



Rijksorganisatie voor Ontwikkeling,
Digitalisering en Innovatie
*Ministerie van Binnenlandse Zaken en
Koninkrijksrelaties*

AI is geen Science Fiction

ODI Kennissessie over AI

20 juni | 10:00-11:30

Joost van der Burgt (Rijks ICT Gilde)

*met introductie van Richard Vielvoije,
Directeur Rijksorganisatie ODI*



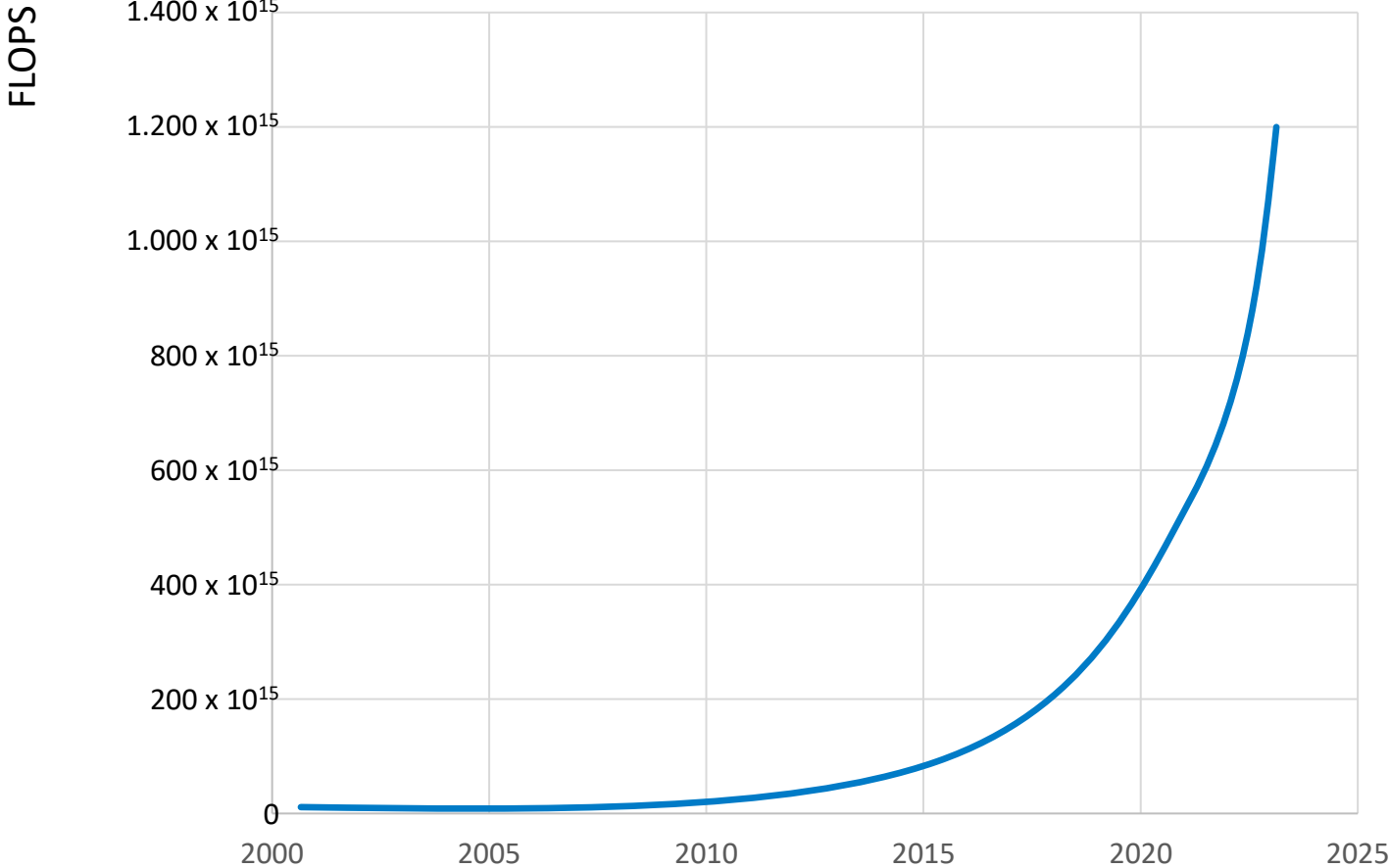
*“vanwege toegenomen
rekenkracht en data ...”*

