



SOCIAL ROBOTS IN SOCIAL INSTITUTIONS

August 16-19 | University of Helsinki

Memory slices by Anna Strasser
**DISCLAIMER: JUST MEMORIES – AIMING FOR CORRESPONDENCE
WITH REALITY BUT CANNOT GUARANTEE IT.**

Sven Nyholm

A New Control Problem? Humanoid Robots, AI, and the Value of Control

TESLA BOT | ERICA | SELF-DRIVING CARS | KILLER ROBOTS

direct or indirect | more or less robust | multi-dimensional control over AI

Who is controlling whom?

FORMS OF CONTROL	POSITIVE	NEGATIVE
instrumental	control over means to important ends	cases in which control is somehow counter-productive
in itself (non-instrumentally)	self-control (Kant – inner worth, Stoic, Rachels – personal autonomy Waldron – human dignity) virtuosity / mastery control over environment (Nussbaum)	control over persons (slavery)

Can control over AI as a form of self-control?

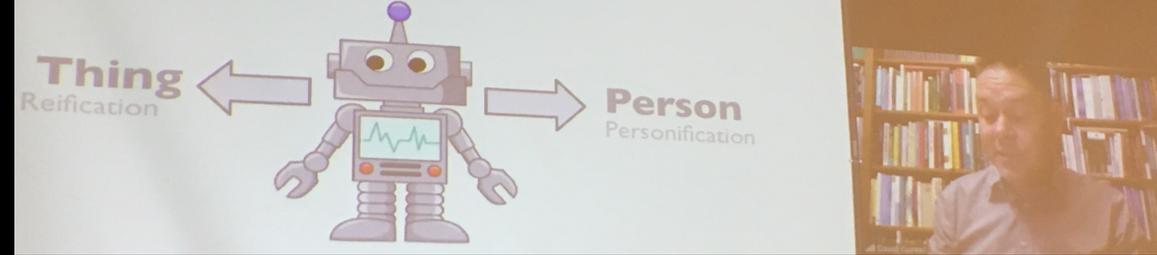
- extended minds
- AI systems as extensions of our (group) agency



What negative aspects become relevant if control over AI is similar to control over agents that at least represent or symbolize moral persons?



Persons, Things, or Otherwise

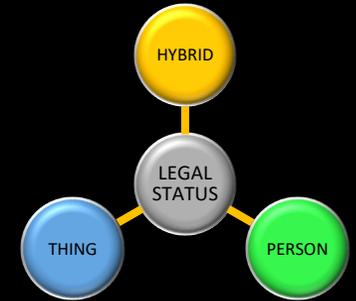


ORGANIZER: DAVID GUNKEL: THE PLACE OF SOCIAL ROBOT IN SOCIAL INSTITUTIONS

- *distinction between persons & things becomes flexible*

Diana Mădălina Mocanu: *Humans With, Not Versus Robots*

- *Can we attribute liability to the hybrid?*
economic benefit | control capacities ... examples: Rosa (robotic arms) / Tommy / Nao



Jesse Pagter: *Should We Speculate About Robots?*

- *internal: one should be self-critical | external: one needs to demonstrate why and how speculative thinking can be fruitful*

Dane Leigh Gogoshin: *Challenging the Premises of the Responsibility Gap*

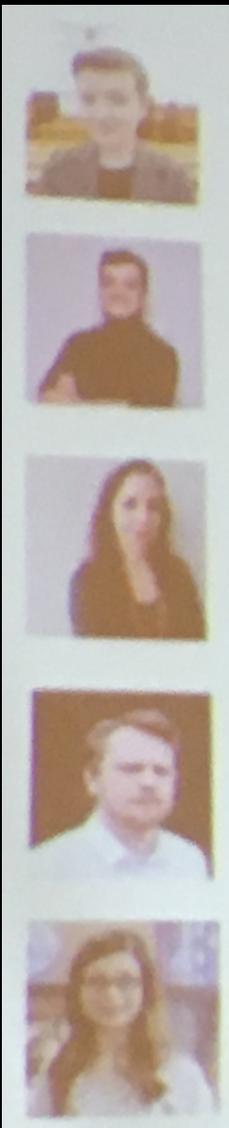
- *Why we can not hide behind the team?*
revisit questions regarding accountability OR substantive responsibility

Maciej Musiał: *If Robots Were Persons, What Kind of Persons Could and Should They Be?*

