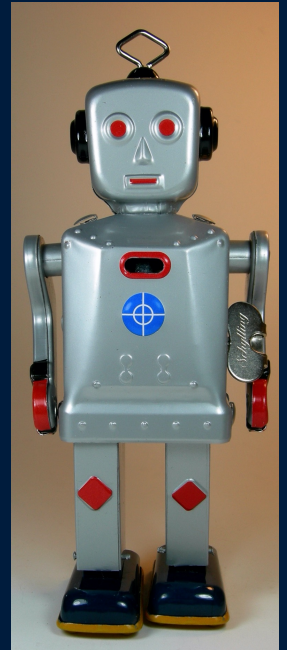


# Regels voor robots

Filosofisch café Amersfoort  
25 september 2023

Dr. Katleen Gabriels  
[k.gabriels@maastrichtuniversity.nl](mailto:k.gabriels@maastrichtuniversity.nl)



# Overzicht

- 1. Technologie en waarden**
- 2. Wat is artificiële intelligentie?**
- 3. Voordelen: altijd beter?**
- 4. Ethische problemen**
- 5. Slotbeschouwingen**

# 1. Technologie en waarden

<https://www.youtube.com/watch?v=QmA0nC4tZVM>

**Elk product dat door mensen gemaakt is, is  
'made with morality'**

[https://www.youtube.com/watch?v=YJjv\\_OeiHmo](https://www.youtube.com/watch?v=YJjv_OeiHmo)

<https://www.indy100.com/article/disturbing-anti-homeless-architecture-ryan-brown-twitter-8189081>

<https://www.youtube.com/watch?v=TNwhIHqM60g>



## 2. Wat is artificiële intelligentie?

# Wat is intelligentie?

- Max Tegmark: “het vermogen om complexe doelstellingen te verwezenlijken”

Logisch denken

Inzicht

Planning

Emotionele ontwikkeling

Zelfbewustzijn

Oplossen van problemen

...

- Sociaal
- Analytisch
- Wiskundig
- Emotioneel
- ...
- Motivatie en creativiteit als onderdelen van intelligentie

# Wat is artificiële intelligentie?

A Proposal for the  
**DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE**

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

# Machines intelligent maken (1)

- De mens heeft *zicht* -> AI: beeld- en patroonherkenning
- De mens heeft *taal* -> Natural Language Processing (NLP)
  - Bv. Siri, Alexa
  - Stemherkenning
  - Large Language Models, bv. ChatGPT (generatieve AI)
- De mens kan *emoties* herkennen -> emotion recognition algorithms
- Opgelet: onnauwkeurigheden! Taal en emoties zijn lang genegeerd als onderdelen van intelligentie



Schrijf een K3 lied over carnaval.



Verse 1:

Laten we samen carnaval vieren  
Met onze vrolijke kostuums aan  
We dansen door de straten  
Met onze vrienden, elk jaar weer

Refrein:

Kom maar mee, kom maar mee  
Laten we samen dansen  
Kom maar mee, kom maar mee  
Laten we carnaval vieren

Verse 2:

We gooien pepernoten in de lucht  
En vangen ze weer op  
We lachen, we zingen, we dansen  
Want carnaval is hier weer op



[https://www.theguardian.com/technology/2022/oct/25/information-commissioner-warns-firms-over-emotional-analysis-technologies?CMP=Share AndroidApp Other](https://www.theguardian.com/technology/2022/oct/25/information-commissioner-warns-firms-over-emotional-analysis-technologies?CMP=Share_AndroidApp_Other)

# Machines intelligent maken (2)

- Robots zijn ‘containers’ en de *belichaming* van AI
- Zelfrijdende wagen: sensoren, camera’s (‘ogen’), microfoons (‘oren’) om te leren van ervaringen
  - Data = input
  - Gegevens worden geanalyseerd om een beslissing of voorspelling te maken (output)
- Menselijke intelligentie ‘wint’: veel complexer dan machine-intelligentie

## Nieuwe AI-zomer dankzij

- nieuwe technologische ontwikkelingen en applicaties;
- toegenomen opslagcapaciteit, snelheid en rekenkracht;
- en big dataprojecten die immense datasets aanleggen waar algoritmen mee getraind worden:
  - datagebaseerde AI
  - zelflerende algoritmen

