



# Ethiek van kunstmatige intelligentie (AI)

---

Overwegingen van  
EU, UNESCO & COMECE





Europese Commissie

---

# Witboek over kunstmatige intelligentie - *een Europese benadering op basis van excellentie en vertrouwen (19 februari 2020)*



Brussel, 19.2.2020  
COM(2020) 65 final

**WITBOEK**

**over kunstmatige intelligentie - een Europese benadering op basis van excellentie en  
vertrouwen**



## 'Europa ligt achter op het gebied van kunstmatige intelligentie'

Door NU.nl

18 jun 2018 om 12:43  
Update: 5 jaar geleden

 41 reacties

 Delen

**De Europese Unie dreigt achter te blijven liggen op het gebied van kunstmatige intelligentie, waarschuwen meerdere onderzoekers.**

# Nieuwe kansen, vertrouwen als voorwaarde

---

*“Europa (kan) een wereldleider worden op het gebied van innovatie in de data-economie en de toepassingen ervan.”*

*“Europa in staat stellen een van de meest aantrekkelijke, veilige en dynamische data-agile economieën ter wereld te worden.”*

## Gemeenschappelijke Europese benadering

- voldoende schaalgrootte
- voorkomen versnippering

## Voordelen voor burgers, bedrijven en openbare instellingen

- efficiënter, goedkoper, duurzamer, veiliger.

# Eerste definitie (2018): Kunstmatige intelligentie...

---

verwijst naar

- systemen die intelligent gedrag vertonen door hun omgeving te analyseren en – in zekere mate zelfstandig – actie te ondernemen om specifieke doelstellingen te verwezenlijken.
- Op AI gebaseerde systemen kunnen uitsluitend uit software bestaan en actief zijn in de virtuele wereld (bijvoorbeeld stemgestuurde assistenten, software voor beeldanalyse, zoekmachines en systemen voor spraak- en gezichtsherkenning), maar AI kan ook in hardwareapparaten worden geïntegreerd (bijvoorbeeld geavanceerde robots, zelfrijdende auto's, drones of toepassingen van het internet der dingen)."

# Aanscherping definitie (2019)

---

“Systemen op basis van kunstmatige intelligentie (AI) zijn:

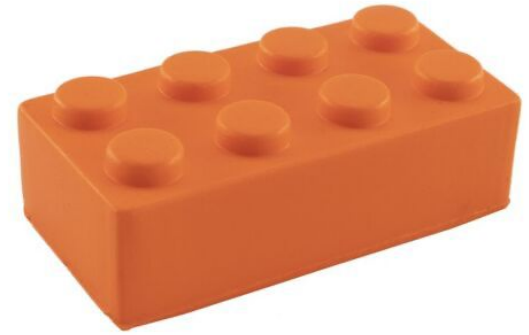
door mensen ontworpen softwaresystemen (en mogelijk ook hardware-systemen)

- die, met een complex doel, in de fysieke of digitale dimensie werken
- door via gegevensverzameling hun omgeving waar te nemen,
- de verzamelde gestructureerde of ongestructureerde gegevens te interpreteren,
- te redeneren op basis van de uit deze gegevens verkregen kennis of de verkregen informatie te verwerken
- en te beslissen met welke handeling(en) het gestelde doel het best kan worden bereikt.
- AI-systemen kunnen gebruik maken van symbolische regels of een numeriek model leren en kunnen hun gedrag ook aanpassen door te analyseren welke invloed hun eerdere handelingen op de omgeving hebben.”

# Bouwstenen

---

- ECOSYSTEEM van EXCELLENTIE
- ECOSYSTEEM van VERTROUWEN
  - garantie naleving EU-regels ter bescherming grond/consumentenrechten
  - een “mensgerichte” benadering
    - *“Vertrouwen kweken in mensgerichte kunstmatige intelligentie”*







-EXCELLENCE-

# Ecosysteem van Excellentie

---



- samenwerking van lidstaten en internationale gremia
  - bundelen van investeringen
  - aandacht voor maatschappelijk en ecologisch welzijn
  - UNESCO, OESO, VN
- centrum voor onderzoek, innovatie en deskundigheid voor AI
  - bundelen van netwerken
  - bundelen van expertise
- bevorderen vaardigheden van burgers, bedrijven en instellingen
  - m.n. gezondheidszorg en vervoer
- verbeteren toegang tot en het beheer van data
  - verbeteren toegang nieuwe data
  - verantwoord beheer volgens FAIR-beginselen (Findable, Accessible, Interoperable, Reusable)

# Eerbiediging grondrechten

---

- menselijke waardigheid
- pluralisme
- inclusie
- non-discriminatie
  - geslacht, ras, etnische afkomst, religie of overtuiging, handicap, leeftijd, seksuele geaardheid
- bescherming privacy
- bescherming persoonsgegevens\*
- vrijheid van meningsuiting
- vrijheid van vergadering
- recht op eerlijk proces en consumentenbescherming