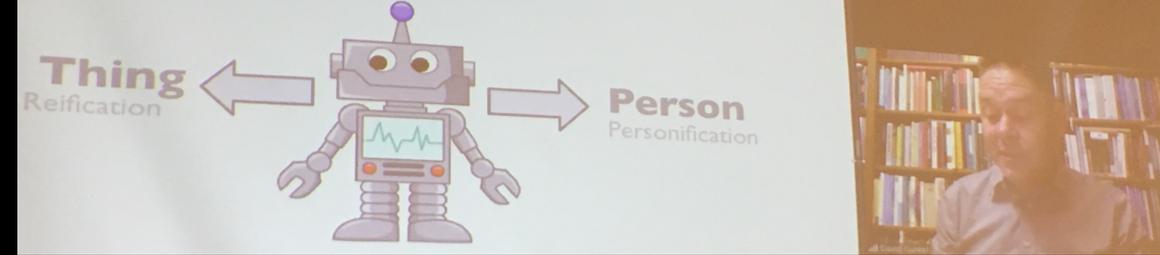
- *How should we design and develop robot persons that are moral agents to care for their well-being?*

Aybike Tunç: *Legal Personhood for Artificial Intelligence: Can it and should it be Conferred?*

- *Everything can be a legal person, but should artificial person really get legal personhood?*



Workshop 1: Persons, Things, or Otherwise



- Kamil Mama: *How Should the Law Treat Attacks on Police Robots?*
 - *crimes against the functioning of the state and local government institutions (Poland) authority*
- Henrik Skaug Sætra: *Requirements for the Inclusion of Robots in Social Institutions*
 - *possible*
- Anne Gerdes: *It's Time to Make a Luddite Turn – We're Confronted with Neo-Tayloristic Vampire Robots*
 - *being afraid of the way we use it*
- David Gunkel: *Robots Should Not be Slaves*
 - *We should not normalize any institution of slavery!*



Virginia Dignum:

Responsible AI: From Principles to Action



WHAT IS AI? most pictures are blue | pictured as an entity | agency | there is a user | data

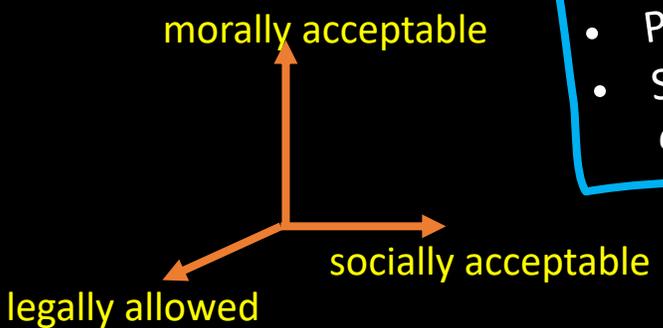
stereotypes
a rational agent
OPTIMIZATION : EFFICIENCY :
RATIONALITY : AGENCY : AUTONOMY
western / masculine

ADD social grounds for AI
FEMINIST THEORY

- acceptance & trust
- power structure (reinforcement / visualization)
- representation (bias / inclusion)
- other cultures

important concerns

- DATIFICATION (we are more than data)
- POWER (who is in charge, whose purposes)
- SUSTAINABILITY (costs: societal & environmental impact of AI)



ART

→ ASKING THE WHY QUESTIONS!

Piercosma Bisconti & Luca Possati:

Sociomorphing and an Actor-Network Approach to Social Robotics



implicit assumption:

the more anthropomorphism we can reach in HRI, the more effective the robot will be in "being social."

alternative

→ a non-anthropocentric approach to HRI is the one of "sociomorphing" Aarhus University

ANT: semiotics of materiality that is symmetrical with respect to human and non-human agents

possibilities offered by

ACTOR NETWORK THEORY (ANT)

- a valid framework to re-think the conceptual couple anthropomorphizing / sociomorphing → reinterpreting the dualism of the two via ANT

QUESTION:

- What kind of negotiation process and social practices can be developed in HRI, given the notion of sociomorph interactional networks?

A research agenda: sociality as a property of a system

I) Sociality is a property of a system. It does not pertain to single actors but to their system of relationships.

II) Therefore, we are not interested on the sociability of the robot *in se*, but on how much the robot can increase the sociality of the system.

III) What kind of behaviours increase the sociality of the system can be studied through ANT, with the concepts of spokesperson and negotiation.

IV) This might lead to a theoretical reformulation in interaction studies and in the evaluation criteria for social robotics.

