



Algocratie: de computers aan de macht?!



Meine Koeslag

Institute for Sustainable Process Technology



Mira Vegter

Wageningen University & Research

Zaalvoorzitter: Taco Kluwer, Agentschap Telecom

Symposium: Opstap naar weerbaarheid

SAAI – Sociale Acceptatie AI in Industrie



Institute for
Sustainable
Process Technology



This project is co-funded with subsidy from the
Topsector Energy by the Ministry of Economic
Affairs and Climate Policy.

Introductie SAAI – Sociale Acceptatie AI in Industrie

- Institute for Sustainable Process Technology
- Wageningen University & Research
- Radboud Universiteit Nijmegen
- Nouryon



Afbeelding met persoon, binnen, zitten, vrouw

Automatisch gegenereerde beschrijving



Institute for
Sustainable
Process Technology

Radboud University



Nouryon



WAGENINGEN
UNIVERSITY & RESEARCH

Vertrouwt u AI?

- Luistert u altijd naar uw navigatie?



Vertrouwt u AI?

- Luistert u altijd naar uw navigatie?
- Hoeveel veiliger/beter moet AI scoren t.o.v. de gemiddelde mens, voordat u een advies van de AI accepteert?



Vertrouwt u AI?

- Luistert u altijd naar uw navigatie?
- Hoeveel veiliger/beter moet AI scoren t.o.v. de gemiddelde mens, voordat u een advies van de AI accepteert?
- Svenson (1981) Illusory superiority: For driving skills, 93% of the U.S. sample and 69% of the Swedish sample put themselves in the top 50%; for safety, 88% of the U.S. and 77% of the Swedish put themselves in the top 50%.



Introductie

Acceptatie nieuwe technologie

- Nieuwe technologie stuit historisch op weerstand:
 - Industriële revolutie:
 - Angst voor baanverlies (Luddieten)
 - Sabbotage: Gooi je sabot (klomp) in de machine!
- Vandaag: AI staat uit in de fabriek



Introductie

Artificial Intelligence voor Industrie

Beloften

- Productie kwaliteit
- Hogere arbeidsproductiviteit
- Predictive maintenance
- Snellere R&D
- Nieuwe/persoonlijke producten
- Hogere grondstof efficiëntie
- Nieuwe business modellen/inzichten
- Automatisering gevaarlijk/onprettig/**saai** werk
- ...

Angsten & Weerstanden

- Baanverlies
- Run-away
- Gebrek aan transparantie
- Discriminatie
- Verlies autonomie
- Afhankelijkheid
- Herverdelingsvraagstuk
- Dat hebben we in 1990 al geprobeerd...
- China
- Skynet / Rise of the machines
- ...



Overwegingen

- Fairness
- Accountability
- Explainability
- Auditability



'Soft Impacts' van Technologie (Swierstra & te Molder 2012)

- “hard” impacts
 - Vaak in termen van risico's
 - objectief, rationeel, feiten, neutral
- “soft” impacts
 - Sociaal/ politiek / ethisch, vaak in termen van waarden
 - subjectief, emotioneel, partisan



'Soft Impacts' van Technologie (Swiertsra & te Molder 2012)

- “hard” impacts
 - Vaak in termen van risico's(objectief, rationeel, feite, neutral),
- “soft” impacts
 - Sociaal / ethisch, vaak in termen van waarden (subjectief, emotioneel, partisan,)

En dus niet
belangrijk?



'Soft Impacts' van Technologie (Swiertsra & te Molder 2012)

- “hard” impacts
 - Vaak in termen van risico's (objectief, rationeel, feiten, neutral),
- “soft” impacts
 - Sociaal / ethisch, vaak in termen van waarden (subjectief, emotioneel, partisan,)

En dus niet
belangrijk?

Klopt dit
onderscheid wel?



Technologie verbindt ons met de wereld - het 'medieert'

Denk aan jij aan het verleden in zwart / wit ? (tv)

Ben jij altijd bereikbaar? (app-en)