<https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>

<https://www.nature.com/articles/nature21056>

[https://www.youtube.com/watch?v=\\_sBBaNYex3E](https://www.youtube.com/watch?v=_sBBaNYex3E)

# Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

Signatures

**33712**

Add your  
signature

Published

March 22, 2023

<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

<https://edition.cnn.com/2023/07/12/tech/elon-musk-ai-company/index.html>



# 3. Voordelen: altijd beter?

# Altijd beter...?

- Optimalisering (effectiviteit)
- Snelheid
- Verzamelen, analyseren en verwerken van een enorme hoeveelheid data
  - Geheugen
- Consistentie
- Objectiever

# 4. Ethische problemen

1. 'Van 'wie' leren zelflerende algoritmen?
2. Personalisering
3. Kwetsbaarheid
4. Privacy

# *1. Van 'wie' leren zelflerende algoritmen?*

<https://www.reuters.com/article/us-newzealand-passport-error-idUSKBN13W0RL>

[https://www.nemokennislink.nl/publicaties/  
hoe-kunnen-we-leren-van-  
discriminerende-algoritmes/](https://www.nemokennislink.nl/publicaties/hoekunnenwelerenvandiscriminerendealgoritmes/)

[https://www.amnesty.nl/actueel/toeslagenaffaire-  
-is-mensenrechtenschending-zegt-amnesty-  
international](https://www.amnesty.nl/actueel/toeslagenaffaire-is-mensenrechtenschending-zegt-amnesty-international)

“Algorithms are opinions embedded in code” (Cathy O’Neil)

<https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report>

## Meer om te vragen :

---

Wat voor auto rijdt max prive? ▼

---

Wie heeft de meeste Grand Prix gewonnen? ▼

---

Wat is het bezit van Max Verstappen? ▼

---

Wat is de leeftijd van Max Verstappen? ▼

---

[Feedback](#)



## 2. *Personalisering*

- Content (Facebook) & search (Google, sinds 2009)
  - Surfgedrag beïnvloedt algoritmen: eerst de ads, daarna het nieuws, daarna de hele 'flow of information' (gradueel proces)
  - Stijgende nieuwsconsumptie via sociale media
  - Om mensen langer op platform te houden
  - Gemanipuleerde algoritmen (bv. FB 2013)
- Distraction by design

- **Behavioral microtargeting**

- O.b.v. informatie over onlinegedrag doelgericht benaderd worden met advertenties
- Alphabet Inc.: Fitbit, Android, YouTube, et cetera

<https://www.buzzfeednews.com/article/alexkantrowitz/how-the-retweet-ruined-the-internet>

<https://www.fastcompany.com/90132364/nest-founder-i-wake-up-in-cold-sweats-thinking-what-did-we-bring-to-the-world>

# ***3. Kwetsbaarheid – ‘onlife’***

✓ Ongewenste e-mail ★

⚠ Items in Ongewenste e-mail worden na 10 dagen verwijderd

Van

Onderwerp

Ontvangen ∨

Vandaag

○ Bol New [Je bent gekozen!](#) Win een Bol.com cadeaukaart! Hoe shopt Nederl... 17:16

○ 📧 WhatsApp [💎❤️Katleengabriels ik was me aan het omkleden 💎❤️ Ik stuur je ...](#) 16:10

T Tinder

[Someone matched with you on Tinder!](#) It's a Match! Someone ma... 7:30

✓ Ongewenste e-mail ★

⚠ Items in Ongewenste e-mail worden na 10 dagen verwijderd

Van

Onderwerp

Ontvangen ∨

Vandaag

B Bol.com [Gefeliciteerd, u komt nu in aanmerking om te winnen](#) Let op: nu t... 12:56

I Iemand probeert je account te hacken 🇸🇰⚠️!