Robin Zebrowski:

Dual Aspect Presence: Intercorporeality for Thee but Not for Me



dual aspect presence (= feeling of being present with other beings)

lack of a full experience of presence with others because of zoom: no eye contact because of camera-placement, no phenomenological facts about experiences of co-presence.

feel a presence of a being even though it's only a robot interacting with some social robots, I may feel as if I am in the presence of a real being

foundationally related & jointly are oddly paradoxical

explored approaches:

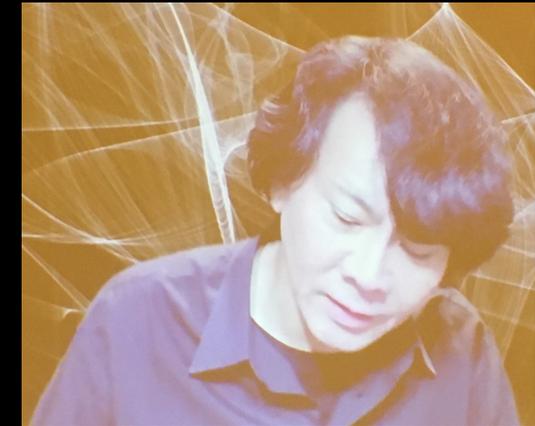
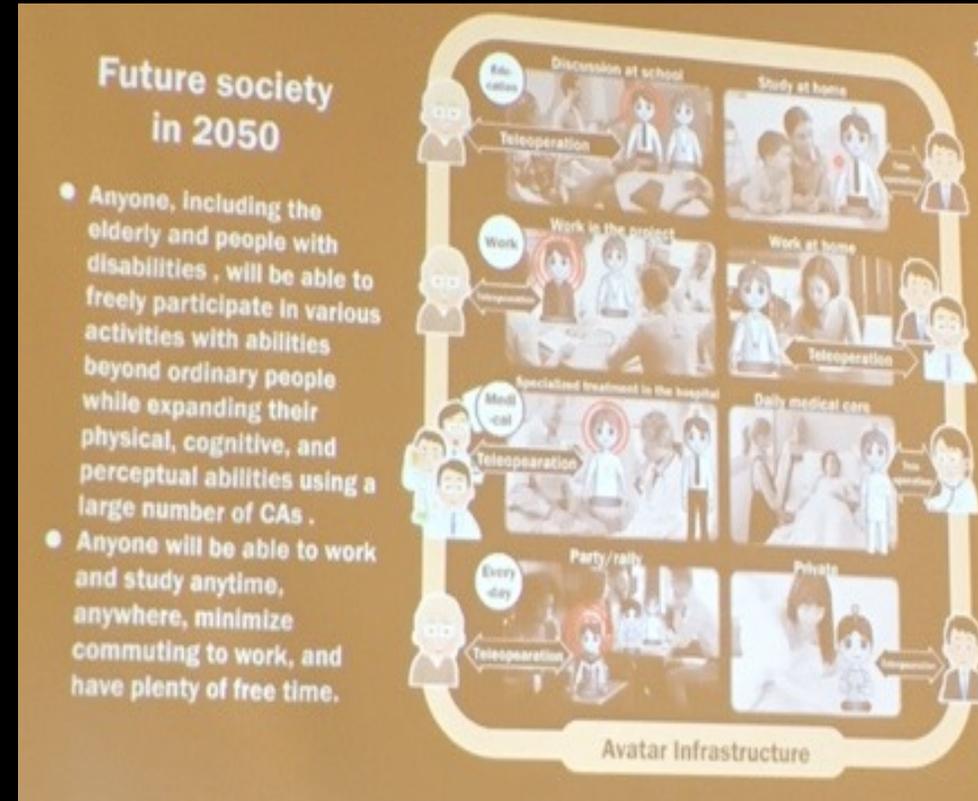
NEUROPHENOMENOLOGY, INTERCORPOREALITY, AND SOCIOMORPHING

