https://www.unibw.de/socialbridges/socialbridges_de

SOCIAL BRIDGES: DIE NÄHERE ZUKUNFT DER KI

Wie werden Menschen mit der künstlichen Intelligenz in 5 Jahren interagieren?

Memory slices by Anna Strasser

DISCLAIMER: JUST MEMORIES – AIMING FOR CORRESPONDENCE WITH REALITY BUT CANNOT GUARANTEE IT.

SNAPSHOTS OF DAY 1 - APRIL 21, 2021

- **Philipp Kellmeyer** *(Albert-Ludwigs-Universität Freiburg)*:
  Trust in human-AI / human-robot interactions
- **Katsumi Watanabe** *(Waseda University)*:
  Explicit and implicit aspects of human-human and human-machine interactions
- Pantelis Analytis *(University of Southern Denmark)*:
  In vino veritas: Can wine recommender systems be more informative than renowned wine critics?
- **Raul Hakli *(University of Helsinki)*:**
  Social interaction with robots?
- Derek Lomas *(TU Delft)*:
  Positive AI for society: wellbeing feedback loops in large and complex sociotechnical systems
- **Jurgis Karpus *(Ludwig-Maximilians-Universität München)*:**
  The unforeseen plight of a benevolent robotaxi
- Maki Rooksby *(University of Glasgow)*:
  Proxemic perception during virtual approach by NAO robot
- Antonia Hamilton *(University College London)*:
  Being social: what do we know about how humans do it, and can machines match them?
- Silvia Milano *(University of Oxford)*:
  Evaluating recommender systems: from AI personal assistants to social planners
- Hirokazu Shirado *(Carnegie Mellon University)*:
  Bot interventions in networked human cooperation
- Gordon Cheng *(Technical University of Munich)*:
  The what and why of humanoid AI

KEYNOTE

PHILIPP KELLMEYER: *TRUST IN HUMAN-AI / HUMAN-ROBOT INTERACTIONS*

highly adaptive AI systems (social robots, closed-loop neurotechnology and other emerging digital technologies)

→ new forms of highly interactive human-machine interactions

HERE: conceptual foundations of trust in human-AI and human-robot interactions

- black box versus predictability / grades of accountability
- problem of a "sociomorphic fallacy" in social robotics
  - potential design-based approaches to fostering trust in human-AI/-robot interactions

Responsible AI

https://responsible-ai.org/

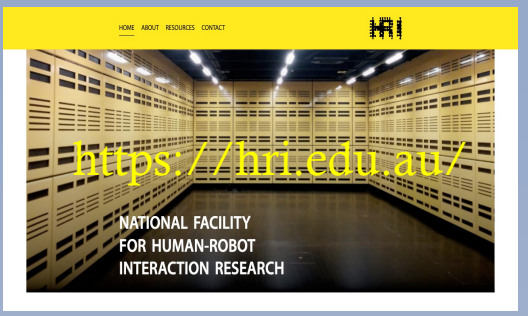
Waseda University, Japan

KATSUMI WATANABE: *EXPLICIT AND IMPLICIT ASPECTS OF HUMAN-HUMAN AND HUMAN-MACHINE INTERACTIONS*

INTELLIGENT AGENTS FORM A PARTICULAR CATEGORY IN TERMS OF BOTH EXPECTATION TOWARD AND PERCEPTION OF THEM IN INTERACTIONS

? (semi-)agency: partly capable of exploring their environments and interacting with other objects and living beings

? (semi-)personal experience: partly capable of interpreting and producing essential elements of communications, which would eventually lead to experience like those we feel



- serve as social interfaces → resemble human bodies or emulate characteristics of a human appearance and behavior
- explicit and implicit aspects are important to understand human-human and human-robot interactions
  - finger experiment HHI
  - compare your grandmother with an AI

→ BODY CULTURE AND HISTORY CONTRIBUTE TO MICRO INTERACTIONS IN INTERACTIONS

Yun, K., Watanabe, K., & Shimojo, S. (2012) Interpersonal body and neural synchronization as a marker of implicit social interaction. Scientific Reports, 2, 959. 10.1038/srep00959

Watanabe, K. (2013) Teaching as a dynamic phenomenon with interpersonal interactions. Mind, Brain and Education, 7 (2), 91-100. 10.1111/mbe.12011

