Hybrid Intelligence

AI systems that collaborate with people instead of replacing them

Frank van Harmelen
Vrije Universiteit Amsterdam

Creative Commons License
CC BY 3.0:
Allowed to copy, redistribute
remix & transform
But must attribute

Why HI?

The classical automation perspective on AI

"My guess for when we will have **full autonomy** [in cars] is approximately three years" (Elon Musk, **2015**)

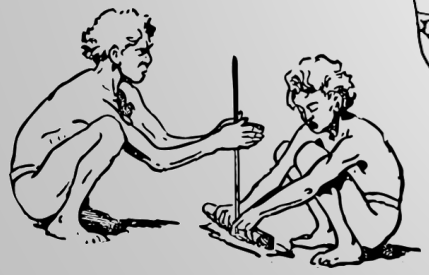
"a highly-trained and specialised radiologist may now be in greater danger of **being replaced by a machine** than his own executive assistant" (Andrew Ng, The Economist, **2016**)

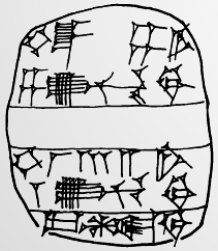
"People should stop training as radiologists now. It's just completely obvious that within 5 years, deep learning is **going to do better than radiologists**"
(Geoffrey Hinton, The New Yorker, **2017**)

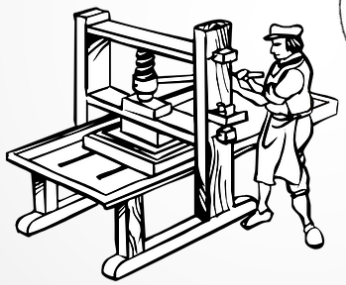
The problem with the automation perspective

