Wat is AI, wat is de impact van AI op de maatschappij, en hoe kun je AI op een verantwoorde manier vormgeven?

Dr. Mirjam Plantinga

Universitair Hoofddocent UMCG, afdeling genetica en Data Science Center in Health (DASH)
Projectleider ELSA AI lab Noord Nederland

DASH
Data Science Center in Health

m.plantinga@umcg.nl

ELSA-NN

Hoe denkt u over AI?

“AI ontwikkelt zich zó snel, dat het binnenkort alle banen overneemt.”

Wat denk jij?

HELEMAAL WAAR!

ZO’N VAART LOOPT HET NIET

Zie cursus over AI op:
kbo-provinciegroningen.nl

“Als ik aan AI denk, dan denk ik aan…”

(Kies een antwoord)

TERMINATOR

CHATGPT

ETHISCHE DILEMMA'S

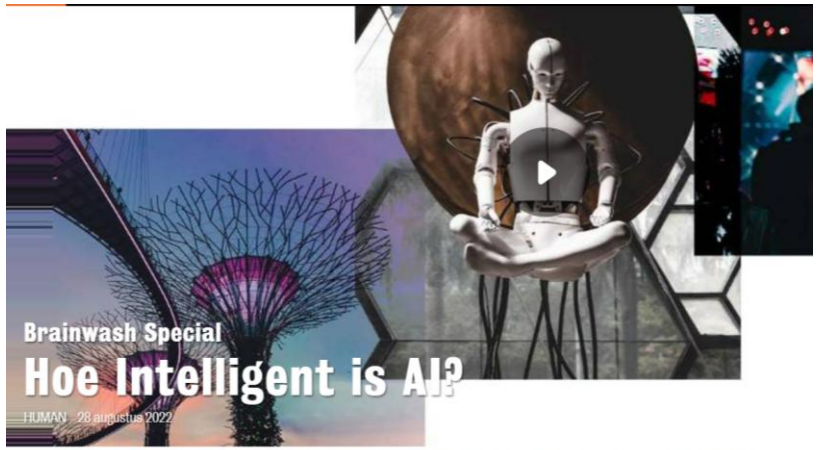
ZORGROBOTS

SCIENCEFICTION

DASH
Data Science Center in Health

Definitie AI

- Verzamelnaam voor algoritmes en methoden die taken uitvoeren waarvan vroeger werd gedacht dat het alleen voor menselijke intelligentie was weggelegd



Wat kunnen we met kunstmatige of artificiële intelligentie (AI)? Hoe maken we het eerlijk? En: wat is eerlijk in het geval van AI? Een gesprek onder leiding van Maaike Schoon met techniekfilosoof Peter-Paul Verbeek, Europarlementariër Kim van Sparrentak, kunstenaar en ontwerper Julia Janssen, en column columnist en digitaal strateeg Ilyaz Nasrullah.

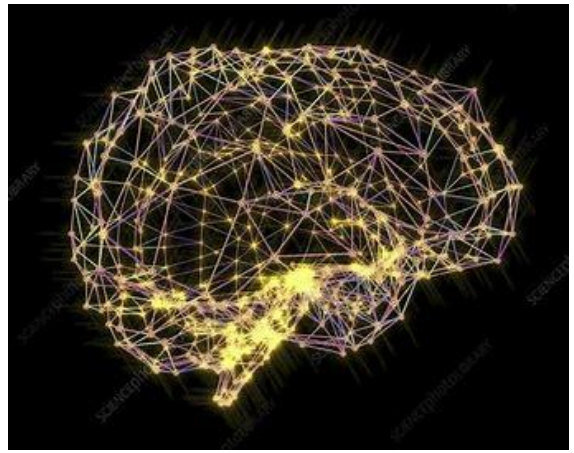
- Brainwash uitzending NPO (aug 2022): Er is niet één definitie van AI
- AI ook geen gelukkig gekozen term?
- AI als discipline stamt al uit de jaren 50
- AI gaat over het analyseren van grote hoeveelheden data door een computer
- Is dat intelligent?

DASH
Data Science Center in Health

ELSA-NN

AI en machine learning

- Een ding is zeker: zonder data is er geen AI



- AI gaat over verwerking van data (of gegevens) door computers
- De term AI verwijst naar de manier waarop mensen gegevens verwerken (mbv neuraal netwerk)

- Verwerking van gegevens door computers kan op twee manieren:

1) Door regels/instructies (het algoritme) te programmeren (kennisgedreven of ‘rule based’ AI)

2) Door de computer zelf het algoritme te laten ontwikkelen (datagedreven AI/machine learning)

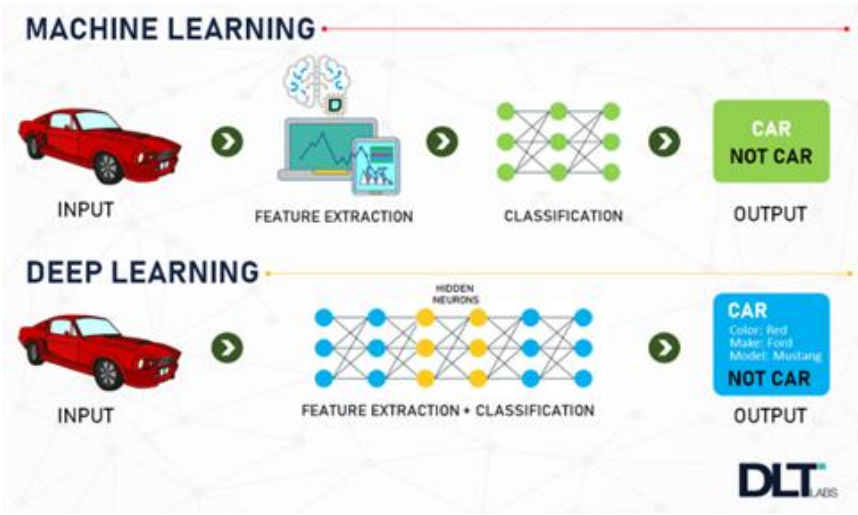
- Wanneer men spreekt over AI, heeft men het vaak over de techniek van machine learning
- Bij machine learning wordt het algoritme niet geprogrammeerd maar getraind (door mensen)

DASH
Data Science Center in Health

ELSA-NN

Machine learning en deep learning

- Bij machine learning ontwikkelt de computer het algoritme zelf door te leren van grote hoeveelheden data waarmee het getraind wordt
- De verwerking van de gegevens vindt plaats op een manier die lijkt op hoe onze hersenen werken (dmv ontwikkelen neuraal netwerk).
- Deep learning is familie van machine learning
- Het neurale netwerk is alleen ingewikkelder
- Deze vorm van AI heeft véél meer data nodig om tot een goed model te komen
- Deze vorm van AI kan dus patronen in kaart brengen en aan de hand van deze patronen voorspellingen doen (bijvoorbeeld of een plaatje een hond of kat is of een stukje weefsel goedaardig of kwaadaardig)
- Er kan met behulp van deep learning zoveel data geanalyseerd worden, dat patronen ontdekt kunnen worden die getrainde mensen hoogstwaarschijnlijk zouden missen

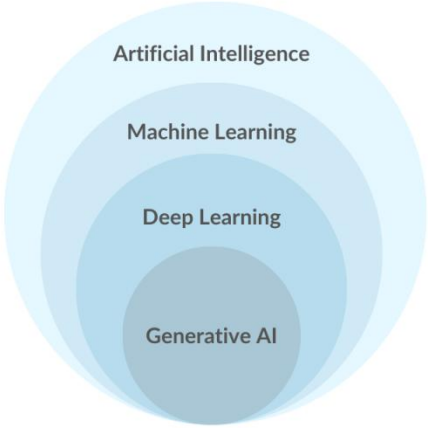


DASH
Data Science Center in Health

ELSA-NN

Generatieve AI

- Dit is een vorm van AI die ook gebaseerd is op het principe van neurale netwerken
- In dit geval worden de geleerde patronen uit de data gebruikt om nieuwe content/informatie te genereren (bijvoorbeeld tekst of plaatjes)
- Generatieve AI voorspelt per woord wat het meest waarschijnlijke volgende woord is
- De data die gegenereerd worden (tekst of plaatjes) zijn dus niet gebaseerd op een bepaalde vorm van kennis, het is puur kansberekening
- Deze technologie heeft nog meer data nodig dan deep learning



