

# MEMORY SLICES BY ANNA STRASSER

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



## Tuesday, August 18

14.15-14:30 | Welcome  
**Introduction to RP2020**

14.30-15.45 | Plenary 1  
**Selma Šabanović**

16:00-19:00 | Workshop 1 | Clodic et al.  
**Social Models for Social Robotics**

19:30-21:30 | Workshop 2 | Gunkel et al.  
**Should Robots Have Standing?**

## Wednesday, August 19

9.00-10.15 | Plenary 2  
**Robert Sparrow**

10:30-11:15 | Discussion Session 1  
**THEORY I**

11:30-12:15 | Discussion Session 2  
**THEORY II**

15:00-15:45 | Discussion Session 3  
**ETHICS I**

16:00-16:45 | Discussion Session 4  
**ETHICS II**

17:00-18.15 | Plenary 3  
**Shannon Vallor**

## Thursday, August 20

9.00-10.15 | Plenary 4  
**Alan Winfield**

10:30-11:15 | Discussion Session 5  
**MORAL ROBOTS I**

11:15-12:00 | Discussion Session 6  
**MORAL ROBOTS II**

12:15-14:45 | Workshop 3  
**Trust in Robots and AI**

15:00-15:45 | Discussion Session 7  
**METHODS**

16:00-16:45 | Discussion Session 8  
**CASES**

17:00-18:15 | Plenary 5  
**Aimee Van Wynsberghe**

19:30-21:30 | Workshop 4  
**Think & Perform Tank**

## Friday, August 21

9.00-9:50 | Discussion Session 9  
**DESIGN**

9:55-10:40 | Discussion Session 10  
**PERSPECTIVES**

10:45-13:45 | Workshop 5  
**Robots in Religious Contexts**

14:00-15:15 | Plenary 6  
**John Danaher**

15:15-15:30 | Conclusion  
**Outlook to RP2022**



# DAY 1: Tuesday, 18 August

- Introduction
- Plenary 1: Selma Šabanović
- Workshop I
- Workshop II

# Designing "Companion Artifacts": The Relational Construction of Culture & Technology in Social Robotics

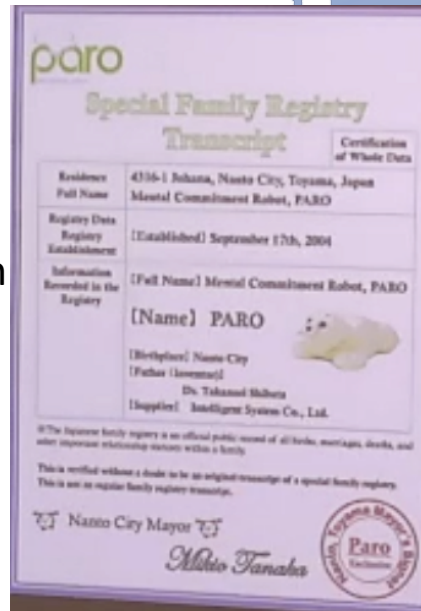
*co-evolving relationship*

**Social robots:**  
designed to coexist with  
people & learn through  
their interactions

**We:**  
are expected to develop ways of  
behaving, communicating, and  
organizing that support robots

*explore social robots as "companion artifacts"*

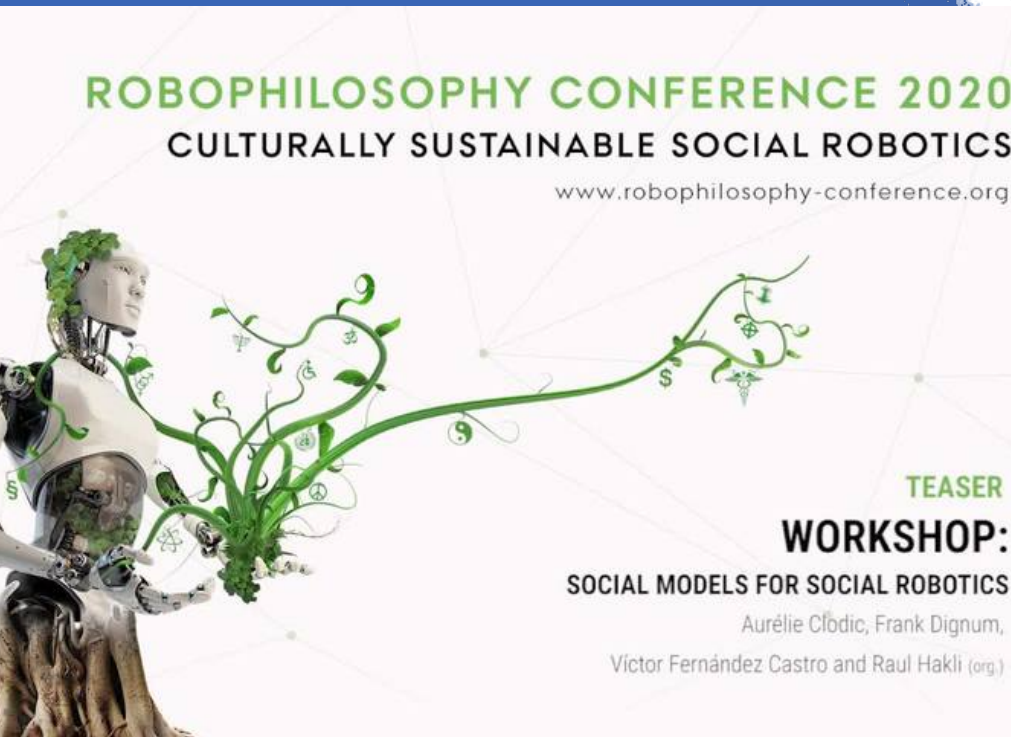
- how our concepts of self, cultural practices, social organizations, and sociotechnical infrastructures are co-constructed with existing and imagined social robots