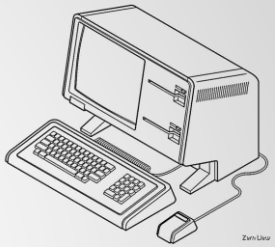
Replacing humans?...☹

An alternative perspective on AI: Augmenting humans









Collaboration in Art

The Hybrid Intelligence Center

20m€, 7 Uni's, 50 faculty, 10 years,
now going for 5 years

30 PhD/PDs in first cohort (2020-2023)
20 more in second cohort (2023-2026)
25 more about to be hired (2025-2028)

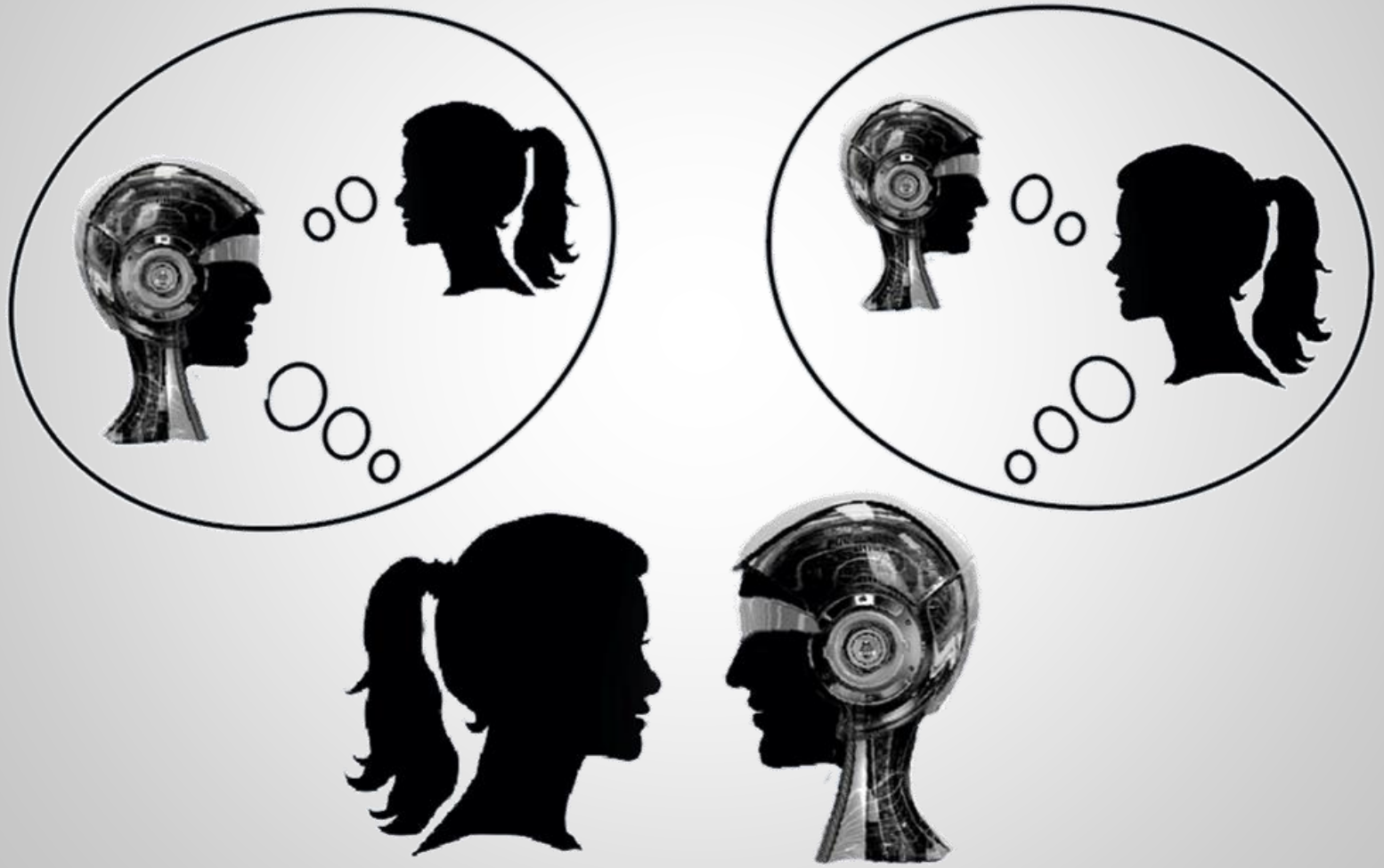
https://www.hybrid-intelligence-centre.nl/

UNIVERSITEIT TWENTE.

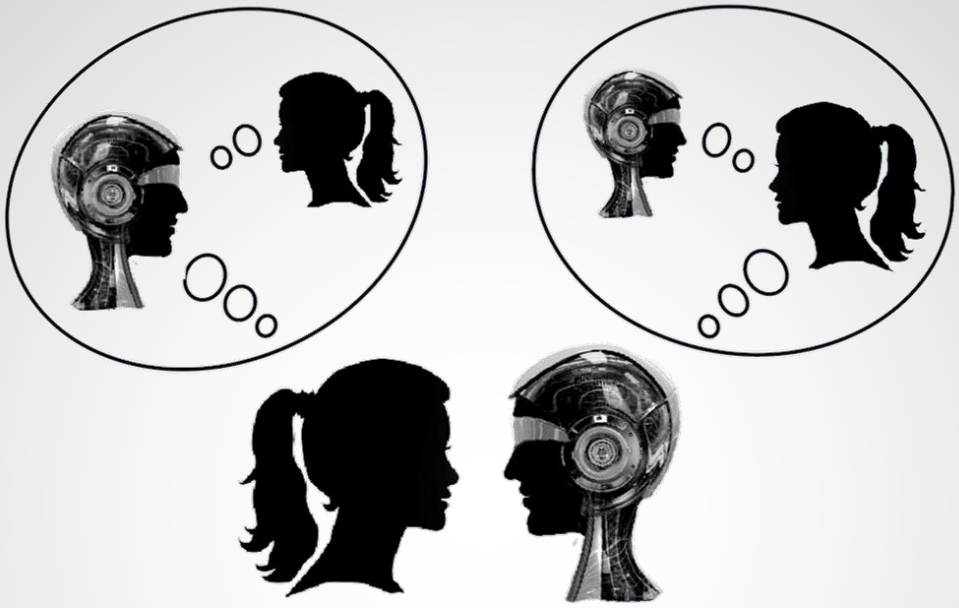
VU VRIJE UNIVERSITEIT AMSTERDAM TU Delft UvA university of groningen Universiteit Leiden Universiteit Utrecht

