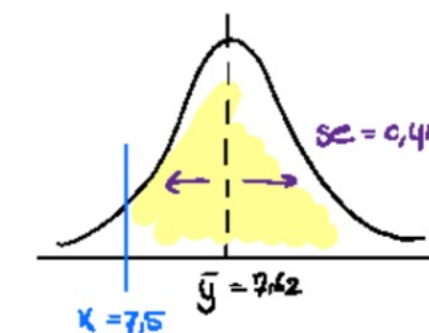
CONTROLE
ALT DELETE

???

$n = 502$
 $\bar{x} = 7,62$
 $\mu_0 = 7,5$

z-toets
 $H_0 = 7,5$
 $H_1/H_a > 7,5$

! $se = \sigma/\sqrt{n} \Rightarrow$ steekproef = $0,922/\sqrt{502} = 0,411$
 centrale limitstelling



formules
 $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

$z = \frac{7,5 - 7,62}{0,411} = -0,292$

$1 - (-0,29) = 1,29$
 want je wilt $\geq$ weten en er wordt door de tabel altijd naar $\leq$ gekeken

Of het getal zoek je op in tabel a

je kijkt als naar deze "kant"

Table with 4 columns: z, .00, .01, .02, .03. The table contains cumulative probabilities for z values from -3.4 to 3.4.

Table with 4 columns: z, .00, .01, .02, .03. The table contains cumulative probabilities for z values from -3.4 to 3.4.

$\alpha = 0,05$
 is gebruikelijk

je kijkt naar beide decimalen
 $P(z > 1,29) = 0,1015$

$0,1015 > 0,05$; dus H_0 wordt niet verworpen
 Het is aanneemelijk om te zeggen dat de score niet significant hoger is dan 7,5

Nederlandse Technische Afspraak

- **nēn** normalisatie mensen uit het veld, met de expertise en de kennis, komen bij elkaar om vast te stellen wat de 'state of the art' is
- **technische afspraak** beperkte scope voor een bepaald onderwerp, om stappen richting standaarden te zetten
- **in dit geval** afspraken over de statistische methoden om te bepalen welk onderscheid een algoritme maakt.

in dit geval afspraken over de statistische methoden om te bepalen welk onderscheid een algoritme maakt:

- op welke gronden baseert het algoritme de selectie?
- en zitten daar discriminatiegronden bij?
- wordt er dan direct of indirect onderscheid gemaakt?

Als de *statistici* het eens kunnen worden over **hoe deze feitelijke informatie uit het systeem te halen**, en **hoe deze te vertalen** naar taal die de rest van de wereld ook begrijpt, dan kun je het gesprek hebben over welk onderscheid rechtmatig zou zijn.

En: als je ongewenst onderscheid ontdekt, dan kun je het - in ieder geval in de machine - eruit proberen te halen.

Antwoord: NTA

nēn

Nederlandse Technische Afspraak

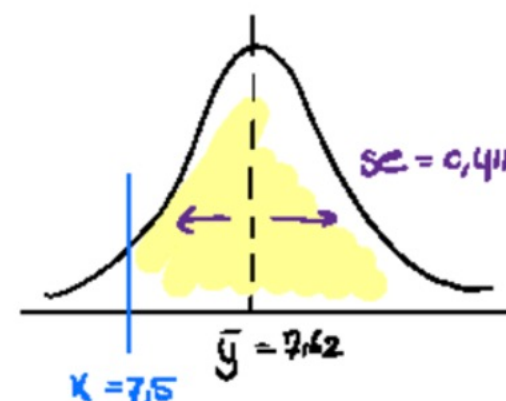
- **nēn** normalisatie mensen uit het veld, met de expertise en de kennis, komen bij elkaar om vast te stellen wat de 'state of the art' is
- **technische afspraak** beperkte scope voor een bepaald onderwerp, om stappen richting standaarden te zetten
- **in dit geval** afspraken over de statistische methoden om te bepalen welk onderscheid een algoritme maakt.

in dit geval afspraken over de statistische methoden om te bepalen welk onderscheid een algoritme maakt:

- op welke gronden baseert het algoritme de selectie?
- en zitten daar discriminatiegronden bij?
- wordt er dan direct of indirect onderscheid gemaakt?

