

Unesco's
Aanbeveling over **De**
ethiek van
kunstmatige
intelligentie



De AI-Aanbeveling in een notendop

De Unesco-Aanbeveling over de ethiek van kunstmatige intelligentie (AI) is een internationale overeenkomst tussen 193 landen, waaronder ook Nederland.

De Aanbeveling spoort overheden, organisaties, bedrijven en individuen aan om AI op een ethische manier te ontwerpen en gebruiken. Zo kan kunstmatige intelligentie bijdragen aan vrede, duurzaamheid, en de bescherming van mensenrechten.

AI wordt op veel terreinen over landsgrenzen heen ingezet. Het is belangrijk dat de hele wereldbevolking kan meedenken over hoe we kunstmatige intelligentie gebruiken.



Aanbeveling over de
ethiek van kunstmatige
intelligentie, aangenomen
op 23 november 2021

Wat is kunstmatige intelligentie?

→ Kunstmatige intelligentie ('AI') systemen zijn in staat om informatie te verwerken op een manier die lijkt op menselijke denkprocessen, zoals redeneren, leren, waarnemen, voorspellen, plannen of besturen.

Waarom hebben we wereldwijde afspraken nodig over AI?

→ Kunstmatige intelligentie is nu al overal. Het heeft enorme potentie voor het vergroten van ons welzijn en voor het bijdragen aan duurzame ontwikkeling. Zo kan AI worden ingezet om de opbrengst van gewassen te voorspellen, waardoor we meer voedselzekerheid creëren. Maar het gebruik van deze technologieën leidt ook tot ethische bezwaren: het kan bijdragen aan het vergroten van genderongelijkheid, of de privacy van mensen aantasten, of het verspreiden van desinformatie vergemakkelijken. Omdat deze onderwerpen grenzen overstijgen, is een internationaal perspectief nodig op het ethisch ontwikkelen en gebruiken van AI-systemen. De in 2021 aangenomen Aanbeveling van Unesco vormt een dergelijke mondiale overeenkomst.

Wat is het doel van de AI-Aanbeveling?

→ De Aanbeveling is opgesteld om het welzijn van mens en milieu te dienen door te zorgen dat tijdens de ontwikkeling, werking en het gebruik AI-systemen 'mensgericht' zijn. Het bevat regels die waarden en principes praktisch toepasbaar maken. De Aanbeveling geeft landen concrete taken op elf beleidsterreinen, waaronder gezondheid, milieu en gendergelijkheid. Zo zorgt de tekst ervoor dat mensenrechten en fundamentele vrijheden worden beschermd in alle stadia van de levenscyclus van een AI-systeem.

Waarin onderscheidt de AI-Aanbeveling zich van andere adviezen?

→ De Aanbeveling focust op onderwerpen die tot nu toe minder aandacht kregen in discussies over kunstmatige intelligentie, zoals (gender) discriminatie, duurzaamheid, en de impact van AI-toepassingen op gemarginaliseerde groepen. Verder is de Aanbeveling de eerste wereldwijde afspraak van zijn soort. Mede daarom legt de tekst nadruk op de AI-behoefte van lage- en middeninkomenslanden. Bovendien presenteert het document de ontwikkeling van mensgerichte AI als een taak voor de gehele maatschappij, die veel verder gaat dan alleen het reguleren van technologie. Iedereen mag meepraten over de rol die kunstmatige intelligentie in onze samenlevingen dient te hebben.

Thema's

De Aanbeveling biedt een ethisch kader op elf verschillende thema's. Nationale overheden hebben de taak de Aanbeveling met input vanuit de samenleving te vertalen naar wet- en regelgeving. Hieronder zijn vijf thema's uitgelicht om een indruk te geven van de terreinen die de Aanbeveling beslaat.



● Gezondheid

Kunstmatige intelligentie wordt al ingezet bij de vroege opsporing van ziekten. Ook kunnen robots van betekenis zijn in het aanpakken van het personeelstekort in de zorg. Maar het gebruik van AI brengt ook risico's met zich mee, bijvoorbeeld bij het delen van digitale patiëntendossiers.

De Aanbeveling:

- Zet aan tot richtlijnen voor mens-robot interacties die mensenrechten in acht nemen en kwetsbaren beschermen
- Roept op tot wereldwijde samenwerking in het gebruik van AI-systemen voor het verminderen van gezondheidsrisico's en virusuitbraken

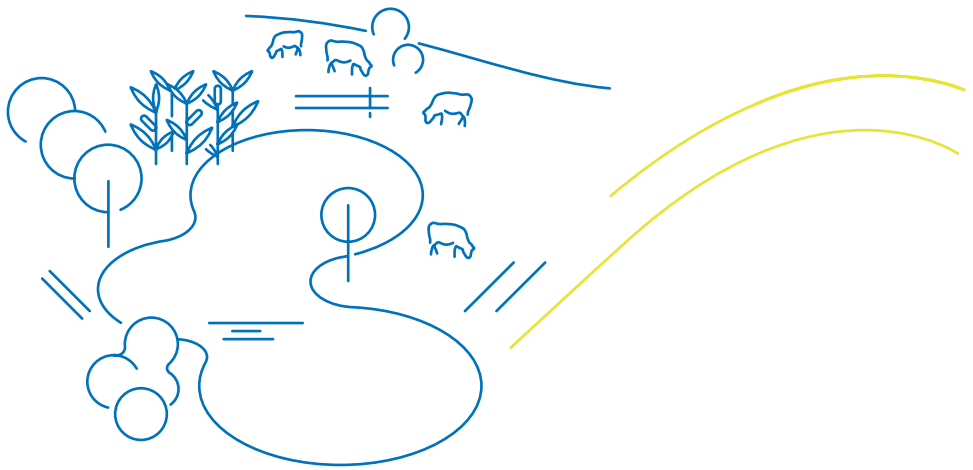
● Cultuur

AI heeft nu al een enorme impact op hoe we cultuur ‘maken’ en hoeveel consumenten willen betalen voor cultuurgoederen. Ook bepalen algoritmen mede onze culturele smaak, bijvoorbeeld door een selectie te maken in de muziek die we luisteren.