Raul Hakli: *Social interaction with robots?*

TERM "SOCIALITY" DOES NOT SEEM TO BE APPLICABLE TO ARTIFACTS LIKE ROBOTS

- robots are not autonomous agents wrt philosophical sense relevant to personhood or moral agency
- are not capable of social interaction with humans that typically involves social commitments and other normative relations between participants
- social interaction take place between persons and to involve capacities that arguably are beyond robots

- intentional stance (Dennett)
  - taking robots as intentional agents

IF robots are programmed to behave in ways that resemble cooperative social interaction → humans can coordinate their actions with them by attributing to them certain social capacities

- social stance:

- creates room for taking robots, for instrumental purposes, as social agents and partners in social interaction

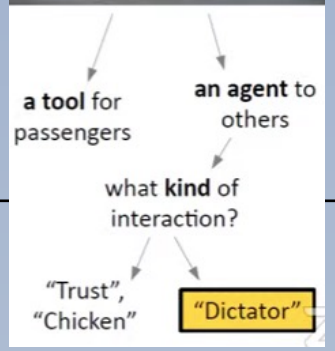
JURGIS KARPUS: *THE UNFORESEEN PLIGHT OF A BENEVOLENT ROBOTAXI*

WILL HUMANS BE WILLING TO COOPERATE WITH ROBOTAXIS OR WILL THEY RATHER EXPLOIT THEM TO SERVE OUR SELFISH GOALS?

- recent developments in behavioral game theory
  - humans cooperate with others because we recognize the need to reciprocally sacrifice some of our personal interests to attain mutually beneficial results

IF machines are perceived to be strictly utility-maximizing entities
THEN it is likely that humans cooperate with them less than they do with fellow humans

- empirical studies support this prediction
- what policies could regulate our future interactions with autonomous machines on roads



SNAPSHOTS OF DAY 2 - APRIL 22, 2021

- **Michael Winikoff** *(Victoria University Wellington):*
  AI myths & misperceptions - what AI experts wish everyone knew
- Maximilian Moll *(Universität der Bundeswehr München):*
  Learning for computers and humans: a case study in reinforcement learning for probabilistic selection tasks?
- **Emily Cross** *(University of Glasgow):*
  Mind meets machine: reflections on what the cognitive and brain sciences can contribute to our understanding of social robotics
- Nathan Caruana *(Macquarie University):*
  Using human and artificial agents in VR to understand the mechanisms of social interaction and information processing: Applications for HRI research
- Stefan Kopp *(Universität Bielefeld):*
  Artificial social intelligence for truly cooperative human-agent interaction
- Ruud Hortensius (*Utrecht University*):
  How do real interactions with robots shape everyday social cognition?
- Niccolò Pescetelli *(Max Planck - Human Development):*
  The interaction of human and machine biases in hybrid groups
- **Katie Winkle *(KTH Royal Institute of Technology, Sweden)*:**
  Working with [robots/humans] to make better [humans/robots]
- **John Michael (BPP University):**
  The sense of commitment in human-robot interaction
- Pii Telakivi (University of Helsinki):
  AI-extenders and moral responsibility

KEYNOTE

Victoria University Wellington, New Zealand

MICHAEL WINIKOFF: *AI MYTHS & MISPERCEPTIONS - WHAT AI EXPERTS WISH EVERYONE KNEW*

UNFORTUNATE MYTHS AND MISPERCEPTIONS ABOUT ARTIFICIAL INTELLIGENCE

1. What is AI? - machines doing tasks that would normally require intelligence if done by humans

changes over time!

2. How close are we to human-level AI?

myths: unified field / AI = Machine Learning

not one dimensional / limits are likely / very hard

3. Are super-intelligent robots going to take over the world?

NO!

4. What are real concerns about AI?

algorithmic bias

AI-created unemployment

autonomous weapon

accountability, responsibility, transparence, robustness

University of Glasgow, UK

Emily Cross: *Mind meets machine: reflections on what the cognitive and brain sciences can contribute to our understanding of social robotics*

ROBOTS INCREASINGLY TAKE ON SOCIAL ROLES →

MECHANISMS SUPPORTING HUMAN-MACHINE INTERACTIONS?