- Inleiding van vrijwel ieder rapport over AI

Rekenkracht



Exponentiële groei rekenkracht



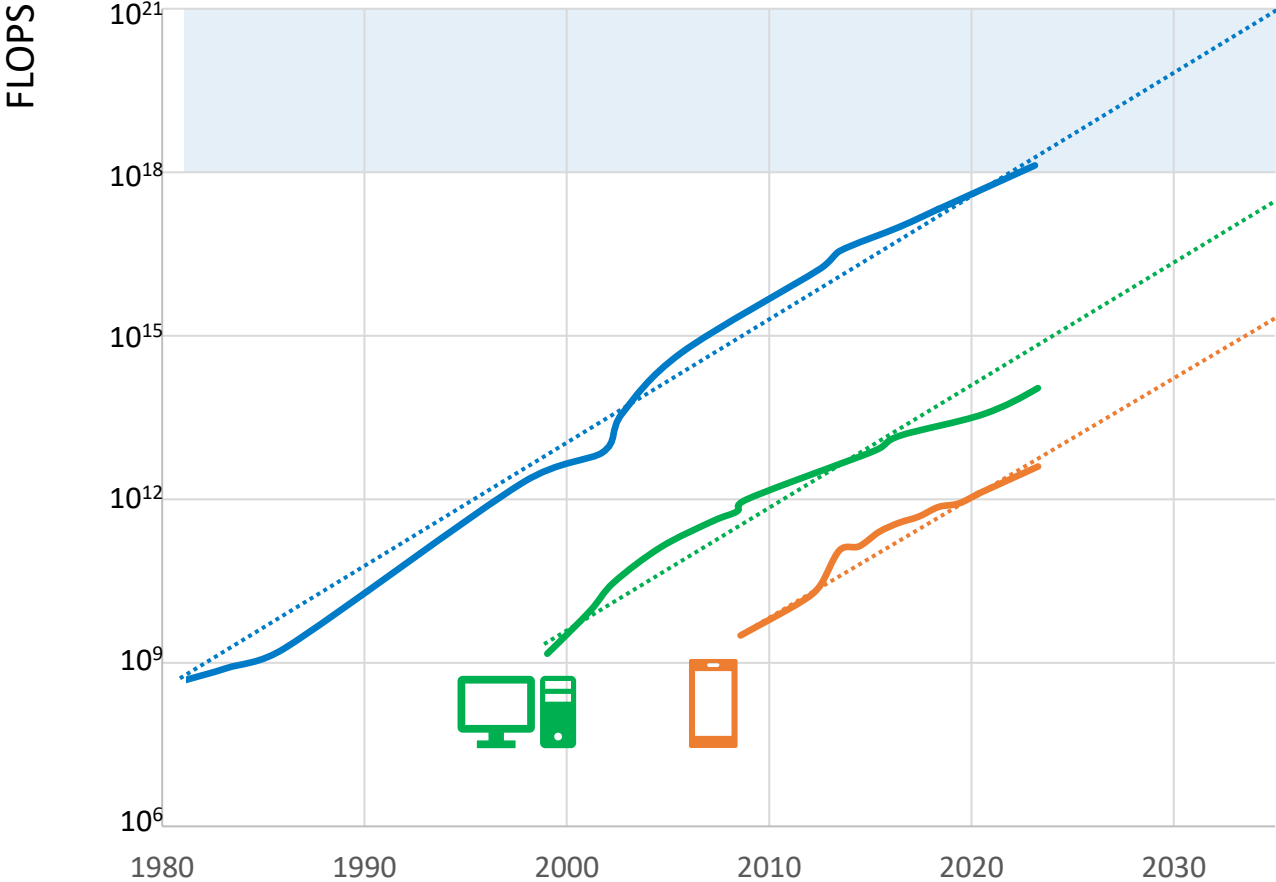
1,200,000,000,000,000,000
floating point operations per second

(of: 12,000 high-end PC's tezamen)

Rekenkracht



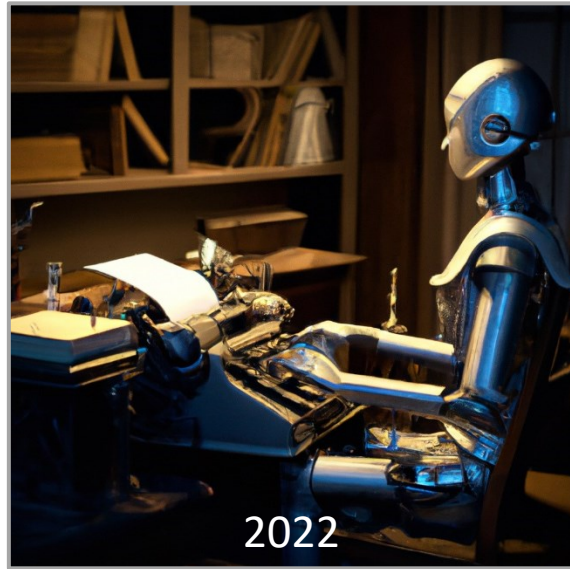
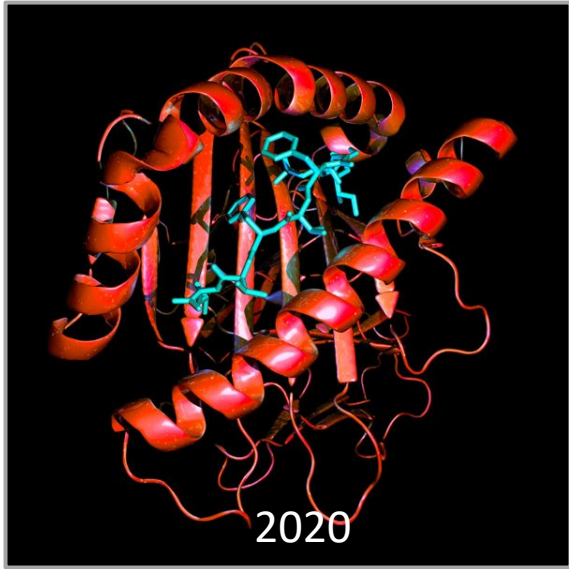
Exponentiële groei rekenkracht



equivalentie menselijk brein

- ▶ peak human ability at "StarCraft" and poker
- ▶ superhuman ability at Go, chess and shogi
- ▶ peak human ability at Go and Jeopardy

- ▶ peak human ability at chess



Impact AI



“Data is the new gold”

- Neelie Kroes (2011)

*“If data is the new gold,
Artificial Intelligence is the
new electricity”*

- Andrew Ng (2016)