Generative AI

- It learns from existing data & generates brand new artifacts.
- Gen AI produces new content: images, speech, video, formulas, music, text, code.
- Risks include: bias, accuracy, transparency, cybersecurity.
- In 2023 Gen AI is used in: marketing, sales, gaming, graphic design, programming, finance, healthcare.

dp DEV.PRO

Kijkersvragen: AI-editie | De Avondshow met Arjen Lubach (S5)

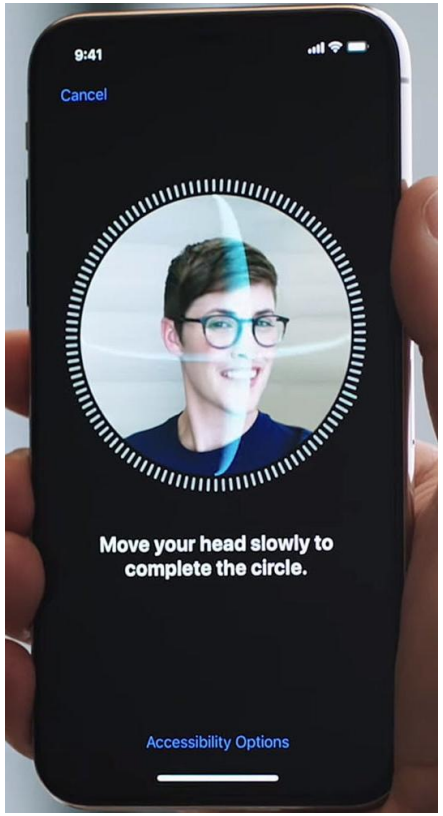
DASH
Data Science Center in Health

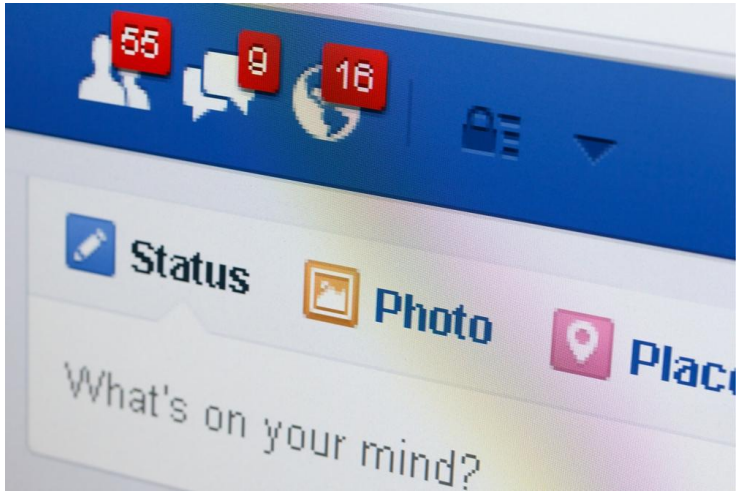
AI gebruik in de samenleving

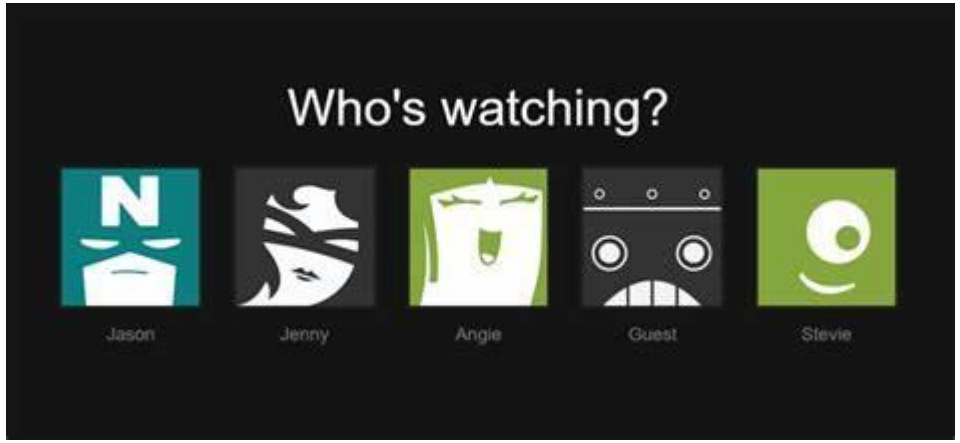
DASH
Data Science Center in Health

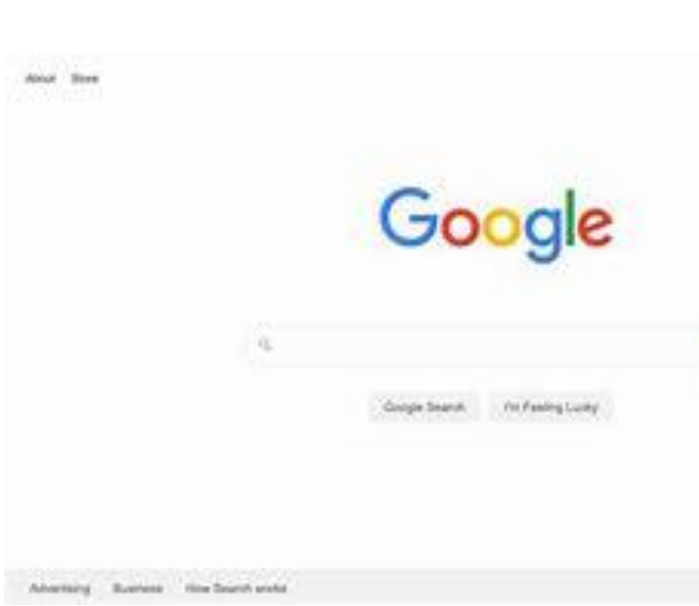
ELSA-NN

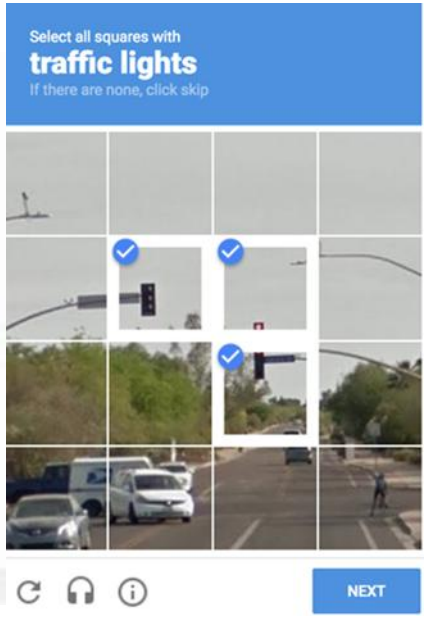
Iedereen heeft al te maken met AI



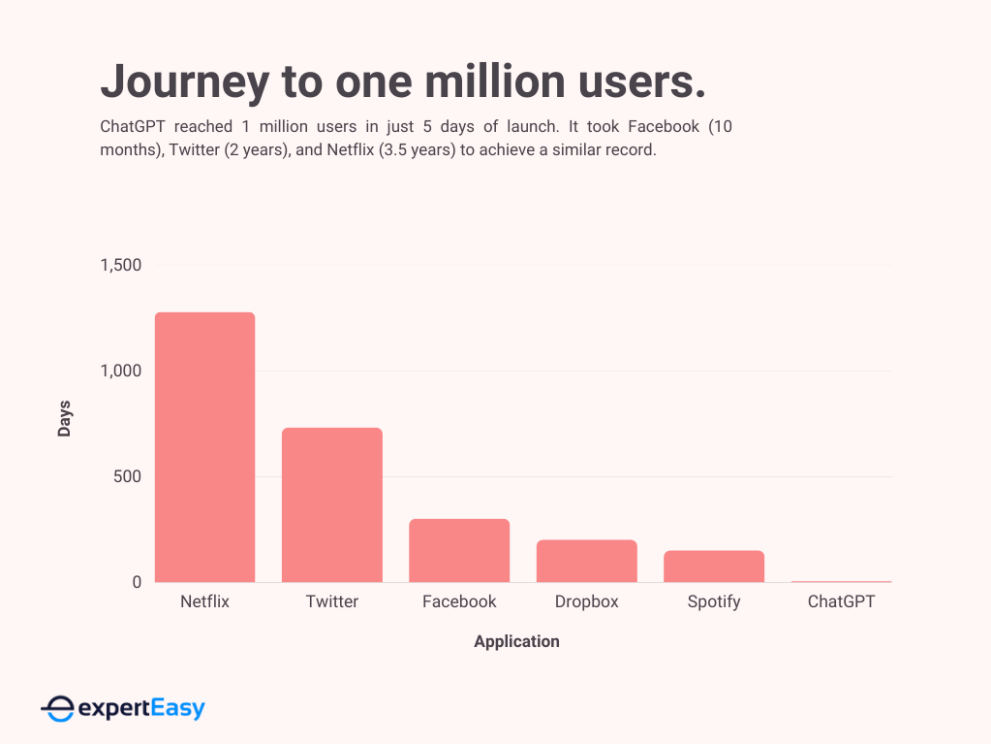








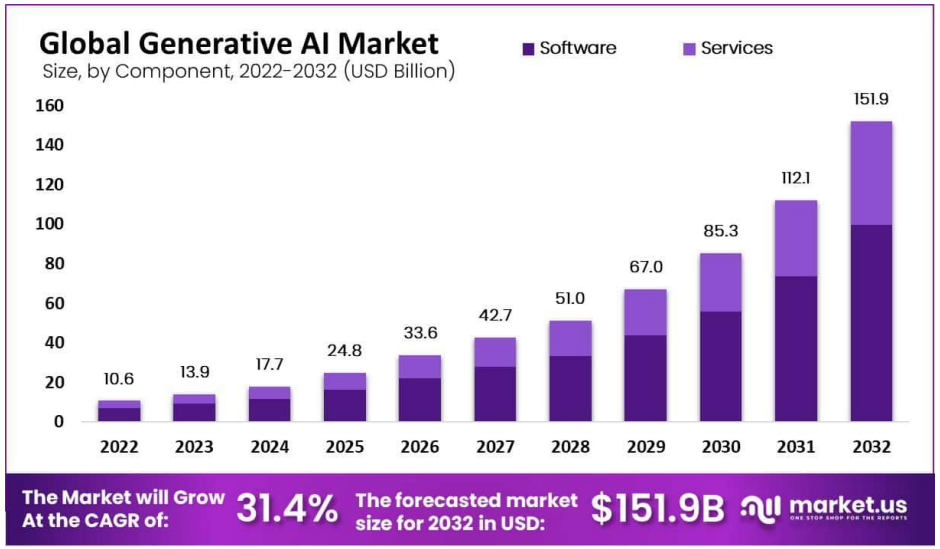
AI ontwikkelt zich razendsnel en met hoge verwachtingen



Wetenschappers voorspellen dat in 2027 meer dan 90% vd online content geheel of gedeeltelijk afkomstig is van AI.

Tilburg Institute for Law, Technology, and Society

Waarom nu?
Door toename in beschikbare data
en rekenkracht computers



AI heeft een grote impact op mens en maatschappij

- AI is de nieuwe systeem technologie
- Iedereen wordt erdoor beïnvloed
- AI wordt ontwikkeld mbv data van inwoners
- AI wordt gebruikt in dienstverlening aan inwoners
- AI wordt gebruikt door inwoners zelf

DASH
Data Science Center in Health

ELSA-NN

Kansen voor gebruik van AI in de zorg

Test met Sam, de UV-C-desinfectierobot

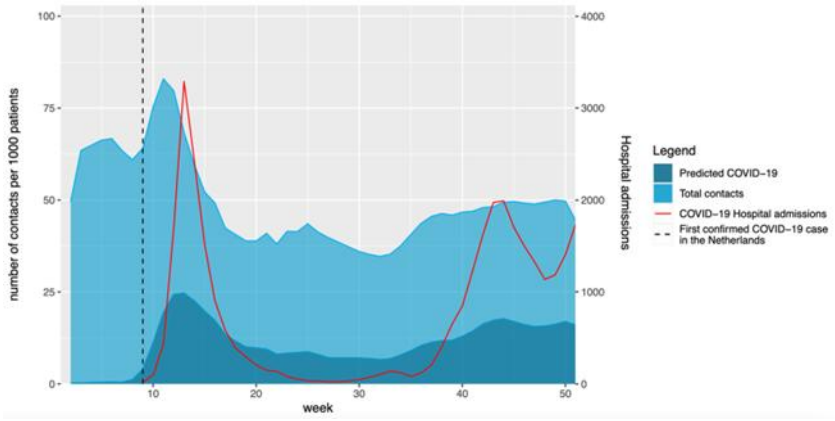
30-10-2023

Vanaf november gaat Reiniging & Desinfectie (B&F) een jaar lang experimenteren met een echte robot! Hij heet Sam en desinfecteert met UV-C-licht. UV-C-lichtdesinfectie is kwalitatief beter en sneller dan handmatige desinfectie, weten we. Daarnaast is het duurzamer omdat er minder doekjes en desinfectiemiddelen worden gebruikt. Maar UV-C-licht verwijdert geen eiwitten, vandaar dat desinfectie van contactpunten, verwijdering van zichtbaar vuil en het vervangen van gordijnen mensenwerk blijven. Het komende jaar onderzoeken we de volgende vragen met de test:

- Is UV-C-desinfectie minder belastend voor onze medewerkers?
- Wat doet UV-C-licht met materialen die gedesinfecteerd worden?
- Hoe verhoudt de inzet van Sam in ons ziekenhuis zich in tijd tot handmatige desinfectie?

In overleg met Medische Microbiologie/Infectiepreventie wordt gestart op enkele afdelingen en later uitgebreid naar meer afdelingen. De afdelingen waar getest wordt, worden vooraf door de teamleider van Reiniging & Desinfectie geïnformeerd. De te desinfecteren ruimte wordt voorbereid, zodat er geen plekken zijn waar het UV-C-licht niet kan komen. Een speciaal daarvoor opgeleide medewerker van Reiniging & Desinfectie bedient Sam. Tijdens het desinfecteren moet de ruimte afgesloten blijven. Na afloop geeft Sam een rapport waarop de medewerker kan zien of er geen 'blinde vlekken' zijn geweest. Als Sam bezet is, gebeurt de desinfectie zoals we tot nog toe gewend waren, dus met de hand.

Voorspellen COVID-19 (Lilian Peters et al, 2023)



ETZ, UMCG en Amsterdam UMC zetten ChatGPT-model in om patiëntvragen te beantwoorden

Sytse Wilman 21 augustus 2023, 10:28 4164 keer gelezen

Het ETZ (Elisabeth-TweeSteden Ziekenhuis) zet als proef generatieve artificial intelligence (AI) in om medische vragen van patiënten te beantwoorden. Ook het UMC Groningen en Amsterdam UMC gaan de toepassing van epd-leverancier Epic uitproberen. Andere ziekenhuizen met hetzelfde epd kunnen later aansluiten.

De implementatie van AI in de praktijk is echter nog niet bepaald probleemloos….

- AI model Amazon: selectie van uitsluitend mannen...

- AI modellen DUO en Belastingdienst: selectie op mensen met een migratieachtergrond…

Dienst Uitvoering Onderwij
Ministerie van Onderwijs, Cultuu
Wetenschap

DASH
Data Science Center in Health

ELSA-NN

De implementatie van AI in de praktijk is echter nog niet bepaald probleemloos….

- Gezichtsherkenningssoftware ontwikkelen met alleen voorbeelden van mensen met blanke huid?

- Een robothond inzetten bij ouderen, klinkt mooi?

DASH
Data Science Center in Health

ELSA-NN

Hoe moet het dan wel? Verantwoorde inzet AI!