- how "Japanese culture" is assembled in relation to social robots
- what it means to "domesticate" robotic technologies
- how community-based methods can incorporate diverse sociocultural values into social robotics.



**PLENARY 1**

# SOCIAL MODELS FOR SOCIAL ROBOTICS

Aurélie Clodic, LAAS CNRS, University of Toulouse (FR) - Frank Dignum, Umea University (SE) - Víctor Fernández Castro, LAAS CNRS, University of Toulouse (FR) - Raul Hakli, University of Helsinki (FI)



## SOCIAL ROBOTICS:

- technologies with potentially profound socio-cultural impact

## CURRENT INTERDISCIPLINARY RESEARCH OF "SOCIAL ROBOTICS" & "HUMAN ROBOT INTERACTION" (HRI):

- not yet equipped with the necessary conceptual tools in order to design interactions between humans & robots
- new approaches are needed
- taking larger ethical principles & socio-cultural perspectives into account

clarify questions arising with this new technology

- *How far can robots go—now and in the future—to fulfill the requirements of full-blown social agents?*
- *How and where do ethical requirements dovetail with the elements (conditions, principles, and procedures) for social agency?*



Helena A. Frijnsa & Oliver Schürer, TU Wien (AT)

## Context-Awareness for Social Robots



engineering approaches to developing context-awareness, social context

- as an object of study in human-centred design
- as a cognitive component of interaction (part of meaning negotiation / sensemaking / theory of mind)

→ definition of context for social robots

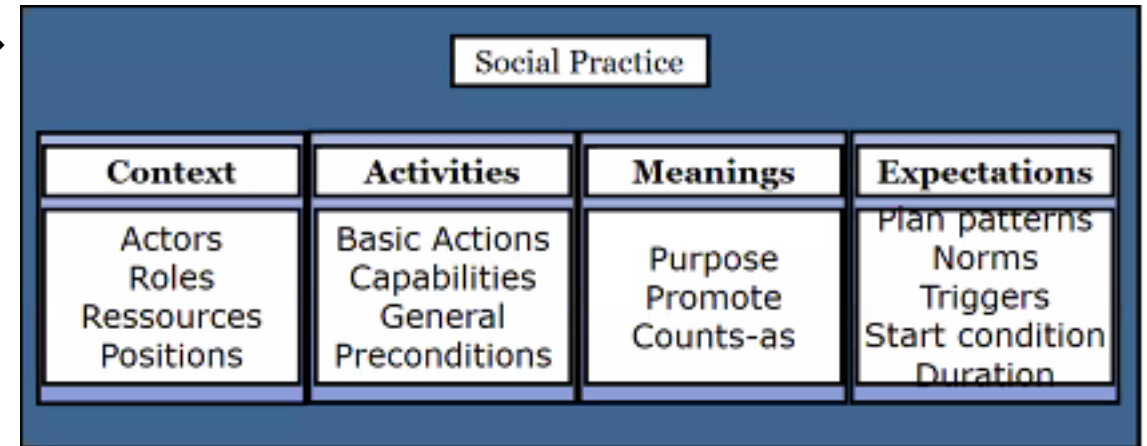
- Context consists of environmental, social (including task-based) and agential features (in short, situational features) that are relevant to predicting, determining and explaining the behaviour of an agent.
- For a system to be context-aware in a cognitive sense, this implies the capability to interpret information and its relevance to behaviour, and negotiate the meaning of actions and events with other agents. Context changes dynamically as (inter)actions open up new possibilities for (inter)actions.