1e

1765

2e

1870

3e

1969

4e

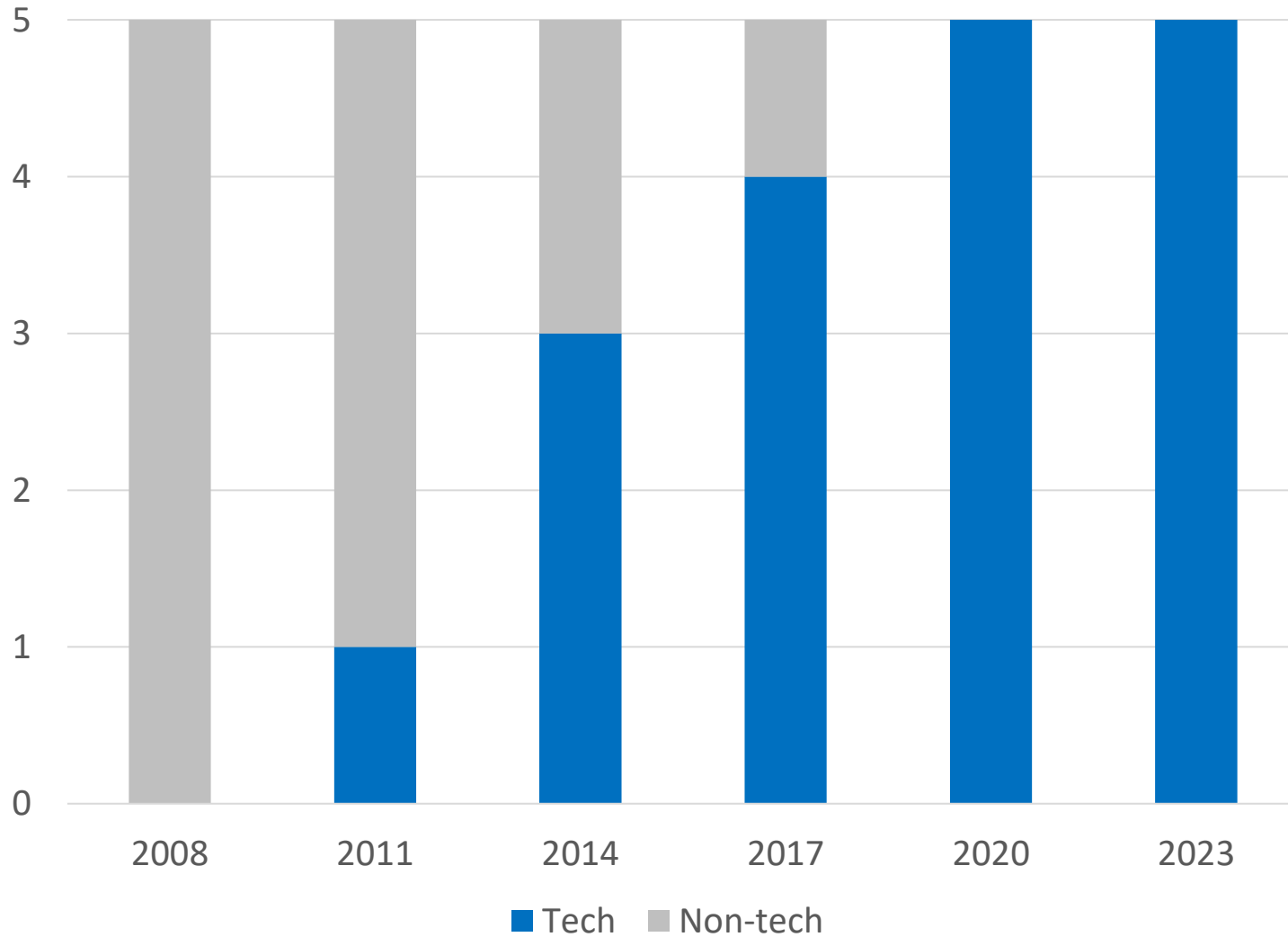
2020



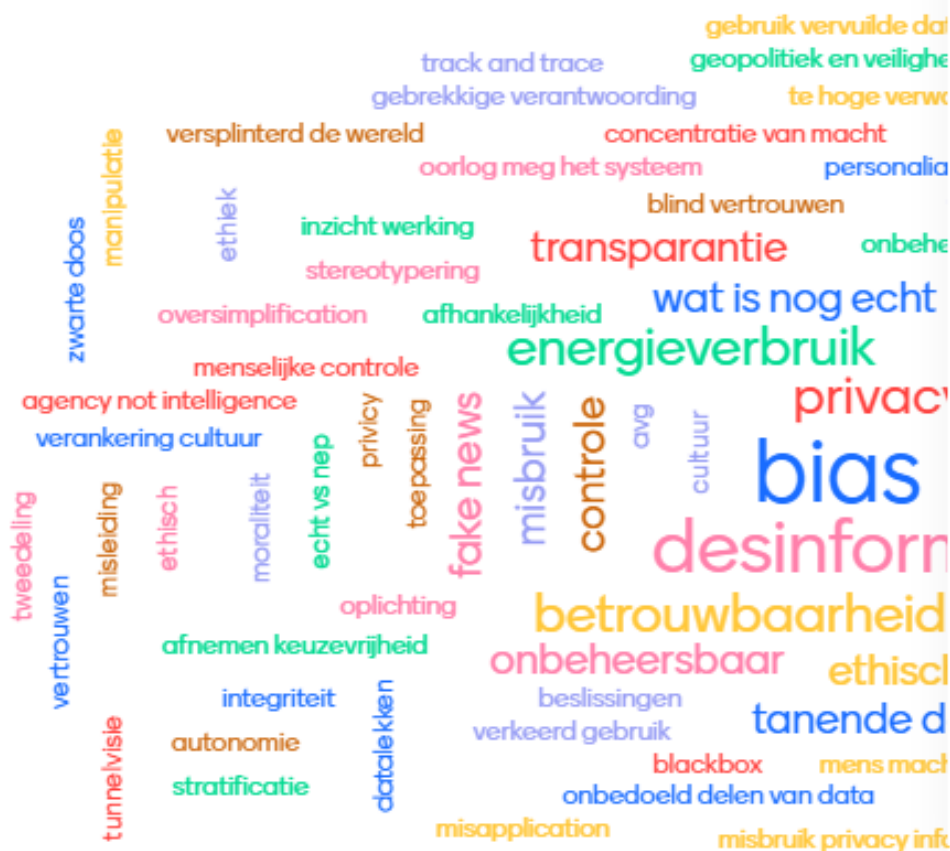
Impact AI



Top 5 grootste bedrijven wereldwijd



Wat zijn risico's van AI?



Ga naar menti.com en gebruik

Content



Account



Content



Design



Settings



Got feedback?

This Menti has results ?

Manage results

Your question ?

Wat zijn risico's van AI?

Add longer description

Entries per participant ?

3

Extras

Let participants submit multiple times ?