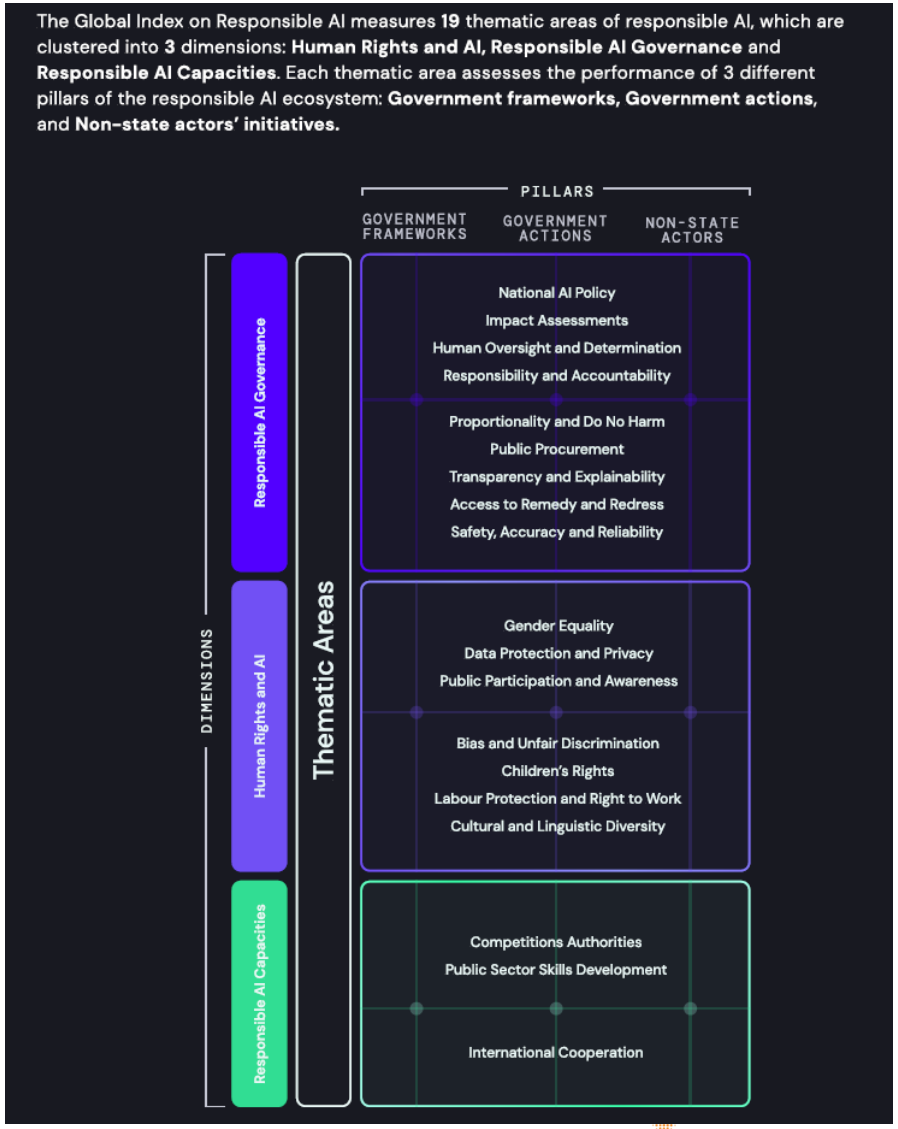
Global Rankings and Scores 1 – 25

RANKINGS	COUNTRY	REGION	INDEX SCORE	PILLAR SCORE			DIMENSION SCORE		
				Government frameworks	Government actions	Non-state actors	Human rights and AI	Responsible AI capacities	Responsible AI governance
1	Netherlands	Europe	86.16	74.33	95.46	91.23	78.74	88.59	91.12
2	Germany	Europe	82.77	72.69	93.00	82.48	80.30	64.03	90.94
3	Ireland	Europe	74.98	81.71	74.17	63.16	84.11	57.60	73.68
4	United Kingdom	Europe	73.12	60.66	80.90	82.48	67.59	66.54	79.62
5	USA	North America	72.81	62.41	79.19	80.87	65.37	84.34	74.75
6	Estonia	Europe	67.61	65.58	81.52	43.86	65.85	58.54	72.01
7	Italy	Europe	61.8	59.99	55.96	77.10	66.34	38.21	66.13
8	France	Europe	57.62	56.23	58.92	57.81	39.84	55.17	72.27
9	Canada	North America	57.39	37.80	80.84	49.68	55.37	57.06	59.07
10	Australia	Asia and Oceania	56.22	34.69	72.44	66.82	71.31	31.62	52.67
11	Singapore	Asia and Oceania	53.77	43.70	70.68	40.09	48.35	52.10	58.54
12	Japan	Asia and Oceania	52.21	28.86	72.54	58.25	41.67	44.19	63.07
13	Slovenia	Europe	49.58	68.89	39.63	30.84	51.97	51.35	47.12
14	Portugal	Europe	49.44	66.06	34.74	45.61	54.47	55.89	43.38
15	Switzerland	Europe	48.11	56.15	50.42	27.41	32.92	54.14	57.92
16	Spain	Europe	45.08	68.02	33.36	22.64	47.59	48.99	41.82
17	United Arab Emirates	Middle East	44.66	46.40	55.62	19.24	40.58	51.15	45.66
18	Brazil	South and Central America	44.42	44.56	40.86	51.26	44.85	35.97	46.90
19	Uruguay	South and Central America	44.09	35.64	54.03	41.12	47.20	37.59	43.84
20	Finland	Europe	43.07	36.45	49.29	43.86	36.58	45.69	47.24
21	Poland	Europe	42.73	63.50	34.97	16.73	43.83	42.15	42.07
22	Romania	Europe	42.11	26.61	55.81	45.71	36.96	34.57	48.63
23	Chile	South and Central America	40.38	33.36	49.18	36.84	42.45	26.30	43.47
24	Belgium	Europe	38.55	57.24	27.13	23.99	40.35	25.11	41.62