Workshop 3: ELSI of the Avatar Symbiotic Society

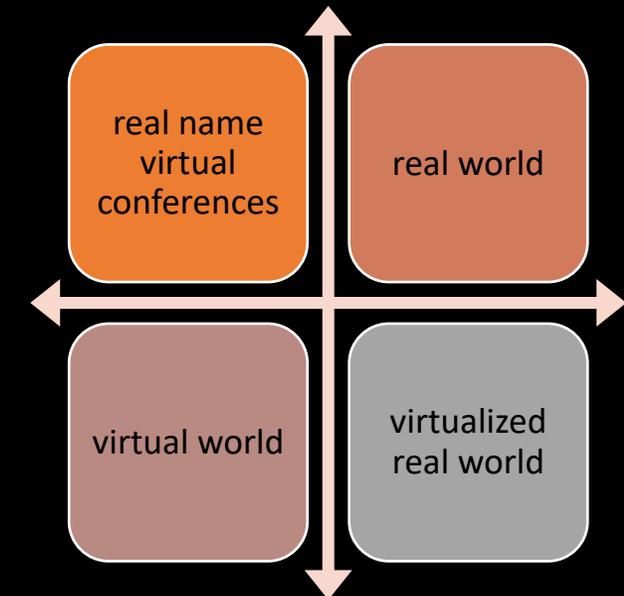
Hiroshi Ishiguro: Realisation of the Avatar Symbiotic Society: The Concept and Technologies

VIRTUALIZE THE REAL WORLD

- realize a society in which people are free from the constraints of body, brain, space, and time
- realize a robot that learns and acts by itself and coexist with people through co-evolution of AI and robots



- how to give robots a sense of presence
- how to make them look and feel alive
- how to enrich human-robot interaction
- how to design a society where humans & robots coexist



Workshop 3: ELSI of the Avatar Symbiotic Society

Takayuki Kanda and Minao Kukita: Moral Computing for Avatars and Its Ethical Challenges

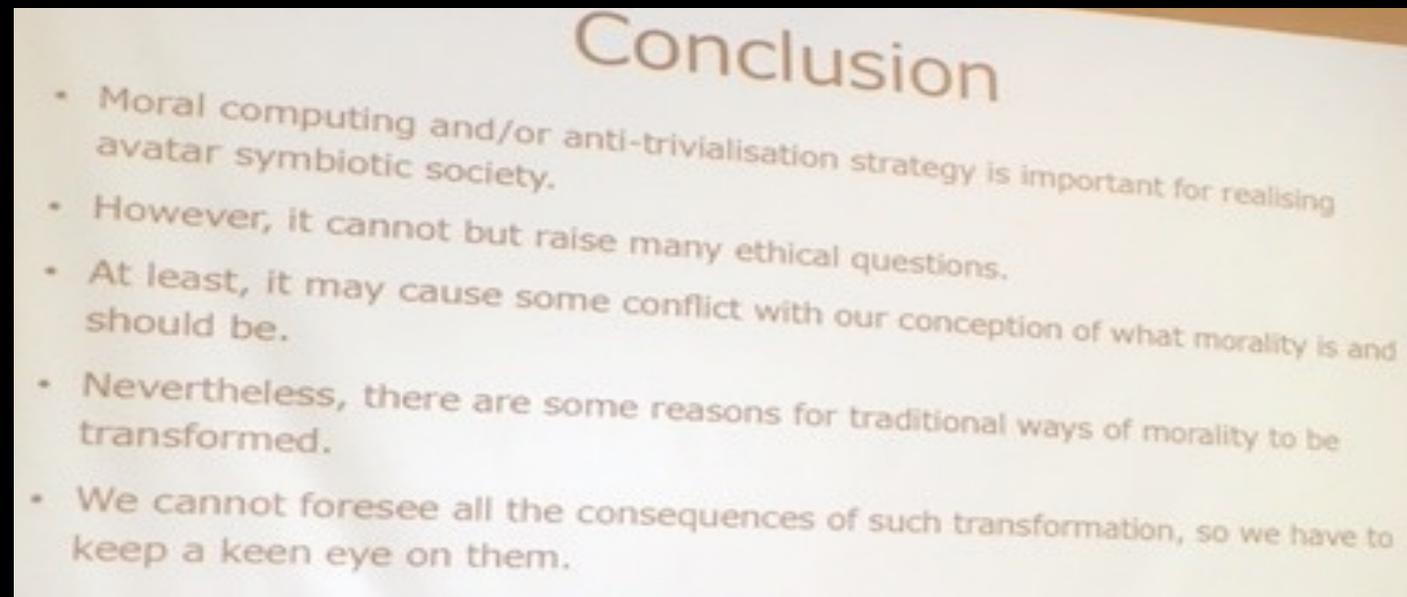
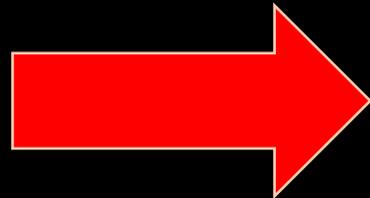