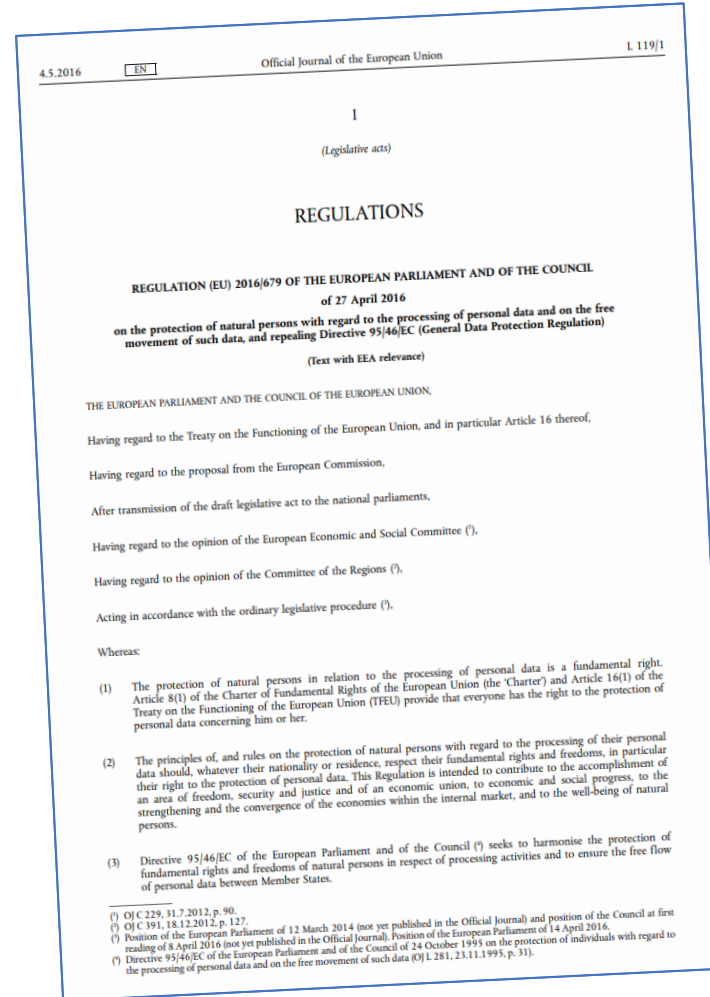
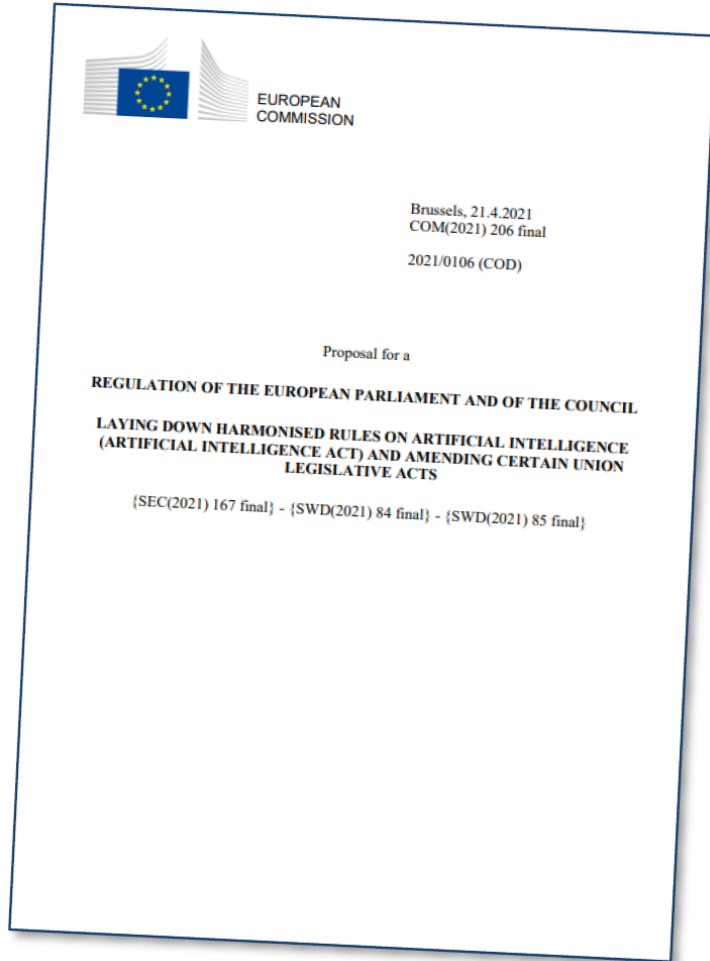