... discuss design implications

**Social behavior is not an extra behavior that is attached on top of the basic robot behavior**

- real social situation where robots interact with people in a dynamic environment

→ robots have to sustain some social relation over time



features real world applications need

not-rational (ability to hold inconsistent beliefs)

combine designed preferences with observation and social change

combine reasoning and learning

context-based pragmatic reasoning

pursue incompatible goals and requirements

Frank Dignum, Umea University (SE)

## Robots Based on Social Contexts



Dane L. Gogoshin , University of Helsinki (FI)

# Robots as Ideal Moral Agents per the Moral Responsibility System



prevailing view = robots cannot be full-blown members of the larger human moral community

**HERE: robots would be ideal moral agents in the way that currently counts**

- robots may fail to meet a number of criteria which human agents meet (at least in theory)
- BUT robots earn a perfect score as far as the conception of behavioristic moral agency at work in our moral responsibility practices goes



Guy Hoffman, Cornell University (US)

# The Social Uncanniness of Robotic Companions



**Social robots offer a promise of relatedness without the limitations of human ability and availability**

- BUT social uncanniness of robotic companions
- social robots violate deeply rooted psychological needs:
  - sense of being unique individuals that can not easily be replaced or replicated
  - fear of dying and of disappearing
  - worry that by relating to robots we may lose our ability to relate appropriately to other humans.

4 lessons:

Conquering our fragility by making perfect stand-ins will come back to haunt us

Authenticity cannot be replicated, and reproductions lose value

Lack of transience will cause us to take things for granted

Clean interaction design teaches us the wrong lessons about human relationships

Víctor Fernández Castro, LAAS CNRS,  
University of Toulouse (FR) / Raul Hakli,  
University of Helsinki (FI) / Aurélie Clodic,  
LAAS CNRS, University of Toulouse (FR)

## What does it take to be a social agent?



### A PHILOSOPHICALLY INSPIRED LIST OF MINIMAL REQUIREMENTS FOR SOCIAL AGENCY

extend intentional stance into a social stance

→ serve as a guideline for social robotics

- no details about the cognitive processes behind sociality
- providing an implementation-free characterization of the capacities & skills associated with sociality

→ 4 requirement:

- reacting to humans / tendency for interaction / normative understanding / normative evaluation & regulation
- potential benefits of understanding them in an instrumentalist way in the context of social robotics

# SHOULD ROBOTS HAVE STANDING?

David J. Gunkel, Northern Illinois University (US)



**Should social robots have moral and/or legal status.**

→ Christopher Stone's investigation of the status of natural objects

- apply this method to technological artifacts that have been designed to elicit and actualize social presence
- interdisciplinary scholars/educators from across the globe, who have extensive experience researching the social impact and consequences of robots,
- → to develop concrete guidelines for responding to and taking responsibility for the design and deployment of culturally sustainable social robots



Autumn Edwards & Chad Edwards,  
Western Michigan University (US)

## Who or What is to Blame? Personality and Situational Attributions of Robot Behavior

### *Fundamental Attribution Error (FAE):*

- over-emphasize dispositional explanations for others' behavior
- under-emphasizing situational explanations people do assign agency, intentionality, blame to robots

→ Do people commit FAE in response to the behaviors of a social robot?

experiment: forced to behave badly is judged as being not responsible

- towards humans: FAE: no choice → dispositional attribution → responsible
- towards robots: no choice → dispositional attribution → not responsible



WORKSHOP 2

Martin Cunneen

University of Limerick (IE)

# Could Autonomous Vehicles become Accidental Autonomous Moral Machines?

## Moral or not moral?

1. autonomous vehicles are moral machines by building their decision architecture on necessary risk quantification
2. autonomous vehicles are inadequate moral machines  
moral inadequacy → significant risks to society

key concepts in Autonomous Vehicle decisionality literature

- reframe the problem of moral machine
  - access meta questions that underlie Autonomous vehicles as machines making high value decisions regarding human welfare and life