[Actie vereist: Verleng uw Norton-abonnement binnen 72 uur!](#) Er i... 12:58

# Phishing en Vishing

<https://www.ad.nl/binnenland/man-door-dochter-opgelicht-via-whatsapp-je-voelt-je-enorm-bescheten~a7ba0744/>

# Wat is echt?

<https://www.bbc.com/news/world-europe-66877718>



# 4. Privacy

# 5. Slotbeschouwingen

- **Technologie is niet neutraal**
  - Verborgen morele dimensies, ook in computertechnologie
  - Technologie vormt mee de context waarin we ons gedragen en beslissingen nemen (vaak onbewust)
  - ‘We shape technology, technology shapes us’
- **Wie technologie ontwerpt, werkt in het ethische domein**
  - Belang van morele verantwoordelijkheid

## - Ethiek ≠ beslissingsboom

“O, als het eens zo eenvoudig was! ...dat er gewoon duistere lieden op aarde rondliepen die arglistig duistere praktijken bedreven, en dat wij hen alleen maar van de overige mensen hoefden af te scheiden en te vernietigen. Maar de streep die het goede van het kwade scheidt, loopt dwars door het hart van ieder mens. En wie zal er een stuk van zijn hart willen vernietigen?”

*De Goelag Archipel*  
Aleksandr Solzjenitsyn

# Ethics by design

een ethisch gefundeerde AI,  
waarbij ethiek een inherent deel van het ontwerp vormt  
van reactief naar proactief

<https://www.mediawijs.be/nl>

<https://www.medianest.be/>

<https://www.data-en-maatschappij.ai/>



<https://www.rathenau.nl>

Katleen Gabriels

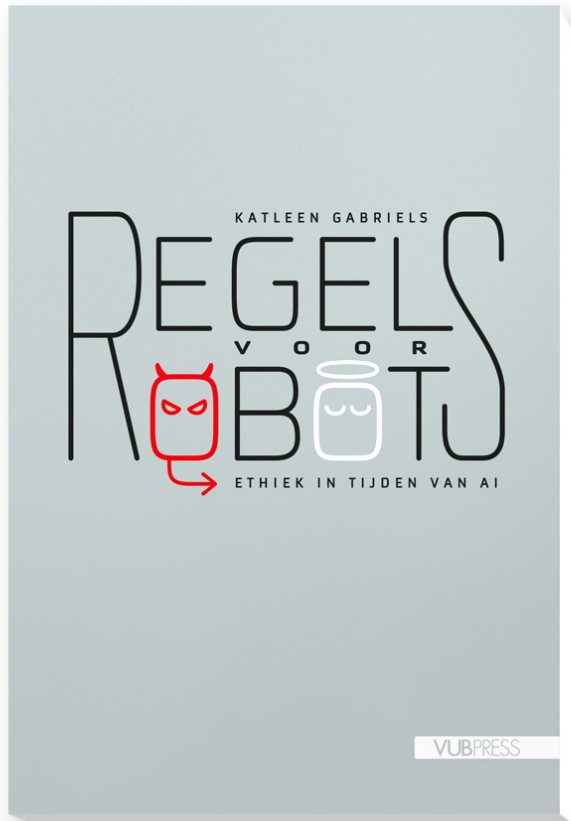


Hoe de digitale wereld  
je leven bepaalt



## INHOUD

7	Woord vooraf
11	<b>OFFLINE, ONLINE, ONLIFE</b> Waarom uitloggen een illusie is
33	<b>RIJDT DE SPOORWEG OVER ONS?</b> De morele, sociale, culturele en politieke dimensies van technologie
65	<b>DE TRANSPARANTE MENS</b> Van jager-verzamelaar tot dataverzamelaar
103	<b>ME, MY SELFIE AND I</b> De cultuur van het ik
137	<b>BIG MOTHER LAAT HAAR KROOST NOOIT MEER LOS</b> Over onzekere ouders, gemonitorde kinderen en bijdehante marketeers
157	<b>MENS SANA IN CORPORE DIMENSO</b> Behoedt een app per dag je voor tegenslag?
195	<b>WERK IS WAAR DE WIFI IS</b> Van koffiedame Arabelle tot robot Alice
221	<b>ONBEZORGD ONLIFE</b> Pleidooi voor mondigheid en ongehoorzaamheid
243	Eindnoten
279	Dankwoord



## INHOUD

Ethiek voor techniek. Waarom ook robots niet aan ethiek ontsnappen	9
Humane technologie. Hoe maken we de machine 'menselijker'?	19
Gekleurde technologie. Macht van het ontwerp	43
Transparantie. Het probleem van de vele handen	65
Superintelligentie. Maakbaar of utopie?	81
Morele machines. Tussen killer robot en moral agent	95
De robot en de filosoof. Tools voor de toekomst	119
Dankwoord	135
Leessuggesties	137
Eindnoten	141