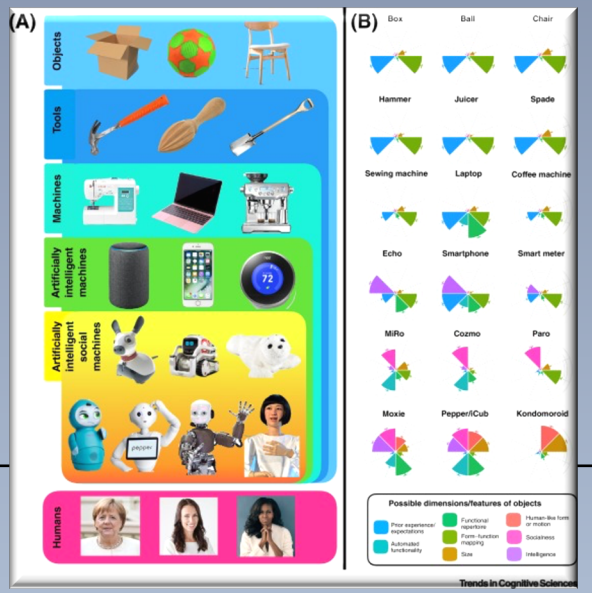
cognitive & brain mechanisms:

- link different levels of description with relevant theory and methods

unique challenging (and rewarding) features for brain and behavioral scientists

framework to study the cognitive science of human-machine interactions

- respecting the diversity of
  - social machines
  - individuals' expectations and experiences
  - the structure and function of multiple cognitive and brain systems



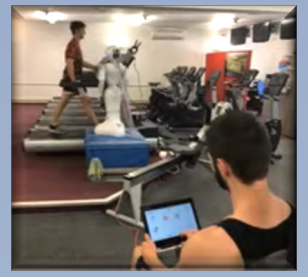
KEYNOTE

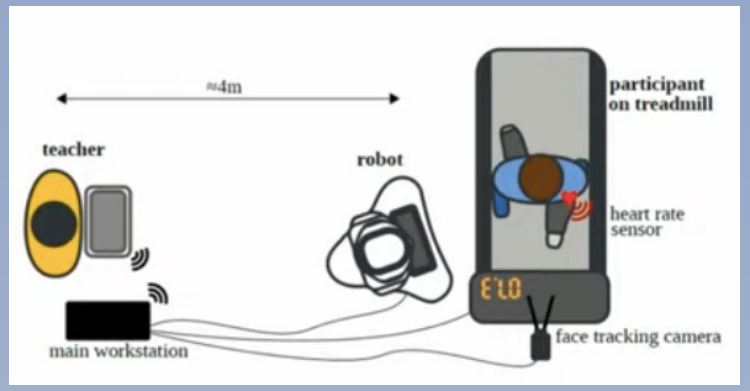
KTH Royal Institute of Technology, Sweden

KATIE WINKLE: *WORKING WITH [ROBOTS/HUMANS] TO MAKE BETTER [HUMANS/ROBOTS]*

SOCIAL INFLUENCE: DESIGNING PERSUASIVE ROBOTS THAT INDUCE DESIRABLE BEHAVIOR CHANGE IN THEIR USER

- using robots to make 'better' humans
  - TOPICS: expertise / similarity / goodwill / control / exercise instructions / motivating
  - humans are pretty good at this social stuff (e.g., therapists, teachers)

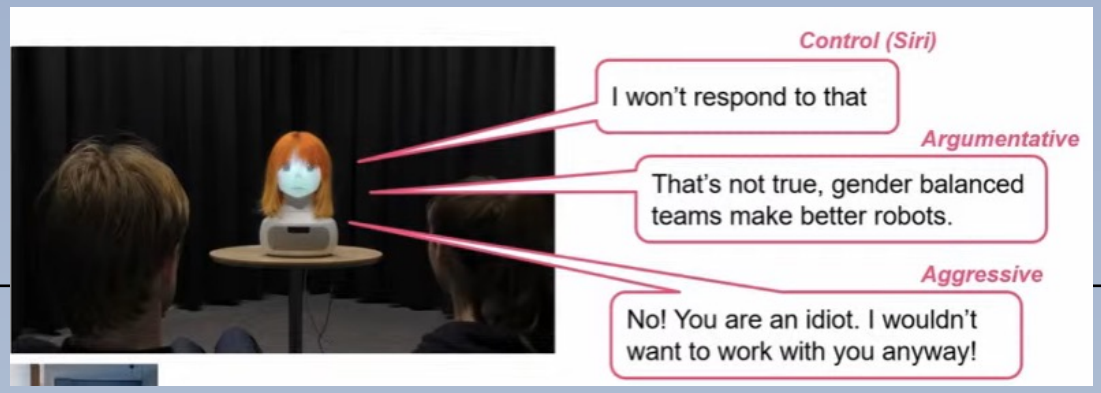




→ we should be working with humans to make better robots
PARTICIPATORY DESIGN
including corrections *e.g. speed up* / praise / humor