# Risico's

---

*sectoren onderscheiden waar aanzienlijke risico's ontstaan met kans op schade*

## waarborgen van grondrechten

- het complexe, zelflerende vermogen (autonoom gedrag) kan leiden tot gebrek aan transparantie hoe AI tot beslissing komt
- lerend vermogen hangt af van de input
- koppelen van databases

## aansprakelijkheid?

- zelflerende producten

## schade

- materieel
  - veiligheid, gezondheid, schade aan bezit
- immaterieel
  - schending privacy en menselijke waardigheid, beperking vrijheid van meningsuiting, discriminatie

# Ecosysteem van Vertrouwen



# Ecosysteem van Vertrouwen

---

- menselijk toezicht
  - voorkomen ondermijning van menselijke autonomie >> passende tussenkomst
- robuust en nauwkeurig
  - betrouwbaar, reproduceerbaar, veilig bij aanvallen, omgaan met fouten/inconsistenties
- privacy en veilig databeheer
- transparantie
  - welke (trainings)data, welk algoritme?
  - over mogelijkheden en beperkingen KI
  - wanneer KI, wanneer humaan contact
- extra eisen bij biometrische identificatie
- \*diversiteit, non-discriminatie en billijkheid
- \*maatschappelijk en ecologisch welzijn



# Deskundigengroep: (2019)

---

## Uitgangspunten:

- betrouwbare KI: wettig, ethisch, robuust
- traceerbaarheid, controleerbaarheid
- openheid, onderzoek en innovatie, ontwikkeling KI-ethiek

## Ethische beginselen

- respect voor menselijke autonomie, preventie van schade, rechtvaardigheid, verantwoording

## Kwetsbare groepen en specifieke risico's

- kinderen, mensen met een beperking, kansarmen, risico op uitsluiting
- mensen herkennen en volgen of beoordeling van burgers
- verborgen KI-systemen
- dodelijke autonome wapensystemen



# COMECE

Commission of the Bishops' Conferences of the European Union







# Eerste reactie op witboek (juni 2020)

---

! gezamenlijk optrekken in Europa in faciliteren, veiligheid, recht en regulatie

! blij met uitgangspunten

- menselijke waardigheid
- bescherming van privacy

**“mensgerichte kunstmatige intelligentie”**

*menselijk leven kent een persoonlijke dimensie en een gemeenschappelijke dimensie*

- *“algemeen welzijn” dienen*
- *“leven van alle menselijke wezens” dienen*

# Oproep 1: verheldering van definitie

---

## Onderschrijven van juridische definitie die

- flexibele ruimte laat voor technische vooruitgang
- nauwkeurig genoeg voor rechtszekerheid

## Verduidelijking van onderscheid menselijke persoon en KI

### *Christelijk perspectief op de menselijke persoon*

- *mens is kwalitatief anders dan andere wezens*
  - *met transcendentale waardigheid, intelligentie en vrijheid*
  - *daarom in staat tot moreel handelen*
- *KI is niet op dezelfde wijze “vrij”*
  - *moreel handelen kan niet beoordeeld worden zoals bij mensen*

# Oproep 2: Ecosysteem van excellentie

---

Excellentie kan alleen als alle stakeholders betrokken worden in ontwikkeling

>> *dus ook de Kerken*

*Ontwikkeling van uitgebreid ethisch discours*

- *relatie tussen mens en KI*
- *invloed KI op vormgeven van samenleving*
  - *informatica, wiskunde, filosofie, ethiek; KUNST?*
- *oprichten van structuren van interdisciplinaire ethische reflectie bij politieke besluitvorming*

# Oproep 3: Ecosysteem van vertrouwen

---

Risico-inventarisatie o.b.v. 2 criteria:

1. sector met significant risico >> *bestaat er zo'n onderscheid?*
2. toepassing zorgt voor verhoogd risico

*voorstel >> per toepassing een risico-analyse maken*

# Oproep 4: voorwaarden

---

- betrouwbaarheid
  - duidelijke (internationale) wetgeving, die zo nodig geupdate wordt
  - transparantie, evt omkering van bewijslast
- veiligheid
  - menselijk toezicht
- openheid m.b.t. algoritmes
- expliciet beschermen van kinderen
- bescherming van persoonlijke gegevens
  - zorgen op profileren o.b.v. geloof