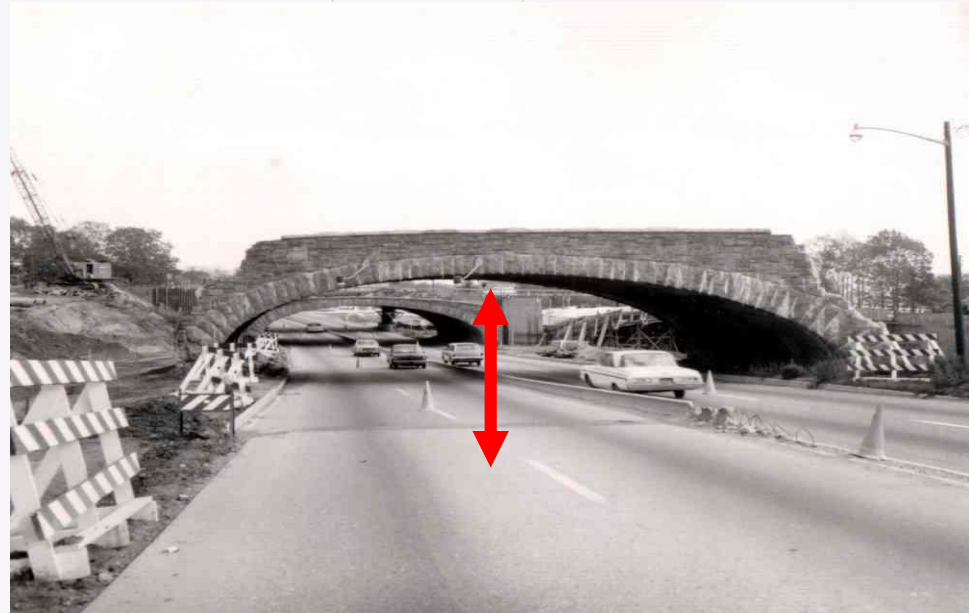
Techno-moral change (Swierstra 2017)



Technologie draagt een script



Technologie draagt een script



Technologie medieert ons handelen en onze waarneming



Project SAAI : hoe kan AI het denken en handelen positief beïnvloeden binnen de proces industrie

negatieve consequenties anticiperen en voorkomen?

- Team Learning Intervention
- Value-sensitive design



Case discussion:

Neem de rol van Project Manager van een fabriek

Industrieel project: ROBOT AUTODIDACT heeft als doel om met deep reinforcement learning industriële robots in staat te stellen zichzelf te trainen. Doordat robots dezelfde taak veelvuldig uit te laten voeren, en bij elke stap te laten leren, worden ze vanzelf voldoende accuraat.

Het doel is dat meerdere robots gezamenlijk leren, zodat wat 1 robot in 8 uur kan leren, kunnen 8 robots in 1 uur. Door snelle leerprocessen hoeven fabrieken minder vaak/lang de productie stil te leggen en kan 1 fabriek meer variatie aan producten bewerken.

Het project ziet het nieuwe onderzoek als een stap naar de toekomst waarin robots meer en meer onderlinge kennisoverdracht kunnen hebben. Robot applicaties met veel repetetief werk is het laag hangende fruit voor deze soort van transfer learning.



Case discussion

- Groups of 5, sort per: **AI Aristocrat, AI Knight, AI Dreamer or AI Doubter**

Angsten & Weerstanden

- Baanverlies
- Run-away
- Gebrek aan transparantie
- Discriminatie
- Verlies autonomie
- Afhankelijkheid
- Herverdelingsvraagstuk
- Dat hebben we in 1990 al geprobeerd...
- China
- Skynet / Rise of the machines
- ...



Case discussion

- Groups of 5, sort per: **AI Aristocrat, AI Knight, AI Dreamer or AI Doubter**
- Zet je dit project op?

Angsten & Weerstanden

- Baanverlies
- Run-away
- Gebrek aan transparantie
- Discriminatie
- Verlies autonomie
- Afhankelijkheid
- Herverdelingsvraagstuk
- Dat hebben we in 1990 al geprobeerd...
- China
- Skynet / Rise of the machines
- ...



Case discussion

- Groups of 5, sort per: **AI Aristocrat, AI Knight, AI Dreamer or AI Doubter**
- Zet je dit project op?
 - Waarom wel/niet?

Angsten & Weerstanden

- Baanverlies
- Run-away
- Gebrek aan transparantie
- Discriminatie
- Verlies autonomie
- Afhankelijkheid
- Herverdelingsvraagstuk
- Dat hebben we in 1990 al geprobeerd...
- China
- Skynet / Rise of the machines
- ...



Case discussion

- Groups of 5, sort per: **AI Aristocrat, AI Knight, AI Dreamer or AI Doubter**
- Zet je dit project op?
 - Waarom wel/niet?
- Welke Angsten & Weerstanden verwacht je?

Angsten & Weerstanden

- Baanverlies
- Run-away
- Gebrek aan transparantie
- Discriminatie
- Verlies autonomie
- Afhankelijkheid
- Herverdelingsvraagstuk
- Dat hebben we in 1990 al geprobeerd...
- China
- Skynet / Rise of the machines
- ...



Discussion

- Is uw beeld over AI verandert in de laatste 45 min?

