



DE ZEVEN HOOFDZONDEN

Wie ben ik?

Prof. Dr. Lode Lauwaert

Hoger Instituut voor Wijsbegeerte

KU Leuven

Ethiek en techniekfilosofie

lode.lauwaert@kuleuven.be



De zeven hoofdzonden

1. Inleiding
2. Zes hoofdzonden
3. Morele verantwoordelijkheid
4. Conclusie



1. INLEIDING



1. Inleiding

Van stoom over elektriciteit tot het internet

“Data zijn de nieuwe olie”

Niet alle AI werkt met data, data niet enkel voor AI

De algoritmische samenleving

1. Inleiding

Targeted advertising: commercieel en politiek

I-Care en preventief onderhoud

Behandelen van halskanker

Gezichtsherkenning, Googs Maps, kredietwaardigheid, Spam, etc.

1. Inleiding

Enkele sterke overtuigingen:

1. AI is goed voor alles
2. *Superintelligence*
3. AI is disruptieve technologie

1. Inleiding

AI ontwricht, net als Uber, Airbnb, Tesla, 3D-printen

Niet oud, wel breed gedragen:

‘In elk geval is het overduidelijk dat de huidige digitale revolutie, die een industriële en een culturele revolutie combineert, onze hele leefwereld grondig wijzigt en alles op zijn kop zet. Veel mensen spreken dan ook over “disruptie”: de technologische revolutie ontwricht de oude wereld en doet die uit elkaar vallen.’
(Thierry Geerts, *Digitalis*, 2018)

1. Inleiding

Reden voor voor twijfel:

Tandenborstel van Oral B, software in tankstation

Redenen voor geloof:

Het wiel en drukpers

Is AI disruptief? Beschrijven en analyseren



2. ZES HOOFDZONDEN



2. Zes hoofdzonden

“Disruptieve technologie” (Cf. Clayton Christensen, 1995)

Economisch, sport, politiek, oorlog, ... en ethiek

Effect van technologie op domein + nieuw + openbreken

Veel interpretaties, veel problemen

“AI creëert totaal nieuwe morele problemen”

2. Zes hoofdzonden

1. Misbruik

Foute doelen: *deep fake* (Obama, DeepNude), *fake news*, terreur, etc.

... en het uitroenen van het menselijk ras

↔ niet nieuw

↔ bestaat niet, kunnen en zullen, ongegronde aanname

2. Zes hoofdzonden

2. Privacy

= controle over de eigen gegevens (welke gegevens en waarvoor?)

= geen *surveillance*

We worden voortdurend gevolgd - Clearview AI - Diabetes

↔ snuffelen in het dagboek van je kind, medische technologie

2. Zes hoofdzonden

3. Bias

Statistisch en moreel vooringenomen (en de band tussen beide)

Amazon (2018), OpenAI

↔ de negativiteitsbias, *lookism*

2. Zes hoofdzonden

4. Veiligheid

Security: beschermd tegen indringers

Safety: zonder lichamelijk schade

Zelfrijdende wagen: terrorisme, Elaine Herzberg (Uber)

↔ weinig of geen technologieën zijn risiceloos

2. Zes hoofdzonden

5. Transparantie

Voldoende informatie + inzicht (*Explainable AI*)

Niet alle AI is ondoorzichtig, niet alle niet-transparante AI is problematisch

Ondoorzichtigheid als moreel probleem: instrumenteel en op zich

↔ solliciteren, sporten, doodstraf

2. Zes hoofdzonden

6. Duurzaamheid

Milieu-ethiek: vervuiling, uitputting, aantasting

Zijn morele problemen morele problemen? Instrumenteel en finaal

Niet *in the cloud*: 1 AI-system = elektriciteitsverbruik van drie kerncentrales/uur

Broekasgassen: 2% van uitstoot, 200.000 kg koolstofdioxide (vs 500 kg per vlucht)

↔ het milieuprobleem is ouder dan AI

2. Zes hoofdzonden

Is AI disruptieve technologie?

= creëert AI totaal nieuwe problemen?

Misbruik, privacy, bias, veiligheid, ondoorzichtigheid en duurzaam

= ernstig en niet te verwaarlozen

Deze morele problemen zijn geen NIEUWE problemen!

Cf. Gradueel vs discontinu



3. MORELE VERANTWOORDELIJKHEID

3. Morele verantwoordelijkheid

Andreas Matthias: *responsibility gap* !

Niemand verantwoordelijk voor fouten door autonome AI

= zevende zonde

= problematisch en nieuw !

➔ Argument pro de disruptiethese

3. Morele verantwoordelijkheid

Autonome technologie

Autonoom 1: zelfstandig (negatief)

Autonoom 2: zelf beslissingen (positief)

Zelfrijdende wagens, Amazon Mechanical Turk, *killer robots*

3. Morele verantwoordelijkheid

1. Oorzakelijke verantwoordelijkheid: deel van keten
2. Morele verantwoordelijkheid: terechte kandidaat voor straf/beloning
3. Rolverantwoordelijkheid: plichten bij een rol

Wat is de relatie tussen die verschillende soorten?

3. Morele verantwoordelijkheid

Waarover gaat de *responsibility gap*?

1. Oorzakelijke verantwoordelijkheid? Neen.
2. Morele verantwoordelijkheid? Je kunt niemand straffen of belonen
3. Rolverantwoordelijkheid? Neen, integendeel. Cf. Hans Jonas (1979)

3. Morele verantwoordelijkheid

Voorwaarden voor morele verantwoordelijkheid:

1. Causale band
2. Autonomie
3. Kennis

3. Morele verantwoordelijkheid

Argument: controle

Autonome AI → verantwoordelijkheidskloof

Morele praktijk: controle is niet noodzakelijk

Gedachte-experiment: epilepsie

→ Morele verantwoordelijkheid zonder controle

3. Morele verantwoordelijkheid

Is er een wél een moreel verantwoordelijke?

Gebruiker, ingenieur, producent, etc.

Drie voorwaarden zijn vervuld!

➔ Er is geen *responsibility gap*



4. CONCLUSIE



4. Conclusie

“AI is disruptieve technologie”

Optimisten en pessimisten

Voorzichtigheid

Geen *zeven* zonden !