KEY ISSUES FOR THE REALIZATION OF AN AVATAR SYMBIOTIC SOCIETY

- how to make human-avatar interactions or avatar-avatar interactions smooth & comfortable
- moral computing detecting & suppressing immoral behaviors of humans (avatar operator / avatar user)

TECHNICAL CHALLENGES & ETHICAL CONCERNS

- like diminishing one's autonomy



Glenda Hannibal & Felix Lindner:

Towards A Questions-Centered Approach to Explainable Human-Robot Interaction (xHRI)



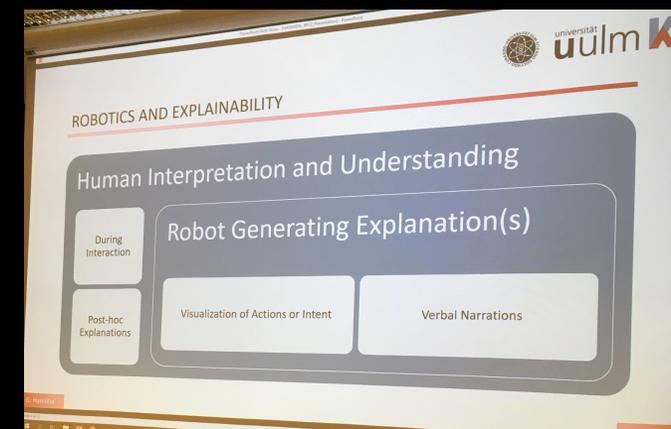
ROBOTS WITH THE CAPACITY TO EXPLAIN THEIR REASONING & BEHAVIOR TO HUMANS DURING OR AFTER HMIs

CHALLENGES:

- burden of interpretation | dynamic exchanges | ethical concerns
- default transparency versus appearance of social agency

QUESTION-CENTERED APPROACH

- explaining for the purpose of transparency should be considered a social activity
- triggered upon request



Session 11: Explainable Robotics



Guglielmo Papagni & Sabine Theresia Koeszegi:

Explaining Intentional and Unintentional Behavior: Social Norms for Explainable Robots

How to distinguish attributing intentions and really having intentions.

1. INTENTIONAL BEHAVIOR – PERCEIVED AS INTENTIONAL
2. UNINTENTIONAL BEHAVIOR – PERCEIVED AS INTENTIONAL
3. INTENTIONAL BEHAVIOR – PERCEIVED AS UNINTENTIONAL
4. UNINTENTIONAL BEHAVIOR – PERCEIVED AS UNINTENTIONAL

		Human perception	
Robot behavior	The robot behaves according to its objective and the user perceives the behavior as intentional	The robot behaves according to its objective but the user does not perceive the behavior as intentional	
	The robot's behavior is accidental but the user thinks it is aligned with the robot's objective	The robot's behavior is accidental and the user perceives it as unintentional behavior	



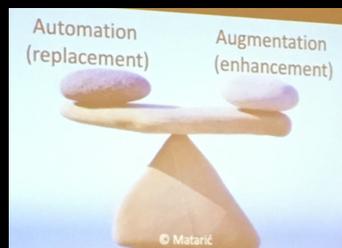
Maja Matarić:

Socially Assistive Robotics - Methods and Implications for the Future of Work and

personalized robots in everyday domains: workplaces, schools, healthcare contexts, and homes

current pandemic: caused and exposed unprecedented levels of health & wellness, education, and training needs worldwide

Socially assistive robotics has the potential to contribute significantly to addressing those challenges