DASH
Data Science Center in Health

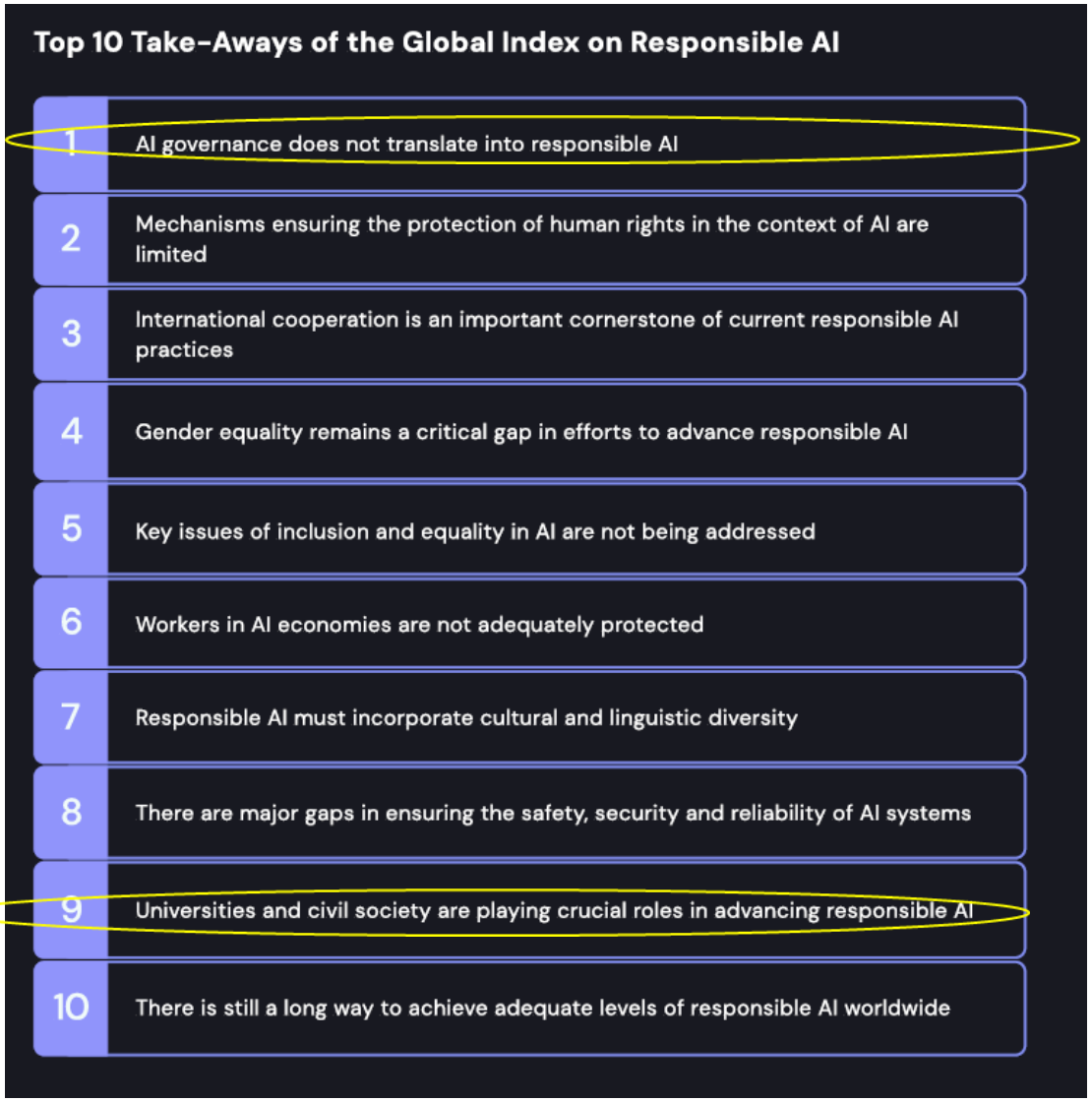
Verantwoorde inzet AI

- 19 thematische aandachtsgebieden
- 3 dimensies:
  - Verantwoorde AI Governance
  - Mensenrechten en AI
  - Verantwoorde AI capaciteiten
- 3 pilaren
  - Overheids raamwerken
  - Overheids acties
  - Acties niet-overheids actoren



DASH
Data Science Center in Health

Verantwoorde inzet AI

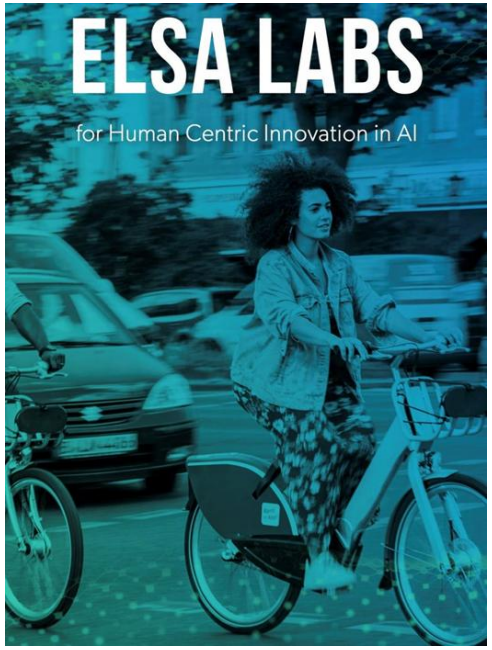


DASH
Data Science Center in Health

ELSA-NN

Verantwoorde ontwikkeling en implementatie van AI

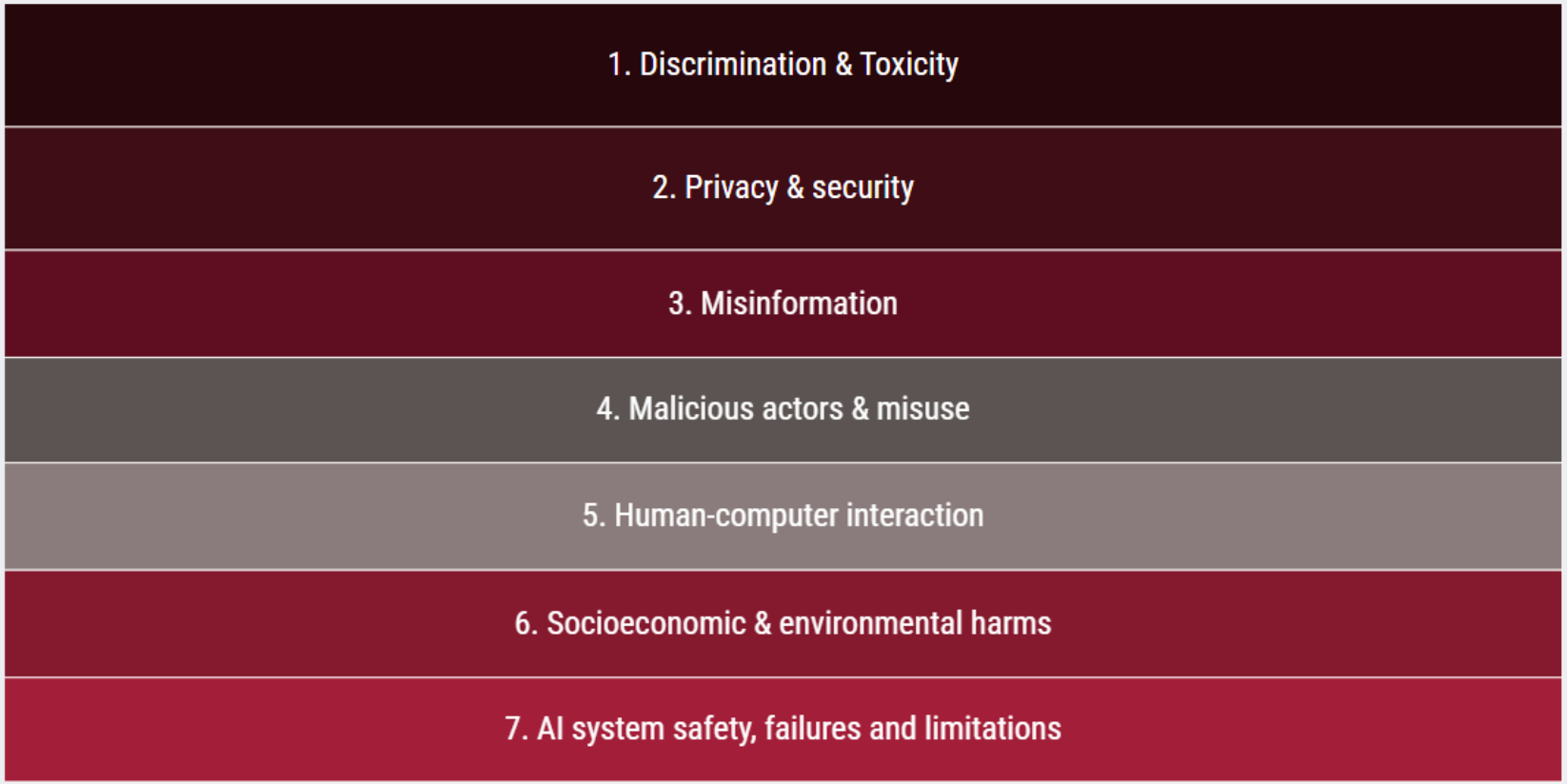
- **Bewustwording** van de impact van AI voor mens en maatschappij en de **verantwoordelijkheid** die hiermee gepaard gaat om verantwoorde inzet van AI te bevorderen
- Niet alleen aandacht voor **kansen**, maar ook voor **risico's** en uitdagingen van AI (**ELSA**)
- **Uitgangspunten** en principes voor verantwoorde inzet van AI **vertalen naar de praktijk**



Risico’s van AI

Domain Taxonomy of AI Risks

The Domain Taxonomy of AI Risks classifies risks from AI into seven domains and 23 subdomains. You can explore the taxonomy (to four levels of depth) in the interactive figure below. Read our preprint for more detail.