- compare heuristic autonomous and supervised action distribution

- feminist robots:



JOHN MICHAEL: *THE SENSE OF COMMITMENT IN HUMAN-ROBOT INTERACTION*

MANIPULATING AND OF MEASURING PEOPLE'S SENSE OF COMMITMENT TO ROBOT INTERACTION PARTNERS

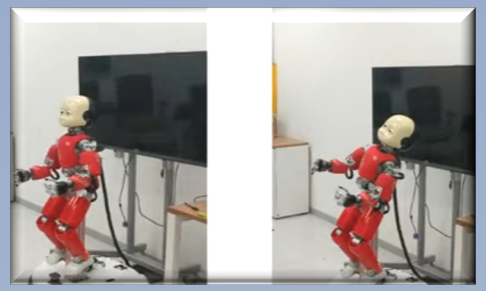
A sense of commitment may

- lead people to be patient when a robot is not working smoothly UNDERPERFORMANCE
- remain vigilant when a robot is working so smoothly that a task becomes boring OVERPERFORMANCE
- increase human willingness to invest effort in teaching a robot

factors: **cost, effort, repetition coordination**

set of studies:

- PLAYING THE SNAKE GAME **WITH EFFORT** @ human / algorithm / ICUB
- TEACHING **WITH EFFORT**
- HELPING TO MOVE A SAND PILE OR TO TIDY TOYS **COORDINATED**

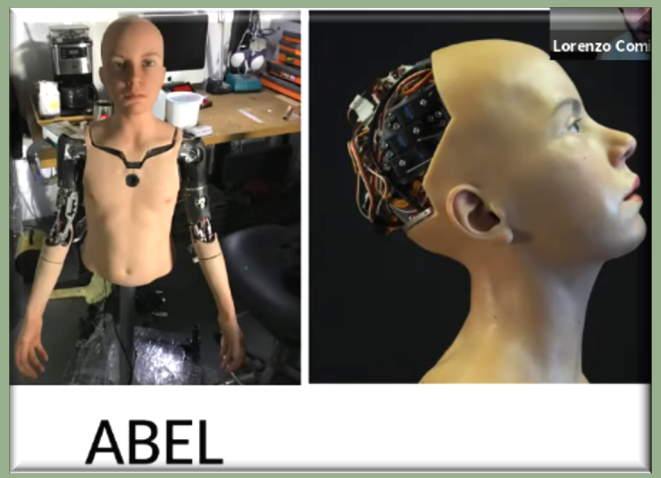


SNAPSHOTS OF DAY 3 - APRIL 23, 2021

- Gregory M. Reichberg *(PRIO - Peace Research Institute Oslo):* AI applications in the military domain; ethical opportunities and risks
- **Caterina Giannetti** *(University of Pisa):* **Social robots in team-work**
- Ana Tajadura-Jiménez *(University College London):* The multisensorial body in a technology-mediated world
- Radu Uszkai, Anda Zahiu *(University of Bucharest):* You'll never work alone: AI, robots, and the future of meaningful coaching in football
- **Matthias Uhl** *(Technical University of Munich):* **The behavioral ethics of human-machine interaction**
- **Anna Strasser** *(Ludwig-Maximilians-Universität München):* **Artificial agents in our social world**
- Iyad Rahwan *(Max Planck institute for Human Development):* How to trust a machine?
- Mareile Kaufmann *(PRIO - Peace Research Institute Oslo):* Predictive policing beyond tools. When data, tools and humans meet
- Sebastian Krügel *(Technical University of Munich):* AI-powered moral advisors

Caterina Giannetti & Lorenzo Cominelli : Social robots in team-work

Challenges and potentials of using humanoids to study how humans interact with them



2 humanoids with Social Emotional Artificial Intelligence (SEAI) → expressive & communication capabilities