Remove question



AI en de overheid als wetgever

Internationaal



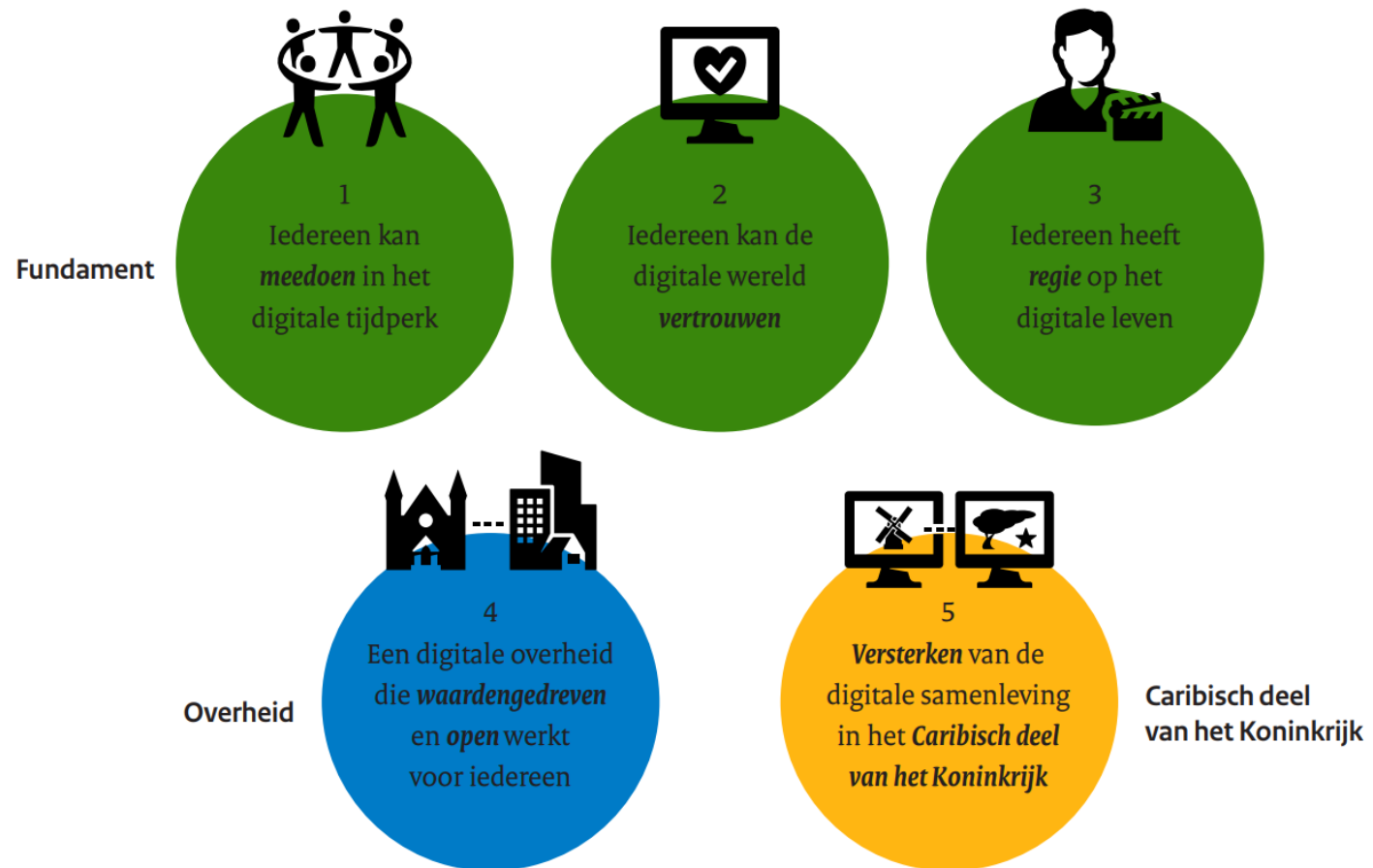
G7 calls for developing global technical standards for AI

By Kantaro Komiya and Supantha Mukherjee

May 20, 2023



Werkagenda *Waardengedreven* Digitaliseren



Werkagenda Waardengedreven Digitaliseren



2
Iedereen kan de digitale
wereld **vertrouwen**

2.1 Publieke waarden borgen | Overzicht



Maatschappelijke opgave

Publieke waarden zijn het fundament van een verantwoorde digitale transitie. Digitalisering kan leiden tot schending van publieke waarden online, zoals menselijke waardigheid, autonomie en non-discriminatie. Concentratie van marktmacht is een actueel gevaar. Welke waarden en normen in de online wereld gelden en hoe die beschermd worden is onduidelijk. Er is geen gedeeld begrippenkader van publieke waarden die belangrijk zijn bij digitalisering. Kinderrechten online, gaming en gokken zijn nog onvoldoende genormeerd. Burgers en publieke organisaties zijn te afhankelijk van private platformen voor publieke activiteiten, zoals onderwijs, onderzoek



Behaalde resultaten

- Het IAMA (mensenrechtentoets) is ontwikkeld voor discussie en besluitvorming rondom de inzet van algoritmen door overheidsorganen
- Op Europees niveau wordt er onderhandeld over nieuwe regels omtrent het borgen van publieke waarden online (DSA, DMA, DGA, DA, AI-act, GDPR)
- Interoperabiliteit van berichtendiensten is naar schatting vanaf 2024 verplicht via de Digital Markets Act (DMA)
- Kindermarketing online is bij grote platforms verboden

- Het IAMA (mensenrechtentoets) is ontwikkeld voor discussie en besluitvorming rondom de inzet van algoritmen door overheidsorganen

verhogen van de online en offline veiligheid van mensenrechtenverdedigers wereldwijd

Werkagenda Waardengedreven Digitaliseren



3 Iedereen heeft *regie*
op het digitale leven

3.3 Algoritmes reguleren | Overzicht



Maatschappelijke opgave

Regie op je digitale leven betekent onder meer dat men erop kan vertrouwen dat digitale systemen, waaronder algoritmes, aan publieke waarden voldoen en dat men begrijpt hoe deze werken. Bij bedrijven en de overheid is dit op het moment geen gegeven.

- Het is voor de maatschappij nu onvoldoende duidelijk welke normen, gebaseerd op publieke waarden, er voor algoritmes gewaarborgd moeten zijn. Zo heeft bijvoorbeeld de belastingdienst bij toeslagen gebruik gemaakt van een zelflerend algoritme dat voor de controle op fraude selecteert op inkomen, met een hogere risicoscore voor lage inkomens



Behaalde resultaten

- Instrumenten digitaal (AI) bewustzijn: AI parade (bij NL bibliotheken)
- Nederlandse inzet in EU AI Act heeft geleid tot sterkere opname van mensenrechten in de risicobeoordeling bij hoge risico AI systemen en aanvullende maatregelen die transparantie en rechten van natuurlijke personen versterken
- ELSA labs (ethical, legal, societal aspects) voor mensgerichte AI gestart waarin samengewerkt wordt aan in de praktijk realiseren van algoritmes die voldoen aan mensenrechten en publieke waarden

- Nederlandse inzet in EU AI Act heeft geleid tot sterkere opname van mensenrechten in de risicobeoordeling bij hoge risico AI systemen en aanvullende maatregelen die transparantie en rechten van natuurlijke personen versterken