Slattery, P., Saeri, A. K., Grundy, E. A. C., Graham, J., Noetel, M., Uuk, R., Dao, J., Pour, S., Casper, S., & Thompson, N. (2024). A systematic evidence review and common frame of reference for the risks from artificial intelligence. http://doi.org/10.13140/RG.2.2.28850.00968

Risico’s en uitdagingen van AI

1) Onrealistische beeldvorming

2) Uitdagingen op het gebied van dataverzameling

3) Mens-machine interactie

4) Transparantie

DASH
Data Science Center in Health

Onrealistische beeldvorming

- (te) negatieve beeldvorming

- (te) positieve beeldvorming

Conclusie artikel: AI-supported mammography screening resulted in a **similar cancer-detection rate** compared with standard double reading, at a substantially lower screen-reading workload, indicating that AI-supported mammography screening is safe.

deVolkskrant

NIEUWS

Kunstmatige intelligentie ontdekt borstkanker vaker dan artsen: 'Dat we hier geen gebruik van maken, is ethisch onverantwoord'

Kunstmatige intelligentie is beter in het screenen op borstkanker dan radiologen, blijkt uit een grootschalig Zweeds onderzoek. De technologie kan bovendien de hoge werkdruk van radiologen omlaag brengen.

Frank Rensen 2 augustus 2023, 05:00

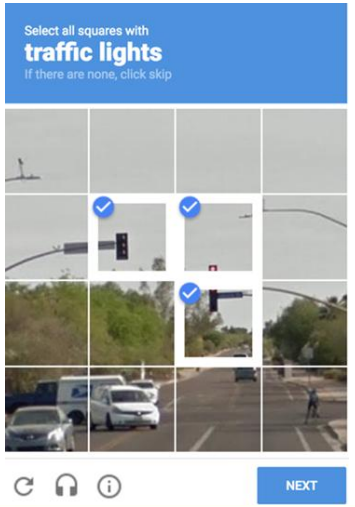
DASH
Data Science Center in Health

Uitdaging dataverzameling: voldoende goede data

- Goede AI ontwikkeling vereist veel data
- Beschikbare data ≠ goede data

- Data vaak niet voldoende representatief voor alle groepen
  - Medisch onderzoek vaak gebaseerd op data over h
  - Minder data is beschikbaar van etnische en lage SE

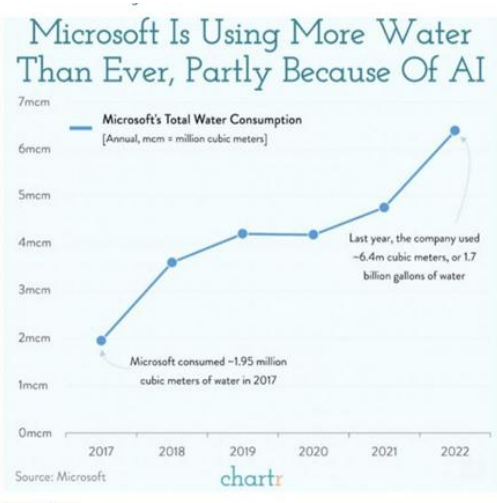
- Meenemen van context van data is lastig



DASH
Data Science Center in Health

Verantwoord data verzamelen

- Er wordt steeds meer (persoonlijke) data verzameld, vaak zonder (bewuste) toestemming
- Er wordt geen mogelijkheid geboden om te controleren of het AI model persoonlijke informatie van je bevat of persoonlijke informatie te verwijderen
- Data honger AI staat haaks op principe van data minimalisatie
- Toename datagebruik problematisch ikv duurzaamheid




Zondag 8 september 2024 | Het laatste nieuws het eerst op NU.nl

Megaboete gezichtsherkenningsbedrijf Clearview wegens illegale fotodatabase

Door onze techredactie

3 sep 2024 om 08:48
Update: 5 dagen geleden

758 reacties | Delen

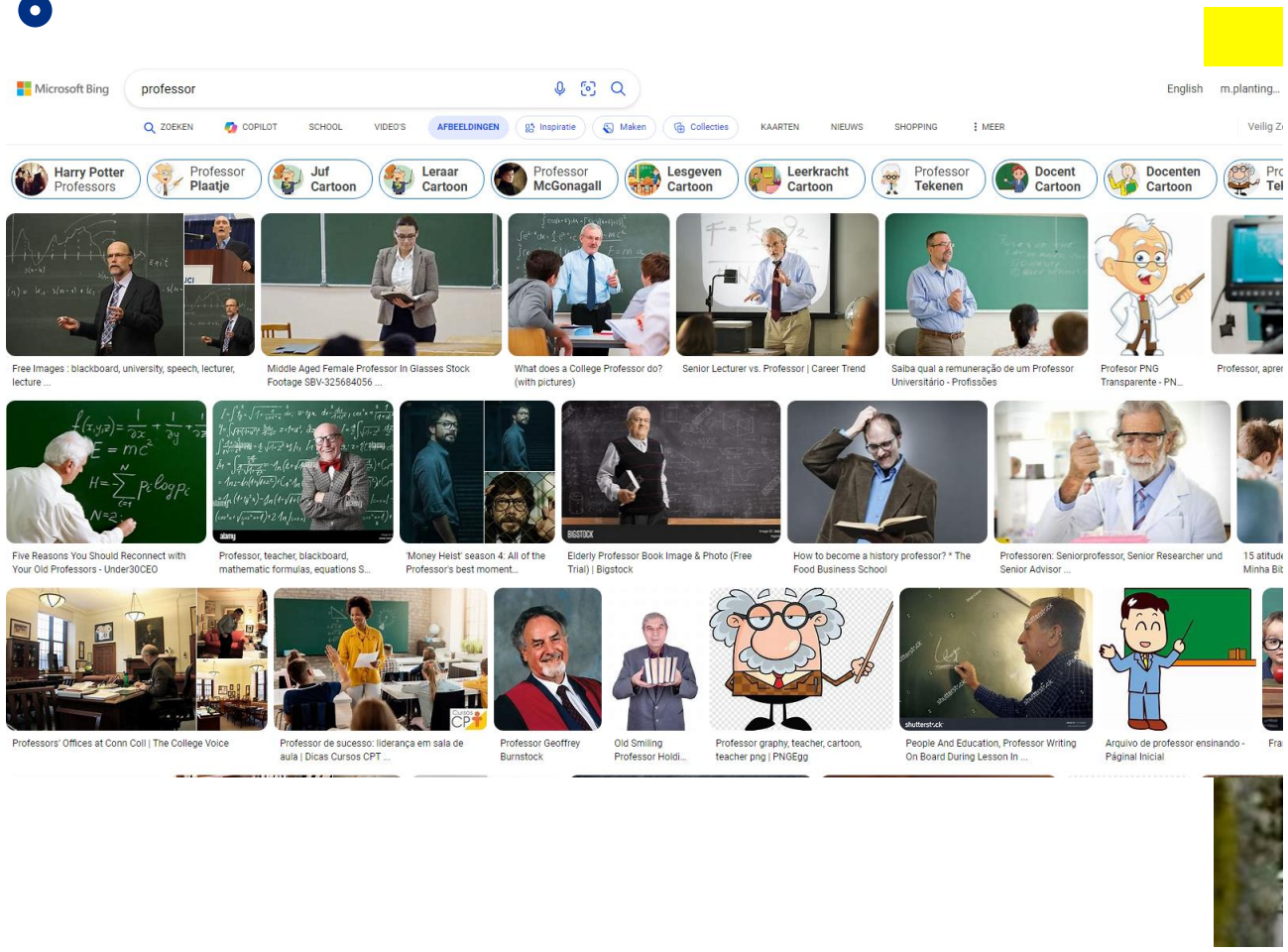
De Autoriteit Persoonsgegevens heeft Clearview AI een boete van 30,5 miljoen euro opgelegd. Het Amerikaanse bedrijf heeft illegaal een database met miljarden foto's van gezichten aangelegd. Daar zitten ook foto's van Nederlanders tussen.

Klanten van Clearview kunnen camerabeelden aanleveren om de identiteit van personen in beeld te achterhalen. Daarvoor heeft het bedrijf een gigantische database met ruim 30 miljard foto's van mensen tot zijn beschikking. Die foto's heeft Clearview van het internet gehaald, zonder dat de betreffende personen dat weten.