$n = 502$
 $\bar{x} = 7,62$
 $H_0 = 7,5$
 $H_1/H_a > 7,5$

$se = \sigma/\sqrt{n} \Rightarrow \text{steekproef} = 0,922/\sqrt{502} = 0,0411$
 centrale limietstelling



formules
 $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

$z = \frac{7,5 - 7,62}{0,411} = -0,292$

$1 - (-0,29) = 1,29$
 want je wilt $\geq$ weten en er wordt door de tabel altijd naar $\leq$ gekeken



Table 1: Standard Normal Probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.3	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.2	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.1	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.9	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.6	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.5	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.4	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.3	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.2	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.1	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-2.0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.9	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.6	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.5	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.4	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.3	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.2	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.1	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-1.0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-0.9	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-0.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-0.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-0.6	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-0.5	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-0.4	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-0.3	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-0.2	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-0.1	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
0.0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001

Table 2: Standard Normal Probabilities (continued)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.2	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.3	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.4	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.5	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.2	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.3	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.4	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.5	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.2	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.3	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.4	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.5	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
3.0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

$\alpha = 0,05$ is gebruikelijk

je kijkt naar beide decimalen
 $P(z > 1,29) = 0,1015$
 $0,1015 > 0,05$; dus H_0 wordt niet verworpen
 Het is onnemenlijk om te zeggen dat de score niet significant hoger is dan 7,5

Als de *statistici* het eens kunnen worden over **hoe deze feitelijke informatie uit het systeem te halen**, en **hoe deze te vertalen** naar taal die de rest van de wereld ook begrijp, dan kun je het gesprek hebben over welk onderscheid rechtmatig zou zijn.

En: als je ongewenst onderscheid ontdekt, dan kun je het - in ieder geval in de machine - eruit proberen te halen.

34 deelnemende organisaties, 47 deelnemers



university of
 groningen



Ministerie van Binnenlandse Zaken en
Koninkrijksrelaties



Centraal Planbureau



Ministerie van Justitie en Veiligheid

TU/e EINDHOVEN
UNIVERSITY OF
TECHNOLOGY



Auditdienst Rijk
Ministerie van Financiën



Centraal Bureau
voor de Statistiek

TNO innovation
for life

 **Algorithm
Audit**



Dienst Toeslagen
Ministerie van Financiën



Wetenschappelijk Onderzoek- en
Datacentrum

RIJKS ICT GILDE

**AMNESTY
INTERNATIONAL**



pm-partners



**CONTROLE
ALT DELETE**

**BITS OF
FREEDOM**

Voor jouw internetvrijheid



Belastingdienst



Ministerie van
Buitenlandse Zaken

College voor
de Rechten
van de Mens

KPMG



expertisecentrum
DATA-ETHIEK



UNIVERSITEIT
VAN AMSTERDAM

Inhoud van de NTA

Wat is die afspraak dan?

Voorlopige structuur van de NTA

volgens de **levenscyclus** van algoritmische profiling

- de initiële ontwikkeling, waarbij vereisten en doelstellingen worden vastgesteld;
- de implementatiefase, waarin de richtlijnen worden toegepast binnen de betrokken organisaties en de nodige aanpassingen worden gemaakt op basis van feedback en evaluaties.

Eenvoudige uitleg en nulvraag

Er wordt steeds gezocht naar de meest eenvoudige uitleg om de toegankelijkheid en bruikbaarheid te waarborgen.

Belangrijk onderdeel van de methode is een 'nulvraag' die organisaties aanmoedigt om na te denken over de **wenselijkheid** van profiling en mogelijke **alternatieven**.

Referentie voor juridische afweging:
Toetsingskader risicoprofiling van het College voor de Rechten van de Mens.

Zoektocht:
verbinding 'statistische feitelijkheden' en 'juridische afwegingen'

Aansluiting bij juridisch kader

scope

- (risico)profiling door de overheid
- met behulp van algoritmen
- die niet reeds gereguleerd zijn door de AI Verordening (omdat het hoog-risico AI-systemen zijn)

Voorlopige structuur van de NTA

volgens de **levenscyclus** van algoritmische profiling

- de initiële ontwikkeling, waarbij vereisten en doelstellingen worden vastgesteld;
- de implementatiefase, waarin de richtlijnen worden toegepast binnen de betrokken organisaties en de nodige aanpassingen worden gemaakt op basis van feedback en evaluaties.

Eenvoudige uitleg en nulvraag

Er wordt steeds gezocht naar de meest eenvoudige uitleg om de toegankelijkheid en bruikbaarheid te waarborgen.