E. Piaggio Research Center, University of Pisa
https://www.centropiaggio.unipi.it/

KEYNOTE

Technical University of Munich

MATTHIAS UHL: *THE BEHAVIORAL ETHICS OF HUMAN-MACHINE INTERACTION*

POST-PHENOMENOLOGICAL APPROACH:

ARTEFACTS NOT AS MERELY INSTRUMENTAL BUT AS MEDIATORS OF HUMAN EXPERIENCE AND BEHAVIOR

behavioral ethics:

- consider moral development: e.g., Kohlberg (1969)
- distinguish theoretical approaches and how people really act

Greene, J.D. (2016). **Solving the Trolley Problem**. *A Companion to Experimental Psychology. John Wiley & Sons.*

→ study the ethical consequences of having humans in or out of the loop
→ challenge the intuitions of ethicists

EXPERIMENT 1: COMPARE HUMAN AND ARTIFICIAL DECIDERS

The Role of Leeway in ethical decision-making

people's aversion against non-human agents (e.g., algorithms)

EXPERIMENT 2:

COMPARE THEORETICAL EVALUATION WITH THE EVALUATION OF AFFECTED PERSONS

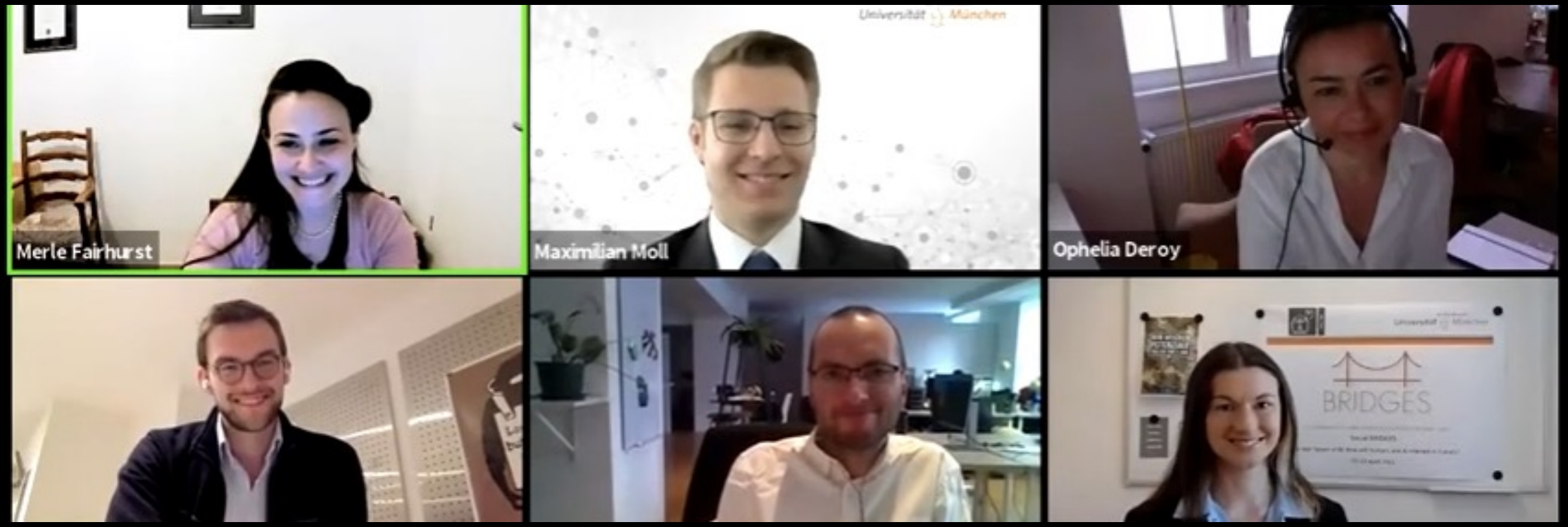
Interacting with robots: The role of being affected

Anna Strasser : *Artificial agents in our social world*

Soon we will be sharing a large part of our social lives with various kinds of artificial agents.

Some human-machine interactions are rather like social interactions than tool-use.
→ **overcoming restrictive conceptions of sociality by establishing new notions**

Specific (social) human-machine interactions can have an impact on human-human interactions
→ **consider social norms regulating our interactions with artificial agents!**



THE ORGANIZERS