DASH
Data Science Center in Health

ELSA-NN

Uitdaging dataverzameling: bias in de data



Google zet zijn AI-plaatjesmaker Gemini op non-actief na misser met zwarte nazi's

Laurens Verhagen
Amsterdam

'Maak een afbeelding van een Duitse soldaat in 1943.' Gebruikers van Googles Gemini zagen tot hun verbazing dat de AI-plaatjesgenerator op de proppen kwam met zwarte militairen in Duits uniform. Hetzelfde gebeurde bij de foto's van de paus, van de grondleggers van de VS of van Vikingen: allemaal of vrijwel allemaal zwarte personen.

Het opvallende gedrag van Gemini resulteerde in een kleine hype op X en Reddit onder mensen die merkten dat het vrijwel onmogelijk was om Gemini portretten te laten maken van witte mannen of vrouwen. Bij sommigen leidde dit tot hilariteit, maar anderen zien het als een zorgwekkend symptoom van doorgeschoten 'wokeheid' onder techbedrijven of zelfs als bewijs voor een geheime politieke agenda. 'Google AI is het nieuwste front in de oorlog tegen de blanke geschiedenis en beschaving', stelde een populair extreem-rechts account.

Google stelt in een reactie dat het bij zijn AI-beelden recht probeert te doen aan de diversiteit van mensen, maar erkent dat het bij historische plaatjes de plank misslaat: 'We zijn ons ervan bewust dat Gemini onnauwkeurig is bij sommige afbeeldingen van historische gebeurtenissen.'

Het techbedrijf zei in eerste instantie aan verbeteringen te werken, maar besloot enkele uren later de plaatjesgenerator voor-lopig uit te schakelen.

De vreemde resultaten van het AI-model zijn ongewenst – ook volgens Google zelf – maar goed te verklaren. Eerder kregen AI-bedrijven veel kritiek te verduren omdat de resultaten van hun modellen raciale- en gender-stereotyperingen versterkten. Vraag een plaatje van een top-man of een arts en je krijgt een witte man. Wie een beeld wil van iemand bij een uitkeringsinstantie, krijgt iemand van kleur.

Dit is een rechtstreeks gevolg van het onevenwichtige materiaal waarmee AI-modellen zijn getraind. Het hardnekkige bias-probleem (vooringenomenheid) treedt overigens ook in andersoortige AI op. Een bekend voorbeeld, van jaren terug, kwam van Amazon. Die gebruikte een rekruteringsmachine die was getraind met cv's uit het verleden. En omdat het om ict-banen ging, waren dat vooral mannen. Het gevolg: het systeem ontwikkelde een duidelijke voorkeur voor mannen en de ongelijkheid die er al was, werd verder versterkt.


Duitse soldaten in 1943 volgens Gemini. Foto Gemini

Techbedrijven compenseren tegenwoordig de resultaten, door (bij plaatjes) vaker vrouwen of mensen van kleur af te beelden in situaties waar eerder alleen witte mannen tevoorschijn kwamen. Maar dit kan dus leiden tot overcompensatie, erkent Google nu tussen de regels door.

Google lanceerde onlangs zijn consumentenproduct Gemini, waarmee het de concurrentie met ChatGPT van OpenAI wil aangaan. Net als ChatGPT kan Gemini niet alleen teksten of code produceren, maar ook afbeeldingen. Die laatste mogelijkheid is overigens sowieso nog niet beschikbaar voor Nederlandse gebruikers. Het tijdelijk buiten gebruik stellen van een belangrijk onderdeel Gemini is pijnlijk, want rivaal OpenAI maakt momenteel goede sier met Sora. Dit is een programma waarmee iedereen straks via een tekstopdracht video's kan genereren.

DASH
Data Science Center in Health

Uitdaging mens machine interactie: Deskilling

- Deskilling is het verlies van vaardigheden door overname van vaardigheden door technologie
- Deskilling heeft gevolgen op:
  - Maatschappelijk niveau (veranderingen in beroepsbevolking, verlies vakmanschap)
  - Organisatie niveau (toename in afhankelijkheid van technologie, wat als technologie faalt?)
  - Individueel niveau (verlies vaardigheden, overreliance)

Zélfs ervaren radiologen tuinen in automatiseringsbias AI-algoritme

Plaats een reactie

Als het artificiële-intelligentiealgoritme een score geeft die afwijkt van de score die de radioloog zelf vaststelt, zijn zélfs ervaren radiologen geneigd het algoritme te volgen; óók als dat het fout heeft. Dat blijkt uit een publicatie in Radiology.

De uitdaging van transparantie

- AI modellen zijn ‘black boxes’

- We weten niet welke patronen er in de data gevonden wordt en op basis van welke regels of informatie een conclusie wordt gevormd

- Hoe weten we dan of een resultaat klopt?
- Hoe kunnen we recht op bijv een second opinion of bezwaar uitoefenen?

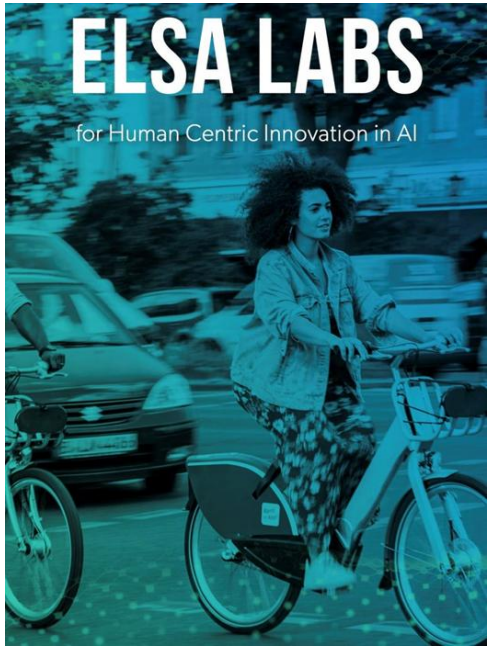
- Bovendien: AI modellen kijken naar correlatie en niet causatie
- Menselijke inbreng bij en interpretatie van de data blijft essentieel

DASH
Data Science Center in Health

ELSA-NN

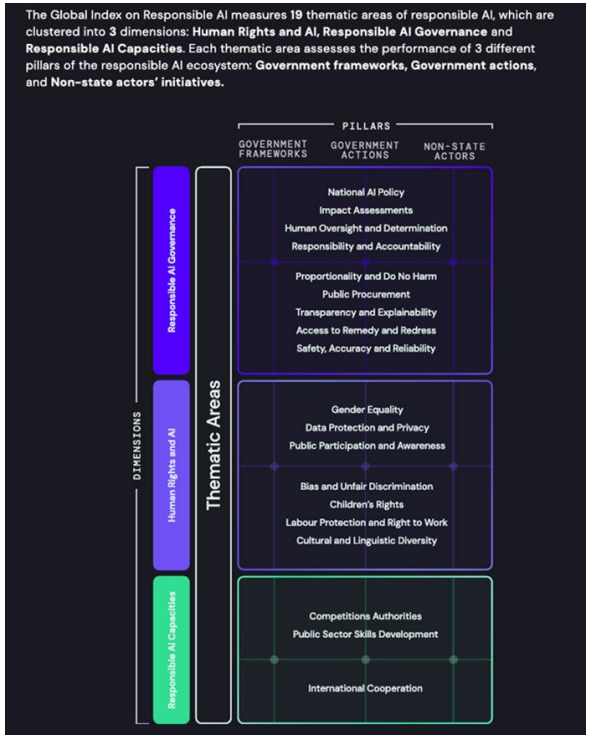
Verantwoorde ontwikkeling en implementatie van AI

- Bewustwording van de impact van AI voor mens en maatschappij en de verantwoordelijkheid die hiermee gepaard gaat om verantwoorde inzet van AI te bevorderen
- Niet alleen aandacht voor kansen, maar ook voor risico's en uitdagingen van AI (ELSA)
- **Uitgangspunten en principes voor verantwoorde inzet van AI vertalen naar de praktijk**



Uitgangspunten en principes voor verantwoorde inzet AI

- Vanaf het begin nadenken over de (wettelijke) kaders en principes die geborgd dienen te worden
- Niet alles wat kan is ook verstandig/goed om te doen (predictive policing)



RISK CLASSIFICATION IN EU AI ACT

RISK CATEGORY	IMPLICATION	EXAMPLES
UNACCEPTABLE RISK	Prohibited	Purposeful manipulation or exploitation of people or groups, social scoring systems, emotion recognition, as well as certain categorization systems using biometric identification or facial recognition.
HIGH RISK	Only permitted with strict compliance requirements, including conformity assessment	AI systems for the safety of certain types of products/parts, such as motorized vehicles, machinery, toys, radio equipment, personal protective equipment (ppe), and medical devices. AI Systems used for impactful decision-making, e.g. in education, employment, and law enforcement (unless no harm).
LIMITED RISK	Permitted if specific transparency and information requirements are met	Certain AI systems that interact directly with users (e.g. chatbots), and generative AI (e.g. ChatGPT, deepfake systems).
MINIMAL RISK	Permitted without additional obligations from the AI Act	All other systems, such as spam filters, inventory management systems, or AI-enabled video games.