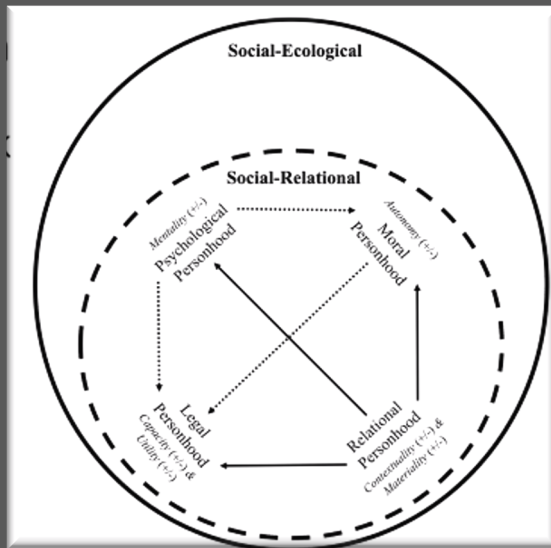


**WORKSHOP 2**

Joshua Gellers University of North  
Florida (US)

# Greening the Machine

## Question: Towards an Ecological Framework for Assessing Robot Rights



## Can robots have rights?

macro-level trends

- robots → human-like in appearance & behavior
- legal systems → recognizing the rights of non-human entities
- interdisciplinary framework for evaluating the conditions under which some robots might be considered eligible for certain rights

❖ critical, materialist, broadly ecological interpretation of the environment

❖ + decisions by jurists establishing or upholding the rights of nature

→ support extension of such rights to non-human entities like certain robots



**WORKSHOP 2**



Anne Gerdes The University of  
Southern Denmark (DK)

# Do We Need to Understand Social Robots to Grant them Rights?



## DELEGATING RESPONSIBILITY TO A SOCIAL ROBOT & GRANTING RIGHTS AS MORAL PATIENT

normative implications of epistemic  
opacity

- *Will we trust a robot that produces morally apt responses to morally challenging situations?*
- *Will we consider it worthy of moral consideration?*

### IF:

- robot's rationale escapes human understanding (non-explainable AI)?
- robot's inner workings equal deep learning networks, which produces black-box (non-interpretable AI)?



WORKSHOP 2

David J. Gunkel Northern Illinois  
University (US)

## The Rights of (Social) Robots

### Should robots have standing?

- recent publications examine the concept of robot rights largely theoretical and speculative
- move debate about the moral & legal standing of social robots out of the realm of theory

→ *What rights social robots need to have in order to facilitate responsible integration of these technologies into our world?*

- human rights  $\neq$  robot rights



Various robotic devices and AI implementations might need some form of social protections; entertaining this exigency is an important component in on-going efforts to test, validate, and even revise the limits of our moral and legal systems

### Should robots have standing?



The very idea of AI or robots being accorded anything approaching moral or legal status is not just wrong-headed thinking but a dangerous development that should be severely curtailed or shut down before it even begins.

**WORKSHOP 2**

## DAY 2: Wednesday, 19 August

- Plenary 2: Robert Sparrow
- Theory I
- Theory II
- Ethics I
- Ethics II
- Plenary 3: Shannon Vallor

TEASER  
**ROBERT SPARROW**  
 MONASH UNIVERSITY (AU)



ROBOPHILOSOPHY CONFERENCE  
 CULTURALLY SUSTAINABLE

# What robots represent ... and why it matters

ENGINEERS NEED TO THINK ABOUT THE POLITICS OF THE MEANINGS  
 THEIR ROBOTS RELY ON AND CONVEY

- robots operate at the level of meaning and not just mechanism
- *embodied nature of robots*: function as icons → convey meanings in a different manner to other media forms
  - more power to shape behavior than films / pictures
- *representational content of social robots*: → difficult ethical dilemmas (sexist and racial bias)



Figure 1. Cluster 1: Robovie MR2, Cosmobot, Autom, Papero, Riba, Nao, Asimo, Atom & Leonardo



Figure 2. Cluster 2: Asoy, ICat, Snackbot, Dynamoid, Ri-Man & Wakamaru



Figure 3. Cluster 3 (Geminoid HI-1 & Ibn Sina) and Cluster 4 (Geminoid DK & HRP-4c).



Figure 4. Cluster 5: PR2, Justin, Robonova, Robosapien, REEM-1390, Wabian, Kobian, HRP2, HRP3, REEM-1, Hwarang



Figure 5. Cluster 6: Mika, Luna, Olivia, Lucas, Kismet, Armar, Popo, Phoebe, Robo-Head, Twendyone

von der Pütten, A. and Krämer, N., 2012, March. A survey on robot appearances. In *7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2012 (pp. 267-268). IEEE.