Belangrijk onderdeel van de methode is een 'nulvraag' die organisaties aanmoedigt om na te denken over de **wenselijkheid** van profilering en mogelijke **alternatieven**.

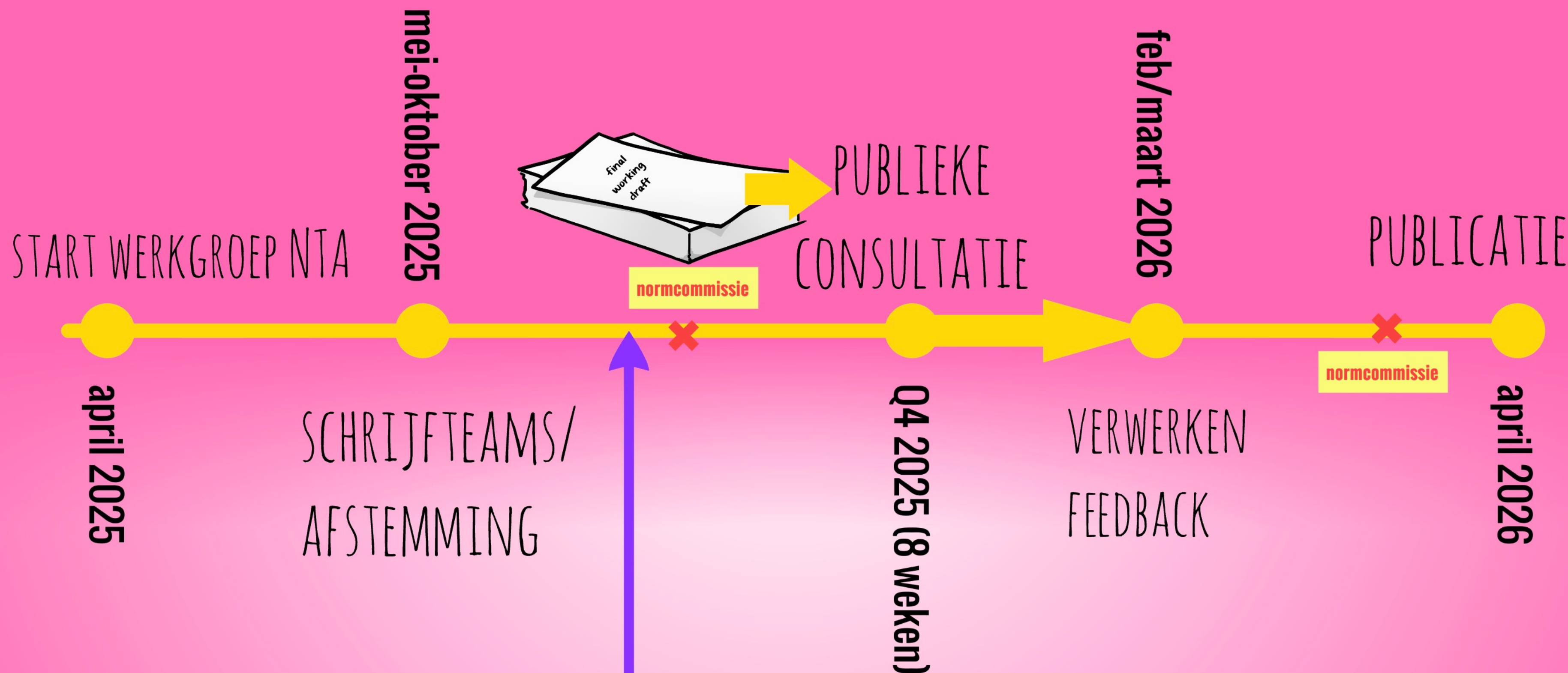
Referentie voor juridische afweging:
Toetsingskader risicoprofilering van het College voor de
Rechten van de Mens.