MANY CHALLENGES

embodiment:
person like / animal like / object like

engagement / feedback
personality / deception

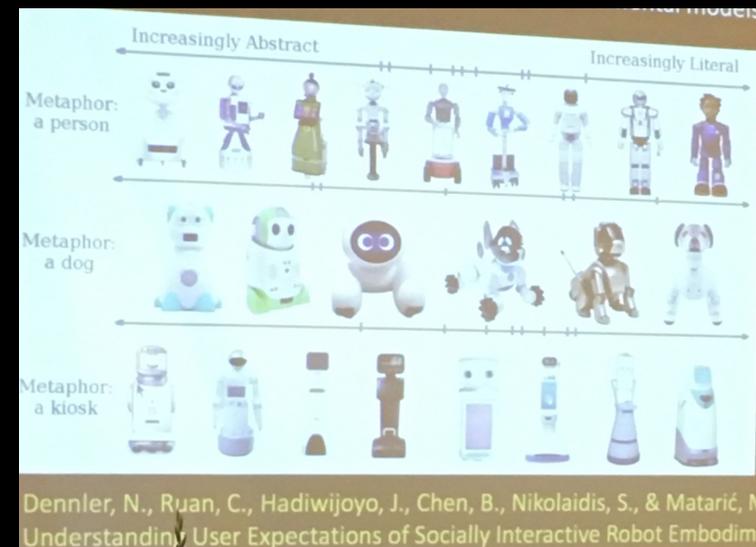
HRI METHODS FOR SOCIALLY ASSISTIVE ROBOTICS

• UTILIZE

- multi-modal interaction data (*eye-gaze, turn-taking*)
- expressive & persuasive robot behavior (*motivate behavior by triggering imitation*)
- include modeling, learning, personalizing user motivation, engagement, coaching
- augmented reality

• GOALS

- monitor, coach, motivate users to engage in health, wellness, education and training activities



<https://interaction-lab.github.io/robot-metaphors/explore/>



ACCOUNT OF THE NATURE (ACTUAL & POTENTIAL) OF ROBOTS
ACCOUNT OF INSTITUTIONAL ROLES AND OF INSTITUTIONS

Seumas Miller: Robots, Institutional Roles and Functions

NORMATIVE-TELEOLOGICAL THEORIES OF INSTITUTIONS

INSTITUTIONS = JOINT ENTERPRISES IN THE SERVICE OF COLLECTIVE ENDS

- multi-layered structures of joint actions
- chains of joint activity
- reproducing themselves by means of roles occupied by human beings

collective ends:=

- food, clean water, health care, education, ...

means:=

- technology (*Robots are increasingly being used in institutional settings*)

TO WHAT EXTENT CAN ROBOTS ADEQUATELY DISCHARGE INSTITUTIONAL ROLES? WHAT ARE THE LIMITATIONS OF ROBOTS IN RELATION TO INSTITUTIONAL ROLES?

What robots can:

- perform tasks formerly performed by humans | perform new tasks | perform jointly with human beings | assist humans | be symbols, proxies, simulacra
- **be organizational role occupants which are technological means in the service of the collective end**

- to act jointly it is not necessary to presuppose
- full-blown intentional moral agency / consciousness / minds

What robots cannot / have not:

- have no moral value qua particulars (*replaceable without moral loss*)
- ultimate-ends are programmed-in (chosen by humans)
- no consciousness, no general intelligence, do not care about, not sensitive to moral properties per se
- **cannot exercise moral judgements which are necessary for having an institutional role**
- **cannot be stakeholder in determining collective ends**
- **cannot be accountable (morally responsible)**



Kirsikka Kaipainen, Salla Jarske, Kaisa Väänänen:

Identifying Opportunities for Social Robots in Youth Services: A Case Study of a Youth Guidance Center

CASE STUDY ABOUT CENTRAL VALUES & NEEDS OF CLIENTS & STAFF IN A YOUTH GUIDANCE SERVICES TO IDENTIFY POSSIBLE OPPORTUNITIES FOR SOCIAL INTERACTIONS ENABLED BY ROBOTS

METHODS:

- context study
- interview with a staff member
- online questionnaire for young people (n=8)

RESULT: alleviating anxiety & enabling participation

PROBLEM: limited resources & time to address all needs especially related to social interaction

Identifying opportunities for social robots: drafting an experience journey of a young person's visit



- possible social robotic concepts
- implications on design process
- institutional practices
- values can be formulated as experience goals to guide the design process

Findings: Context, values and unmet needs

Staff perspective

- Physical context: informal, relaxed space
- Social context: service advisor usually greets clients at the door; other clients may be present in the space
- Employment, education and well-being are the most common reasons for clients' visits
- **Unmet needs:**
 - Conversational support for clients
 - Systematically collected feedback for staff

Youth perspective

- Most common values (n=7): **compassion, encouragement, respect, honesty, safety**
 - E.g. "Compassion manifests in that youth are understood and not pushed forcefully to some direction just because they have to be active and efficient citizens"
- **Unmet needs:**
 - Conversational support
 - Anxiety alleviation
- Two typical experience stories (next slides)