De overheid als gebruiker van AI



SyRI-wetgeving in strijd met het Europees Verdrag voor de Rechten voor de Mens

De
IND maakte zich schuldig aan etnisch profileren

6 mei 2022

Gemeente
verboden algoritme

2 februari 2021

Mogelijke 'vooringenomenheid' bij politie-algoritmes

Fra
moec
18 mei 2022

6 maart 2023

Ondanks waarschuwingen gingen gemeenten door met omstreden fraudedetectiemethode

mi 2022

**Onderzoek: algoritmes overb
oldoen niet aan basiseisen**

3 mei 2022

Rekenkamer: nauwelijks aandacht voor ethiek bij algoritmes overheid

De Rekenkamer waarschuwt Rotterdam: groot risico op discriminatie en profilering door algoritmes

15 april 2021

2 mei 2023

Fiscus gebruikte in toeslagenaffaire algoritmes die mensenrechten schenden

Overheid gebruikt op grote schaal voorspellende algoritmes, 'risico op discriminatie'

29 mei 2019

icht van alg
: 'Gevaar voor onterechte
singen voor burgers'

van



Organisatie

[Belastingdienst](#) (1)

[Gemeente Amersfoort](#) (2)

[Gemeente Amsterdam](#) (8)

[Gemeente Den Haag](#) (26)

[Gemeente Rotterdam](#) (4)

[Gemeente Utrecht](#) (38)

[Kadaster](#) (5)

[Ministerie van Economische Zaken en](#)

[Klimaat, Rijksdienst voor](#)

[Ondernemend Nederland](#) (1)

[Ministerie van Financiën](#) (1)

[Omgevingsdienst](#)

[Noordzeekanaalgebied](#) (6)

[Provincie Noord-Brabant](#) (2)

[Provincie Zuid-Holland](#) (3)

[Rijksdienst Voor Identiteitsgegevens](#)
(12)

[Sociale Verzekeringsbank \(SVB\)](#) (6)

[Uitvoeringsinstituut](#)

[Werknemersverzekeringen \(UWV\)](#) (3)

Signaalmodel erfbelasting (UDA-Erf)

Het algoritme 'Uitnodiging tot het Doen van Aangifte Erfbelasting' wordt gebruikt door medewerkers van de Belastingdienst om vast te stellen of er sprake is van fraude voor de erfbelasting.

TVL Risicomodel

Dit model wordt gebruikt om voor alle aanvragen Tegemoetkoming Vaste Lasten een risicoinschatting te maken, voorafgaand aan een geautomatiseerde of handmatige toelating en uitbetaling van het voorschot.

Berekenen huurtoeslag

Het algoritme huurtoeslag berekent de hoogte op basis van de huurwoning. Hierbij worden de huurtoeslagregels toegepast.

Organisatie	Afdeling
Ministerie van Financiën	Dienst Toeslagen

Geautomatiseerde gezichtsvergelijking bij het RNI-inschrijfproces.

Dit model kan helpen om "look-alike-fraude" te voorkomen.

Afdeling	Beleidssterrein
Directie	Basisregistratie

Preventie & Handhaving

Met dit model voorspellen wij hoe groot de kans is op samenwoont.

Organisatie	Afdeling
Sociale Verzekeringsbank (SVB)	Veld n...

Type algoritme
Zelflerend

Kadastrale kaart next

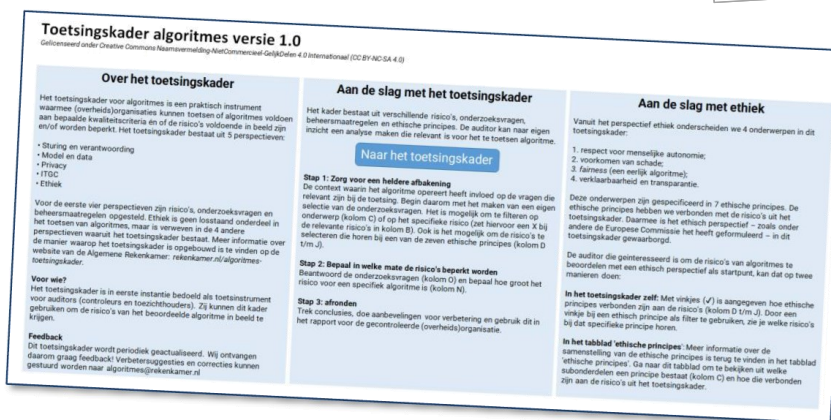
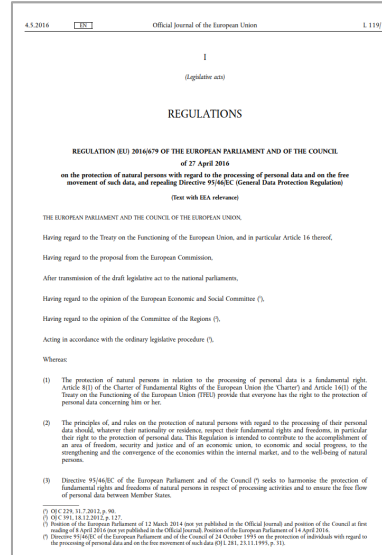
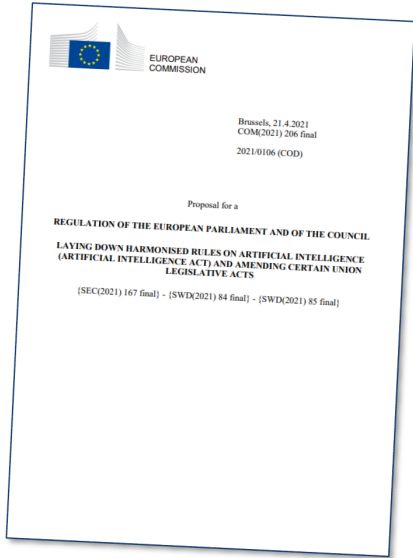
Met artificial intelligence (kunstmatige intelligentie) bouwen we de kadastrale kaart opnieuw op. Hiervoor gebruiken we scans van originele historische veldwerken.