How?
(new research questions)

Theory of Mind



Theory of Mind

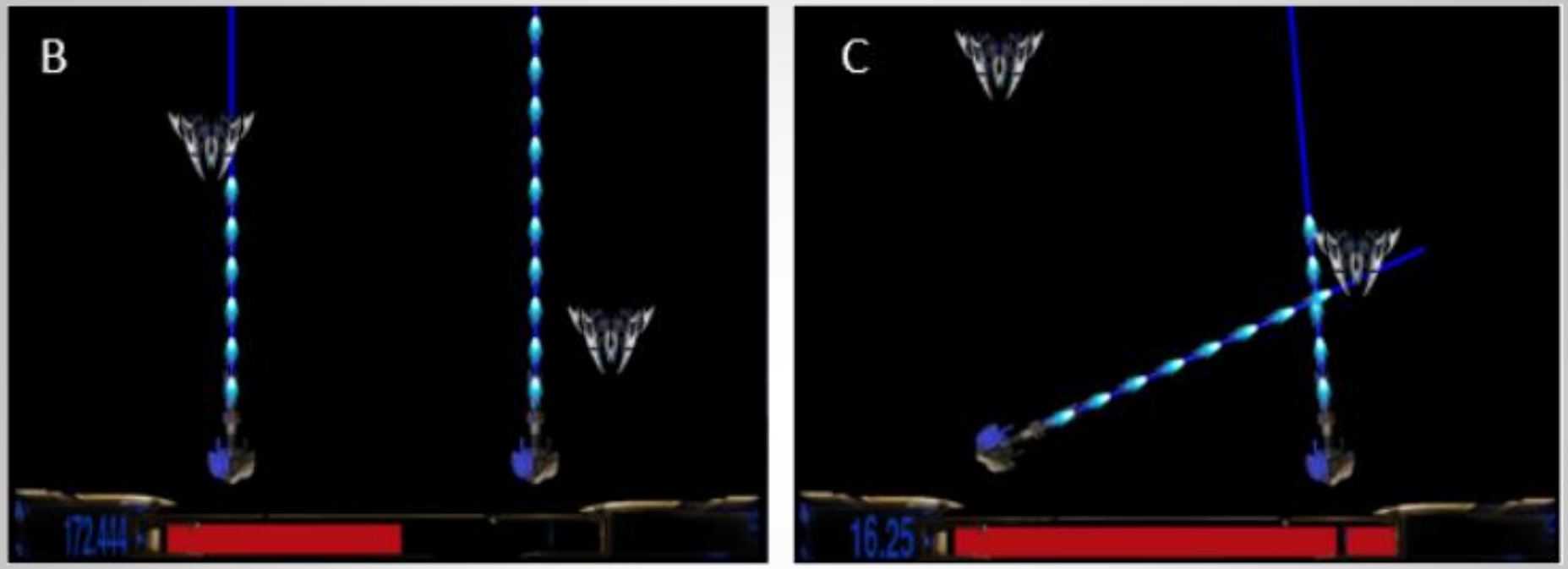


- 2nd order ToM is beneficial in competitive, cooperative, and mixed-motive situations
- software agents with deeper ToM levels give better support to humans on negotiation outcomes.