Laetitia Tanqueray & Stefan Larsson: What Norms Are Social Robots Reflecting? A Socio-Legal Exploration on HRI Developers



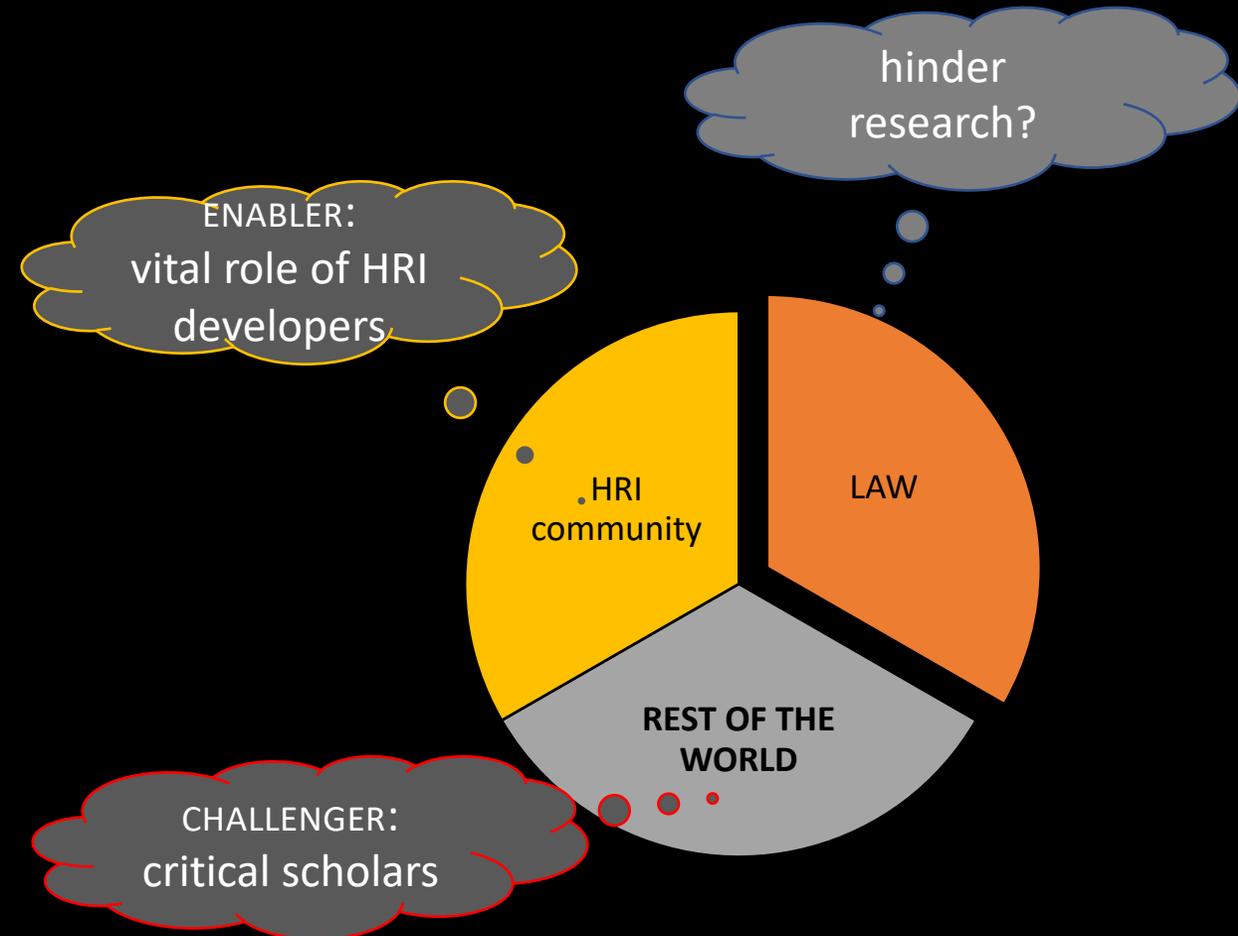
DEVELOPING MIRRORING NORMS (MERGING THEORY FROM SOCIOLOGY OF LAW AND DATA FEMINISM)

study investigating norms driving development in human-robot interaction

METHODS

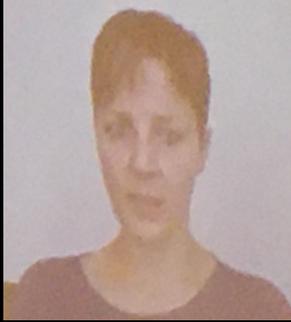
- ethnography of the HRI Conference 2021
- expert interviews
- qualitative coding of data (ethnography content analysis: ECA)

This socio-legal lens enables to pinpoint the lack of clear legal involvement, the reliance on the HRI community to develop, and the normative impact this has on the overall development of social robots.