Beleidssterrein
wonen, werken en leefomgeving



OpenAI
ChatGPT 4.0

Kaders

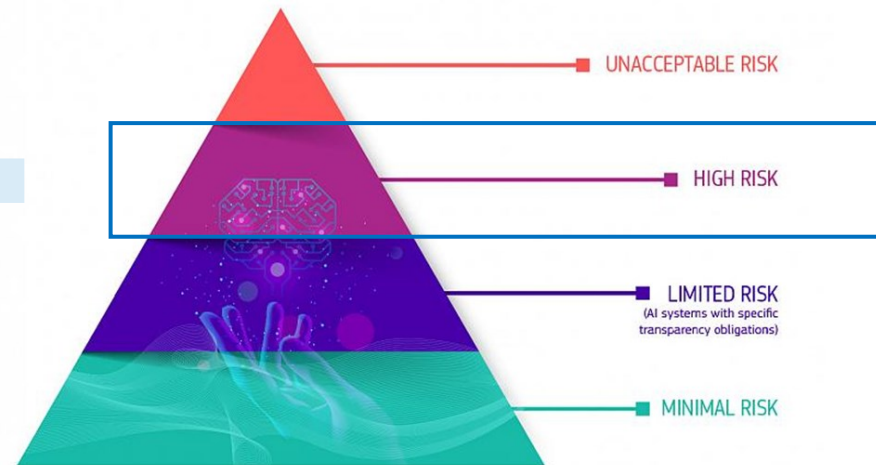
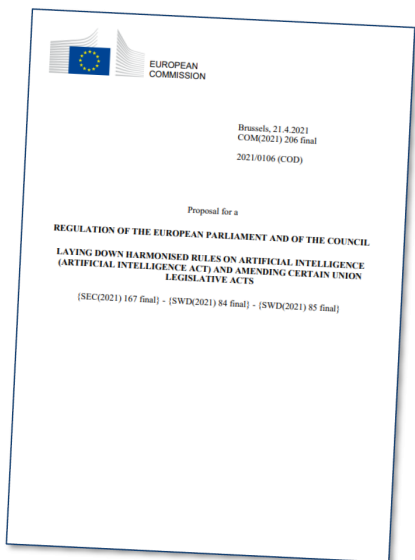


OVERHEID

**VERANTWOORDE
AI**



Kaders



BRON: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>



Artikel 8
Naleving van de voorschriften

Artikel 12
Registratie (logging)

Artikel 9
Systeem voor risicobeheer

Artikel 13
Transparantie aan gebruikers

Artikel 10
Data en databeheer

Artikel 14
Menselijk toezicht

Artikel 11
Technische documentatie

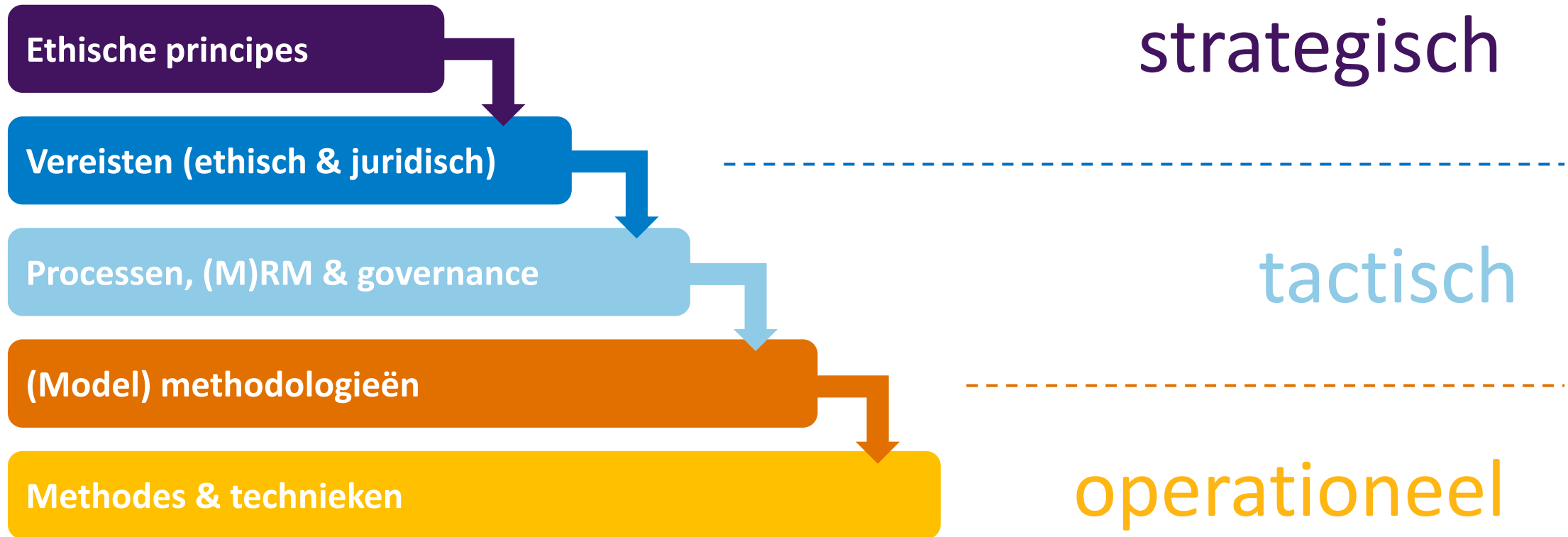
Artikel 15
Nauwkeurigheid & robuustheid



Ministerie van Binnenlandse Zaken en
Koninkrijksrelaties

Implementatiekader

Inbedding





Nieuwe eisen aan mens & organisatie

Impact Assessment
Mensenrechten en Algoritmes



1 Waarom?

Beoogde effecten (doel)

Aanleiding

Vragen, doelen, waarden...

→ Bekijk hoofdstuk 1



2 Wat?

A) Data (input)

Randvoorwaarden

Vragen, organisatie techniek, data, wetgeving...