(de Weerd et al, AI Journal 2013)

Space Canons: **Learning to collaborate**

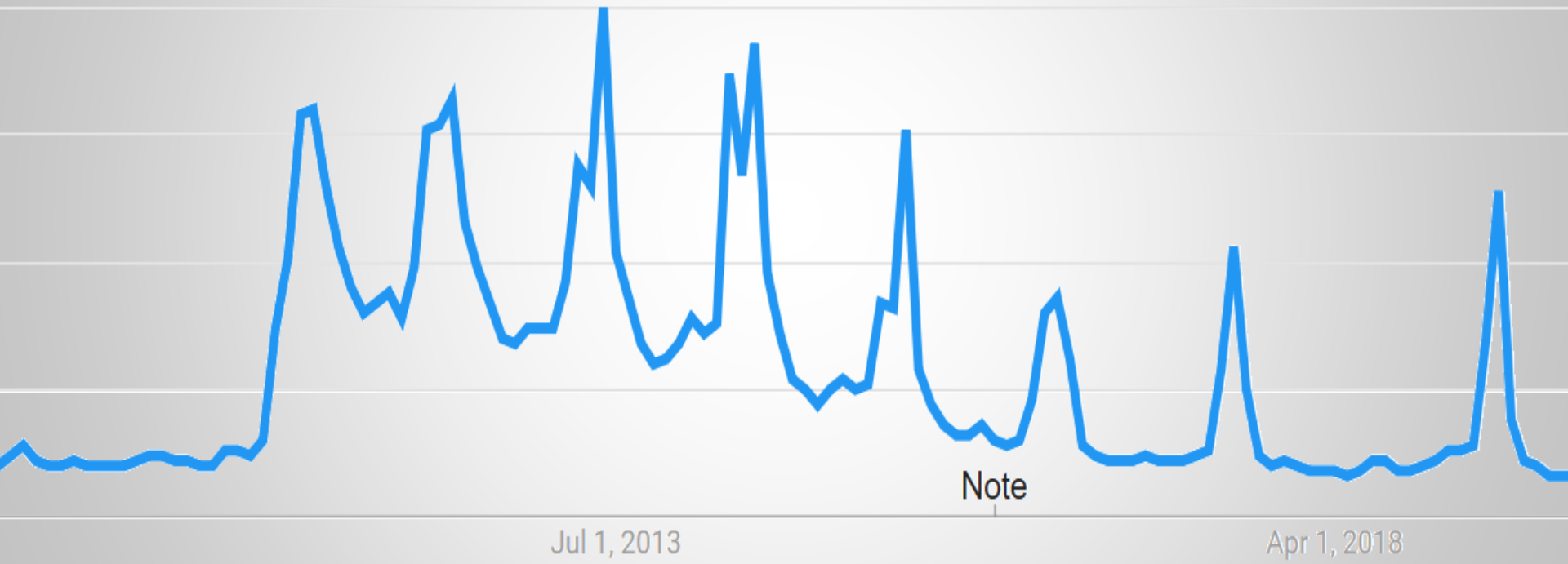
Verma & Acar, HHAI 2022



- Baseline: RL using the game score
- Instruction: RL with human reward for cooperation
- Imitation: Inverse RL from human coop. behaviour

instruction > imitation > baseline > 0

Justifications,
(not: explanation)



Google Trends for “Song of Ice and Fire”

Cooperation needs multi-modal interaction

Interaction beyond language:

Facial expression, Gesture, Posture

Long term human-agent collaboration: How to combat the novelty effect?

Question:

- Effect of personalisation on long term engagement

Experimental setup:

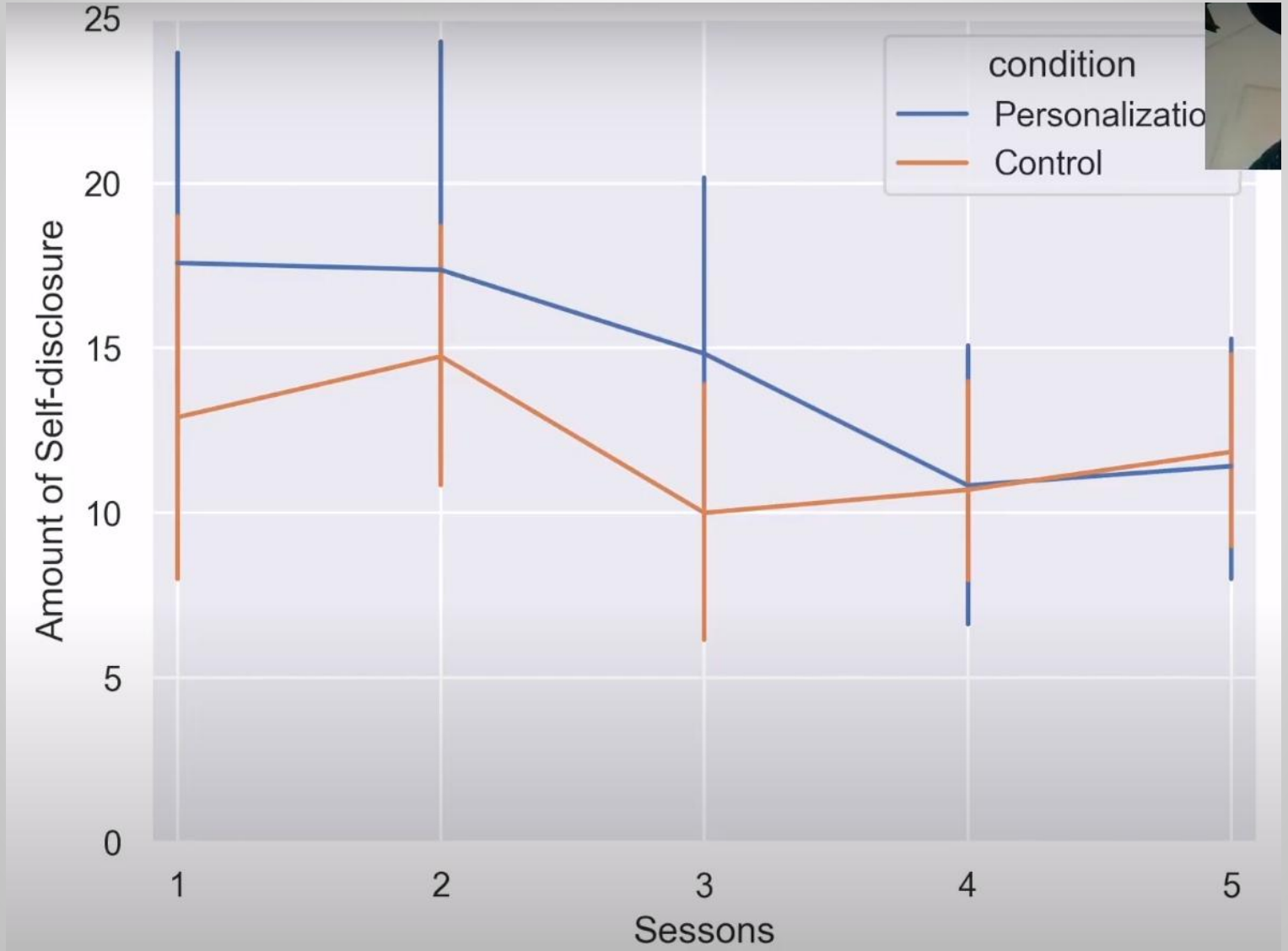
- 46 children (8-10y)
- 5x 15min conversation over 2 months

Metrics:

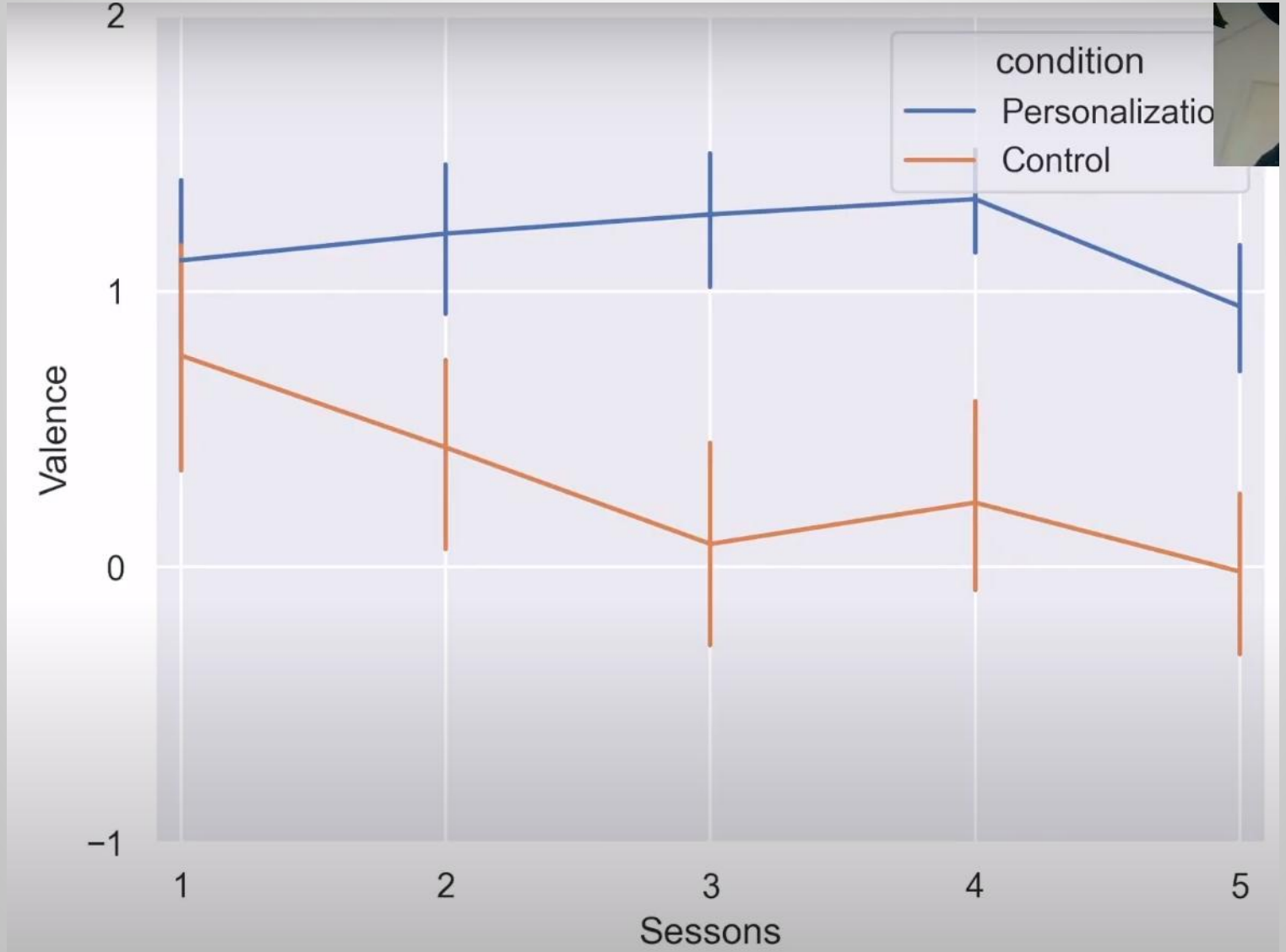
- Willingness to continue,
- self-disclosure,
- emotional valence

Ligthart et al, HRI 2022

Field experiments with children in cancer hospital



Field experiments with children in cancer hospital



HI applications (case studies the HI center)

Meet dr. Dalibor Vasilic, Erasmus Medical Center, Rotterdam

<u>Human:</u>

Microsurgeon

15 yrs of education

Cognitive and Physical

Limitations

Meet Synaptive's Modus V

Machine:

Robotic Microscope

Capabilities:

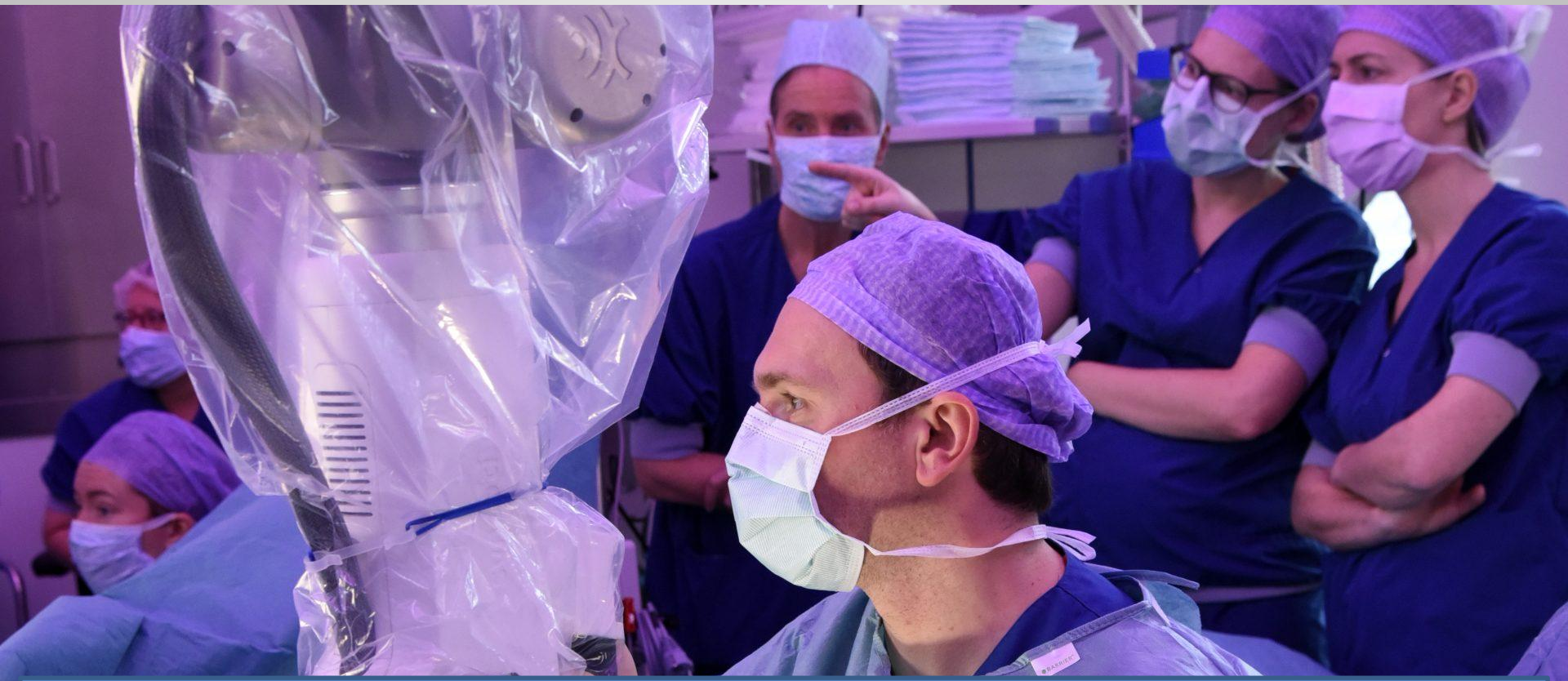
3D robotic arm,

voice recognition

instrument tracking

How Dalibor operates

When they bought the robot



Surgeons of the Depts of Plastic Surgery and Neurosurgery of Erasmus MC are so enthousiastic about the possibilities that, with support of the Executive Board, they have bought the first digital surgery microscope in the country.
“‘The possibilities are endless”, says plastic surgeon dr. Dalibor Vasilic, Erasmus MC, 2020

When they used the robot

It doesn’t adapt to me, I have to adapt to it
It doesn’t know what my plan is,
what I’m going to do next,
what it should do next.
Can HI help us to build a surgical platform?