Karolina Zawieska:

HRI: From Interaction to (Lived) Experience



SHIFT TOWARDS 'INTERACTION EXPERIENCE'

THE WAY WE UNDERSTAND, DESIGN AND STUDY HUMAN-ROBOT INTERACTIONS INEVITABLY CHANGES

preliminary findings of the ethnographic study involving robot developers

→ shift in the HRI field from thinking of human engagement with robots in terms of 'interaction' towards that of 'experience'

WHY?

→ due to an increasing use of robotic systems in the real-world environments

→ long-term user engagement with robots

→ degree of human-likeness in social robots

➤ not just adding a User Experience (UX) perspective to the HRI research

IMPLICATIONS FOR HOW WE UNDERSTAND ETHICS IN ROBOTICS

phenomenological perspective

'lived experience' & 'lived ethics' → 'a good life'

Shift in HRI/social robotics:
User Experience -> Human Experience
Roboethics -> Everyday Ethics
End-user/Customer -> Person



Catrin Misselhorn: Three Ethical Arguments against Killer Robots

WHAT ARE LETHAL, AUTONOMOUS WEAPON SYSTEMS (LAWS) – “KILLER ROBOTS” ?

weapon systems capable of selecting & attacking military targets without human

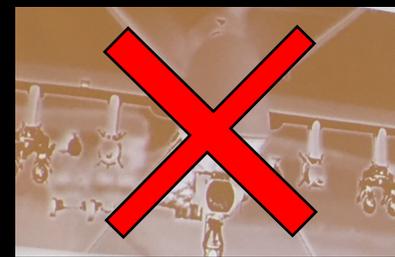
3 objections against LAWS:

- 1. RESPONSIBILITY GAP (*Sparrow*)
- 2. HUMAN AGENCY (*Leveringhaus*)
- 3. MORAL OBLIGATION (*Misselhorn*)

**LAWS have morally bad consequences
&
their use is morally problematic in itself.**

CONDITIONS FOR ATTRIBUTING RESPONSIBILITY

- 1. intentionality
- 2. awareness of the consequences
- 3. right kind of causal control
- 4. free will



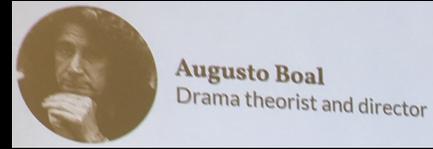
*Is human agency compromised by technology?
Maybe we lost control already?*

Session 20: Robots in Art

Elaheh Sanoubari, Amanda Johnson, John Muñoz, Andrew Houston, Kerstin Dautenhahn:
Using Robot-Mediated Applied Drama to Foster Anti-Bullying Peer Support



Robot-Mediated Applied Drama RMAD



Augusto Boal
Drama theorist and director

- a medium for safely exploring sensitive topics with children

RE-Mind1: an anti-bullying game

- theatrical exercise: spectators are invited to become “SPECT-ACTORS”
- watch the performance twice; the 2nd time: stop the performance & change its direction
 1. observe a bullying scenario between two robots,
 2. intervene by controlling a 3rd robot (the bystander) → practicing intervention strategies

→ engages peers in situated learning

WHY ROBOTS?

- provides a buffer for participants to safely explore sensitive topics in a private setting
- not practicing bullying

ASSIMILATION VERSUS ACCOMMODATION (PIAGET)
→ APPLIED DRAMA IS A NATURAL VEHICLE FOR SITUATED LEARNING

TAKEAWAY
• RMAD is a promising a pedagogical tool



Catherine Botha: Creativity and AI: A Response to Boden

Can creativity be explained?

Two Questions

1. What counts as an artwork? – ontology
2. What (who?) counts as an artist? – creativity



Margaret A. Boden:

3 kinds of creativity making it surprising

1. combinational
2. exploratory
3. transformational



REAL creativity:
autonomy, intentionality, valuation emotion
and consciousness

(H-creativity) & (P-creativity) are added
→ anthropocentric!!

WE DON'T NEED TO SAVE CREATIVITY FOR HUMANS