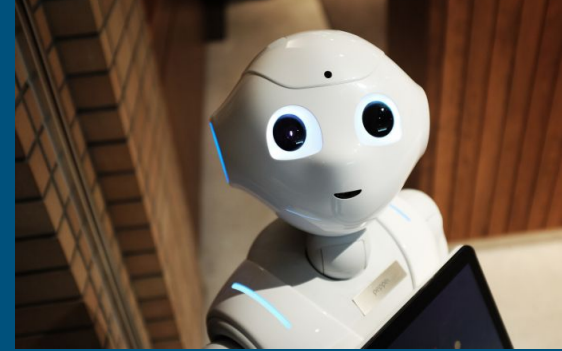
# Oproep 5: specifieke toepassingen

---

- bestrijden van witwassen mag niet leiden tot “controle-samenleving”
  - bijzonder in relatie tot goede doelen en de organisatie van Kerken
  - balans transparantie en privacy en autonomie
- KI in militaire toepassingen
  - oproep tot uitbannen van volledig autonome wapensystemen zonder menselijk toezicht
  - opstarten van internationale onderhandelingen met bindende wetgeving die dodelijke autonome wapensystemen verbiedt
- KI en cyber-security
  - niet alleen nieuwe kansen, maar ook nieuwe kwetsbaarheden
    - snellere en grotere aanvallen op kritische infrastructuur
    - gevolgen voor democratische orde
    - risico's van desinformatie of externe invloed
  - aanbeveling
    - adequate beveiliging, weerbaarheid van kritieke infrastructuren

# Robotisatie (jan 2019)

---



- mogelijkheden, niet per definitie negatief:
  - minimaliseren productie en arbeidskosten
  - verminderen gevaar (politie, defensie, gezondheidszorg)
- “antropologische crisis”: roept vragen op over de identiteit van de mens
  - gelegenheid, geen noodzakelijkheid
  - onrechtvaardigheid vs solidariteit
- **benadrukken van de menselijke persoon en waardigheid, solidariteit**
  - paradox: hoe meer controle door machines, hoe minder handelingsvrijheid en controle
  - “creaturaliteit”: de menselijke persoon - als schepsel van God - heeft vrijheid om overwogen keuzes te maken en als zodanig autonomie.
  - de mens is in staat om orde en betekenis te geven aan de schepping.
- werk als mensenrecht
- rechten van robots?
  - evt “electronic personality” bij hoogontwikkelde robots, vgl dieren.>> idee is verworpen



UNESCO



# Nieuwe ethische vraagstukken door AI

---

“... include, but are not limited to, their impact on **decision-making, employment and labour, social interaction, health care, education, media, access to information, digital divide, personal data and consumer protection, environment, democracy, rule of law, security and policing, dual use, and human rights and fundamental freedoms, including freedom of expression, privacy and non- discrimination.**

[...] These characteristics give AI systems a profound, new role in human practices and society, as well as in their relationship with the environment and ecosystems.”

# Waarden

---

1. respect, bescherming en bevordering van mensenrechten, menselijke waardigheid
  - niemand mag beschadigd raken of onderworpen worden: lichamelijk, economisch, sociaal, politiek, cultureel of mentaal.
2. milieu en ecosysteem
3. beschermen van diversiteit en inclusiviteit
  - ras, kleur, afkomst, sekse, leeftijd, taal, religie, politieke voorkeur, nationaliteit/ethniciteit, sociale afkomst, economische situatie bij geboorte, handicap
4. leven in vreedzame, eerlijke en verbonden samenleving
  - erkenning dat de mens verbonden (“interconnected”) is
  - behoort tot een groter geheel

# Uitgangspunten

---

- proportioneel en transparant
  - menselijk toezicht
  - uitleg en scholing: KI-bewustzijn en geletterdheid
- niet schaden
  - veiligheid en beveiliging & duurzaamheid
    - mens en milieu: ongewenst schade voorkomen en kwetsbaarheden beschermen
- sociale rechtvaardigheid:
  - eerlijkheid en voorkomen discriminatie
  - recht of privacy en gegevensbescherming (uitdrukking van menselijke waardigheid)
- afspraken verantwoordelijkheid en aansprakelijkheid

# Conclusie

---

# Terugblik 1/3

---

- terugblik op de dag
  - zorgtechnologie, ethiek bij AI, antropologie
- ethische afwegingen:
  - menselijke waardigheid (**autonomie, privacy**), veiligheid (wel doen/niet schaden), betrouwbaarheid, duurzaamheid, rechtvaardigheid en solidariteit
  - vooringenomenheid en discriminatie
  - menselijk toezicht
  - behoefte aan regulatie en transparantie



# Terugblik 2/3

---

## antropologie

- eigene van de mens en het menszijn
  - rationaliteit:
  - liefhebben (zelfgave)
  - bewustzijn
  - vrije wil
  - bezield door God
- mens als bezield lichaam
- mens als relationeel wezen, geroepen tot vrijwillige zelfgave
- mens geroepen tot werk en creativiteit
- gebrokenheid
- offer



# Terugblik 3/3

---

- dus AI is niet hetzelfde als de mens
- wel kan het taken krijgen die risico's met zich meebrengen
  - input data
  - algoritmes
  - menselijk toezicht
- ethische toepassingen en regulatie





Dank voor uw aandacht

EU



“Aangezien de mens een  
doorlaat van liefde in actie is,  
kan alleen het delen de mens  
zijn menselijke waardigheid  
tonen.

Joanika Ring