Zoektocht:
verbinding 'statistische feitelijkheden' en 'juridische
afwegingen'

Aansluiting bij juridisch kader

scope

- (risico)profilering door de overheid
- met behulp van algoritmen
- die niet reeds gereguleerd zijn door de AI Verordening (omdat het hoog-risico AI-systemen zijn)



(we zijn dus nu ongeveer hier)

de timeline NTA

volgende stappen

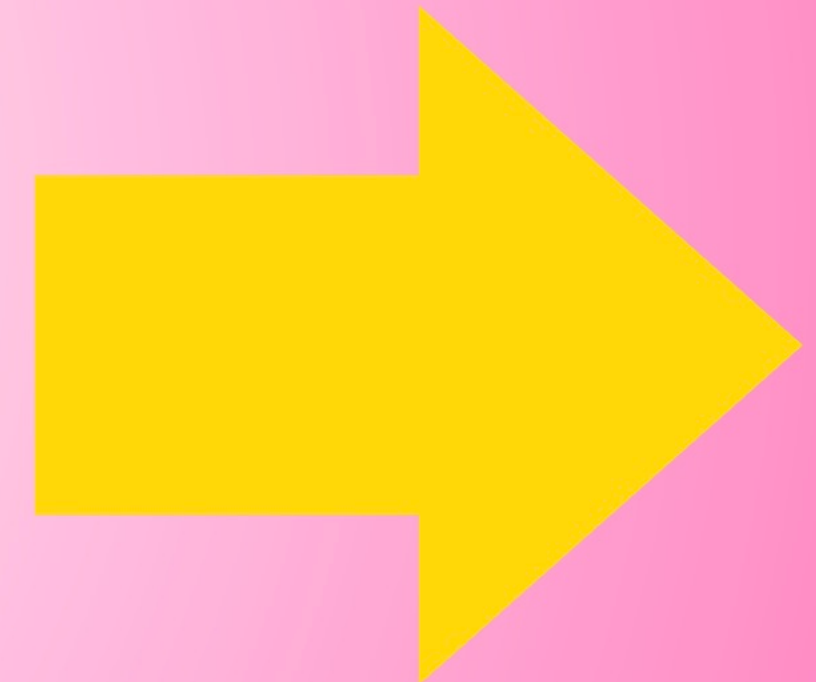
goedkeuring van de
normcommissie

publieke consultatie

commentaren verwerken

vaststelling door de
normcommissie

publicatie



publieke consultatie

te vinden op NEN normontwerpen:

<https://normontwerpen.nen.nl/>

Mocht je op de hoogte gehouden willen worden, geef dat dan na afloop even aan.

The screenshot shows the website normontwerpen.nen.nl. The header includes the NEN logo and the text "normontwerpen". Navigation links are "Home", "Inschrijven", "Neem contact op met NEN", and "Help". A main heading reads "Geef commentaar op NEN normontwerpen." Below this are four numbered steps: 1. Zoek (Search) with a magnifying glass icon, 2. Lees (Read) with an eye icon, 3. Commentaar (Comment) with a speech bubble icon, and 4. We maken normen en met colleg (We make norms and with colleagues) with a person icon. Each step has a brief description. At the bottom, there is a search bar with the placeholder "Zoek" and a magnifying glass icon, and a welcome message: "Welkom bij het NEN-systeem voor normontwerpen. U kunt hier normontwerpen inzien en op eenvoudige manier voorstellen!"

volgende stappen

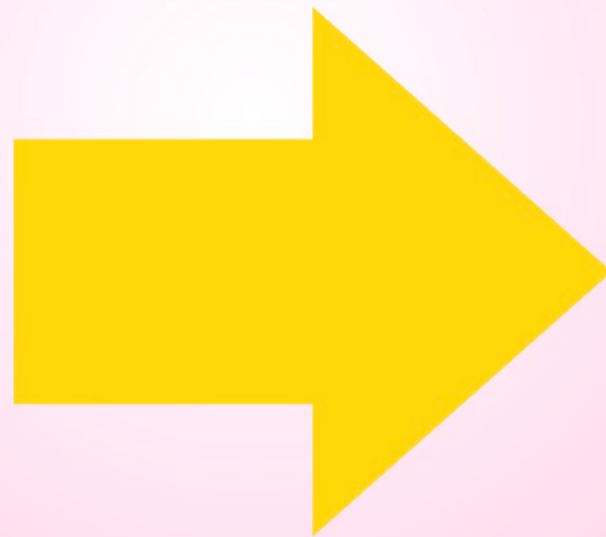
goedkeuring van de
normcommissie

publieke consultatie

commentaren verwerken

vaststelling door de
normcommissie

publicatie



gebruiken

(en doorontwikkelen)

vragen?

Meer informatie:

NEN: Sterre Langeveld sterre.langeveld@nen.nl

RIG: Nelleke Groen nelleke.groen@rijksoverheid.nl 06-50419530