B) Algoritme (throughput)

Randvoorwaarden

Vragen, organisatie techniek, transparantie, uitlegbaarheid

→ Bekijk hoofdstuk 2



3 Hoe?

Implementatie en gebruik algoritme (output)

Randvoorwaarden

Vragen, organisatie techniek, data...

→ Bekijk hoofdstuk 3



4 Mensenrechten

Nagestreefde doelen

Aangetaste grondrechten

→ Bekijk hoofdstuk 4



	IAMA Onderdeel 1	IAMA Onderdeel 2	IAMA Onderdeel 3	IAMA Onderdeel 4
Belangengroep	●			
Bestuur	●			
Burgerpanel	●			
CISO of CIO	●	●		
Communicatieadviseur		●	●	
Data scientist		●	●	
Databeheerder of bronhouder		●		
Domeinexpert (medewerker met domeinkennis op het gebied waar het algoritme toege- past gaat worden)	●	●	●	●
Functionaris Gegevensbescherming		●		
HR-medewerker			●	
Jurist	●	●	●	●
Ontwikkelaar algoritme		●		
Opdrachtgever	●	●	●	
Overige leden projectteam	●			
Projectleider	●	●	●	●
Strategisch adviseur ethiek		●	●	

2 Wat?

A) Data (input)

Randvoorwaarden

Vragen, organisatie techniek, data, wetgeving...

B) Algoritme (throughput)

Randvoorwaarden

Vragen, organisatie techniek, transparantie, uitlegbaarheid

→ Bekijk hoofdstuk 2



CISO of CIO

Communicatieadviseur

Data scientist

Databeheerder of bronhouder

Domeinexpert

Functionaris gegevensbescherming

HR-medewerker

Jurist

Ontwikkelaar algoritme

Opdrachtgever

Overige leden projectteam

Projectleider

Strategisch adviseur ethiek



2A.2

Databronnen en kwaliteit

- Wat voor type data gaat gebruikt worden als input voor het algoritme en uit welke bronnen is de data afkomstig?
- Is de kwaliteit en betrouwbaarheid van de data voldoende voor de beoogde datatoepassing? Leg uit.



2A.3

Bias / aannames in de data

- Welke aannames en bias liggen in de data besloten en hoe wordt de invloed daarvan op de output van het algoritme gecorrigeerd of anderszins ondervangen of gemitigeerd?
- Indien gebruik wordt gemaakt van trainingsdata: is de data representatief voor de context waarin het algoritme ingezet gaat worden?



2B.1

Type algoritme

- Type algoritme: wat voor soort algoritme wordt gebruikt of gaat worden gebruikt? Hoe werkt het? Onderscheid tussen:
 - a) een niet-zelflerend algoritme, waarin de mens de regels specificeert die de computer moet volgen
 - b) een zelflerend algoritme, waarin de machine zelf leert over de patronen in de data (machine learning)
- Waarom wordt voor dit type algoritme gekozen?
- Waarom is dit type algoritme het meest geschikt om de geformuleerde doelstellingen te bereiken?
- Welke alternatieven zijn er en waarom zijn die minder passend of bruikbaar?



2B.3

Accuraatheid algoritme

- Wat is de accuraatheid van het algoritme, en op basis van welke evaluatiecriteria is deze accuraatheid bepaald?
- Is de mate van accuraatheid acceptabel voor de manier waarop het algoritme ingezet zal worden? Hoe wordt het algoritme getest?
- Welke maatregelen kunnen worden getroffen om de risico's van reproductie of zelfs versterking van biases tegen te gaan (bijv. andere samplingstrategie, feature modification, ...)?
- Hoe vaak/erg zit het algoritme ernaast? (bijv. in termen van aantal false positives, false negatives, R-squared, ...)

3 Hoe?

Implementatie en gebruik
algoritme (output)

Randvoorwaarden

Vragen, organisatie techniek, data...

→ Bekijk hoofdstuk 3



CISO of CIO

Communicatieadviseur

Data scientist

Databeheerder of bronhouder

Domeinexpert

Functionaris gegevensbescherming

HR-medewerker

Jurist

Ontwikkelaar algoritme

Opdrachtgever

Overige leden projectteam

Projectleider

Strategisch adviseur ethiek



3.2

De rol van de mens in de beslissing

- Welke rol spelen mensen bij het nemen van beslissingen op basis van de output van het algoritme ('human in the loop') en hoe worden zij in staat gesteld om die rol te spelen?
- Is er nu en in de toekomst voldoende gekwalificeerd personeel aanwezig om het algoritme te beheren, te herzien en aan te passen indien gewenst/nodig?



3.3

Effecten van het algoritme

- Wat zullen de effecten zijn van de inzet van het algoritme voor burgers en hoe wordt rekening gehouden met de 'menselijke maat' bij het nemen van beslissingen op basis van de output van het algoritme?
- Welke risico's voor stigmatiserende, discriminerende of anderszins schadelijke of nadelige effecten zijn er voor de burger en hoe zullen die worden bestreden of gemitigeerd?



3.5

Context

- **Tijd/periode:** wanneer gaat het algoritme ingezet worden? Hoe lang is de periode dat het ingezet wordt?
- **Plaats:** waar vindt inzet van het algoritme plaats? Is dat in een bepaald geografisch gebied, is dat bij een bepaalde groep personen of dossiers?
- Kan het algoritme ook nog worden ingezet als contextfactoren veranderen of als het algoritme gebruikt wordt in een andere context dan waarvoor het is ontwikkeld?



3.7

Evaluatie, auditing en borging

- Is voorzien in goede instrumenten voor evaluatie, auditing en borging van het algoritme?
- Zijn er voldoende mogelijkheden om rekenschap en verantwoording af te leggen over het algoritme?
- Welke mogelijkheden zijn er voor auditors en toezichthouders om (formele) consequenties te verbinden aan de inzet van een algoritme door de overheid (bijv. terugkoppeling van bevindingen, doen van aanbevelingen, budgettaire consequenties, ...)



Contact



www.rijksictgilde.nl



RIG@rijksoverheid.nl



joost.burgt@rijksoverheid.nl