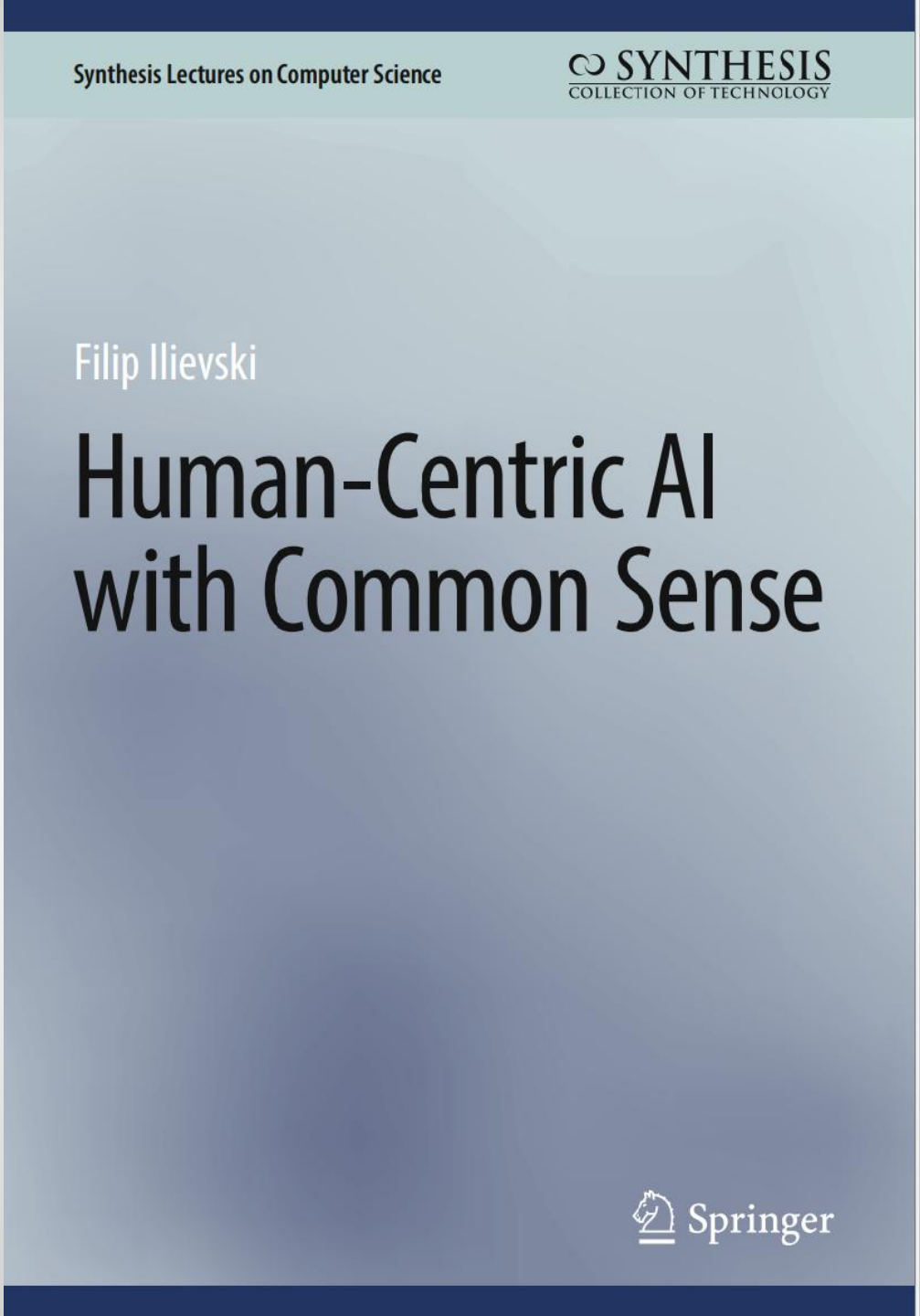
dr. Dalibor Vasilic, conversation, 2022

HI in the classroom

HI in pediatric oncology

HI in the children hospital

Just out:


Synthesis Lectures on Computer Science
SYNTHESIS
COLLECTION OF TECHNOLOGY
Filip Ilievski
Human-Centric AI with Common Sense
Springer