De Aanbeveling:

- Vraagt om samenwerking met technologiebedrijven om een divers aanbod van culturele uitingen te garanderen, ook in ‘kleinere’ talen
- Stimuleert nieuw onderzoek aangaande AI en intellectueel eigendom om te bepalen of, en hoe, AI(-geassisteerde) werken moeten worden beschermd door intellectuele rechten
- Waarschuwt voor de impact die natuurlijke taalverwerkingssystemen, zoals automatische vertalingen en stemassistenten, kunnen hebben op de nuances van menselijke taal en culturele expressie





● Het milieu

Kunstmatige intelligentie draagt bij aan de bescherming van het milieu, bijvoorbeeld door biodiversiteit te monitoren met drone- en satellietbeelden. Maar AI-systemen kunnen ook een negatieve ecologische impact hebben, wanneer ze voor hun berekeningen veel energie verbruiken.

De Aanbeveling:

- Zet aan tot het monitoren van (in)directe milieueffecten bij het creëren en gebruiken van AI-toepassingen, en de naleving van bestaande milieuwetgeving
- Vraagt om het verbieden van AI-systemen waarvan kan worden ingeschat dat zij een disproportioneel negatief effect op het milieu hebben
- Stelt dat de ontwikkeling en acceptatie van technologieën die rampen helpen voorkomen en hun impact kunnen verkleinen, moet worden gestimuleerd



● Gender

AI-systemen kunnen vooroordelen versterken, zoals wanneer vertaalsoftware van neutrale input ('de dokter') genderspecifieke output ('el doctor') maakt. En als een bedrijf in een typische 'mannensector' kandidaten werft op basis van een bestaande dataset met cv's, dan zal het algoritme weinig vrouwen kiezen.

De Aanbeveling:

- Vraagt om het opnemen van een genderactieplan in het digitale overheidsbeleid
- Noemt het belang van investeringen die de loopbanen van vrouwen in de exacte wetenschappen en ICT bevorderen
- Stelt dat moet worden gewerkt aan het tegengaan van genderstereotypen en discriminerende vooroordelen in AI-systemen, bijvoorbeeld door het creëren van diverse, open en betrouwbare ('gouden standaard') datasets



● Onderwijs en onderzoek

De Aanbeveling presenteert onderwijs als de manier om AI-geletterdheid te promoten bij kinderen en volwassenen – zodat zij de technologie snappen, en de menselijke impact ervan begrijpen. Maar juist e-learning levert op scholen ook ethische dilemma's op: bij het gebruik van de persoonsgegevens van leerlingen moet goed worden overwogen wanneer en hoe kunstmatige intelligentie wordt ingezet.

De Aanbeveling:

- Roept op tot de ontwikkeling van educatieve programma's over AI en de impact van kunstmatige intelligentie op mensenrechten voor kinderen
- Vraagt om meer AI-onderzoek waarbij 'exacte' wetenschappers samenwerken met onderzoekers uit de sociale en geesteswetenschappen

Aan de slag voor de ontwikkeling van ethische en mensgerichte AI!

De Aanbeveling van Unesco wil kunstmatige intelligentie die in dienst staat van samenlevingen. De Nederlandse overheid kan hiervoor zorgen door nieuwe wetten en regels in te stellen die zorgen dat AI-toepassingen ‘mensgericht’ zijn en op ethische wijze worden ontwikkeld en gebruikt. Ook ontwikkelaars van apps kunnen met de Aanbeveling aan de slag, bijvoorbeeld door te zorgen dat culturele uitingen ook worden aangeboden in ‘kleinere’ talen, of door te voorkomen dat hun software genderstereotypen bevestigt.

Maar AI is overal en heeft daardoor nu al een enorme impact op de samenleving. Daarom moeten alle burgers inspraak hebben in hoe we deze technologieën nu en in de toekomst willen inzetten. Ook mensen die niet werken in de zorg, de cultuursector of het onderwijs willen tenslotte leven in een maatschappij waarin AI een positieve impact heeft op de mensheid. Er is werk aan de winkel!

Colofon

Uitgegeven door

Nederlandse Unesco Commissie
Prins Willem-Alexanderhof 20
2595 BE Den Haag

Contact

Nederlandse Unesco Commissie
T 070 - 3315484
info@unesco.nl
www.unesco.nl

Redactie

Jon Verriet (bewerking en vertaling),
Anne Diessner, Jeannine Hausmann,
Maximilian Müngersdorff, Davor Orlic

Status

Augustus 2023

Copyright

De teksten van deze publicatie vallen onder de Creative Commons-licentie Naamsvermelding-GelijkDelen 4.0 Internationaal (CC BY-SA 4.0). www.creativecommons.org/licenses/by-sa/4.0/

Design en opmaak

Panatom en teambart

De oorspronkelijke, Engelstalige versie van deze brochure werd uitgebracht door de nationale Unesco Commissies van Nederland, Duitsland en Slovenië, en is beschikbaar via de volgende link: <https://ircai.org/project/the-unesco-recommendation-on-the-ethics-of-ai/>

Met steun van



unesco

Centre
Under the auspices
of UNESCO



International Research Centre
of Artificial Intelligence
under the auspices of UNESCO