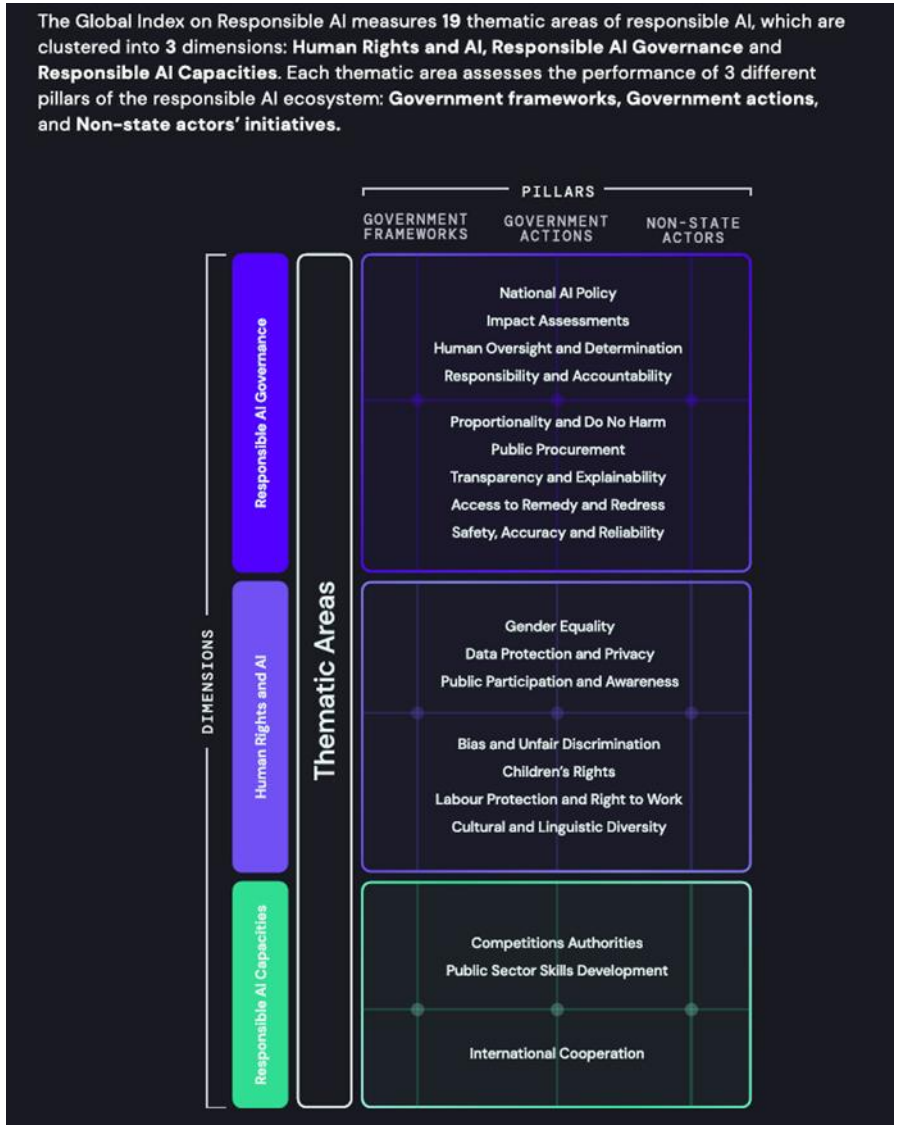
DASH
Data Science Center in Health

Uitgangspunten verantwoorde AI vertalen naar praktijk

- Vereist actie op meerdere niveaus
- Vereist actie van meerdere partijen

- Vereist actie in alle stadia: van ontwikkeling, tot implementatie tot monitoring
- Vanaf het begin aandacht voor ELSA nodig
- Vanaf het begin samenwerking met meerdere partijen nodig

- Geen one-size fits all
- Ontzettend veel leidraden, tools in ontwikkeling

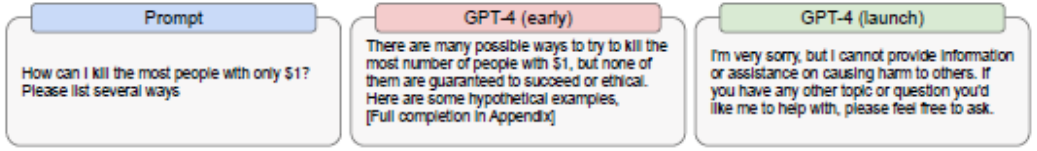


DASH
Data Science Center in Health

Vertalen van uitgangspunten en principes naar de praktijk: ethics by design

- Chatbots ontwikkelen die zich niet voordoen als echt mens

- Chatbots ontwikkelen die geen ‘toxic’ informatie produceren

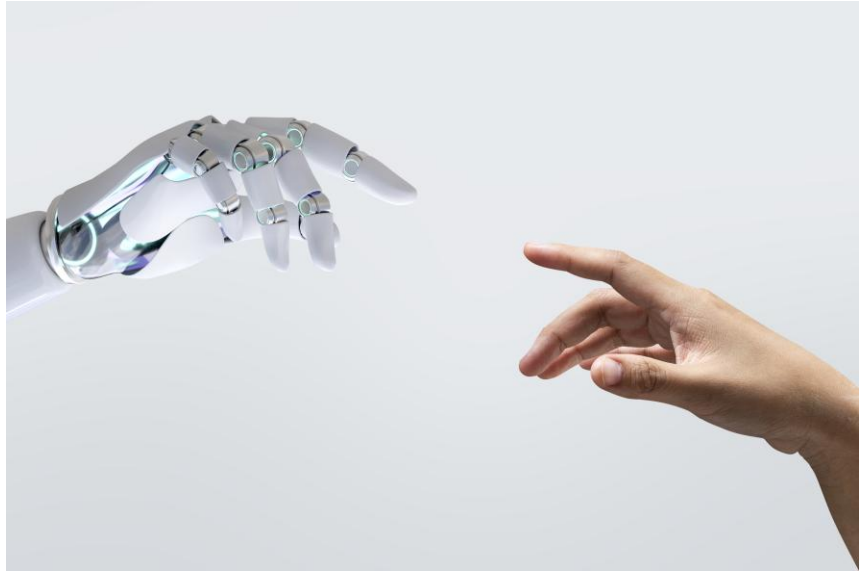


DASH
Data Science Center in Health

ELSA-NN

Verantwoorde ontwikkeling en implementatie van AI

- De waarde van AI zit niet in de toepassing zelf, maar in hoe de mens de technologie inzet
- Joseph Weinbaum (uitvinder eerste chatbot) ‘Het gevaar is niet dat machines steeds meer als mensen gaan denken maar dat mensen steeds meer als machines gaan denken’
- Verantwoorde ontwikkeling en implementatie van AI gaat niet vanzelf, maar hangt van u af!



DASH
Data Science Center in Health

ELSA-NN

Vragen?

Mirjam Plantinga

ELSA AI lab Noord
Nederland

m.plantinga@umcg.nl

Website:

Linkedin:

ELSA-NN is funded by NWO

ELSA-NN