*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

## THEORY I

Conceptual Tools for  
Describing Human-Robot  
Interaction:  
Agency, Intelligence,  
Animacy, Sociality

### **Louis Longin**

- [Towards a Middle-Ground Theory of Agency for Artificial Intelligence](#)

### **Henning Mayer**

- [Robolabs: Building Surfaces of Intelligibility](#)

### **Rebekka Soma**

- [Help, There's a Single Celled Giant in My Back Yard!](#)

### **Johanna Seibt, Malene Damholdt, and Christina Vestergaard**

- [Sociomorphing, Not anthropomorphizing](#)

### **Tom Poljansek and Tobias Störzinger**

- [Trans-Human-Formers](#)

*Q2: In your contribution you suggest new concepts or descriptive tools—which methods and strategy do you use to develop these new descriptive tools and how do you justify choosing these over others?*

*Q3: How would you validate your proposal, what would count as confirmation or disconfirmation of the usefulness or*

Louis Longin

# Towards a Middle-Ground Theory of Agency for Artificial Intelligence

AI systems' ability to achieve higher-level mental states & the ethics of implementation

## **Are AI systems capable of action?**

Anscombe & Davidson →  
intentional mental states

cognitive & computational scientists  
(Beer or Pfeifer) → reduce agency to  
mere behavior

HERE:

gradual concept of agency because both traditional  
concepts of agency fail to differentiate the agential  
capacities of AI systems



# Henning Mayer

## Robolabs: Building Surfaces of Intelligibility

### humanoid robots or chatbots

thinking of them is  
on the surface

their promises are  
on the deep end

- superficial activity data used to make emotions & attitudes available - mostly based on technologies that are not introspective

### analyzing the interconnectivity of intelligence types

- designs of the coupling of algorithmic & hermeneutic intelligences
  - arranged in a way covering different types of intelligence



Rebekka Soma

## Help, There's a Single Celled Giant in My Back Yard!

### PERCEPTION OF ANIMACY

- many abilities tied to animacy are projected onto robotic systems
- non-trivial task of reporting from the experience of perception of animacy in human-robot encounters
  - starting out from a wonder at how it can be that a robotic lawnmower phenomenally appear as “something it is not, doing things it isn't doing”
- Sheets-Johnstone's animate form
  - combining Gibson's ecological psychology with Uexküll's Umwelt theory
  - looking at one perspective on movement, or a false self-movement, as a disguise





# Johanna Seibt, Malene Damholdt, Christina Vestergaard

## Sociomorphing, Not anthropomorphizing

1. Culturally sustainable social robotics requires that we focus on the human side of human robot interactions.
2. We should relinquish presumption that *all* human social interactions are based on 'anthropomorphizing', i.e., the projection of fictional human capacities.
3. Many human social interactions with non-human agents are better understood as sociomorphing, the direct perception of actual non-human social capacities.
4. OASIS allows us to explore how we sociomorph robots, and how this relates to different types of experienced sociality.



# Tom Poljansek and Tobias Störzinger

## Trans-Human-Formers

### **human users can engage in relationships with “social robots”**

- wide-ranging debate about technical, theoretical as well as ethical problems
  - practice-and-perception-theoretical framework (PaP-framework)
- analyze social interaction in general and human-robot-interaction
  - design social robots in way that does not need to mimic patterns of interaction that we are already familiar with from our everyday interactions with other human and animal beings.

functional versus phenomenological perspective (I perspective and me perspective)



*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

## THEORY II

Conceptual Tools for  
Describing Human-Robot  
Interaction:  
Empathy, Imagination, and  
Otherness

**Joanna Malinkowska**

- [The Growing Need for Reliable Conceptual Analysis in HRI Studies](#)

**Matthew Rueben, Eitan Rothberg, and Maja Mataric**

- [Applying the Theory of Make-believe to Human-Robot Interaction](#)

**Peter Remmers**

- [The Artificial Nature of Social Robots](#)

**Chris Chesher and David Silvera Tawil**

- [Robot Handshake in Space](#)

*Q2: Do we need more than neuroscience explanations for why we tend to accept so-called 'social' robots as social agents, and how do you think we could empirically test the veracity or usefulness of your suggestion?*

*Q3: How should we develop the most suitable design for a social robot? Should we try to create something that is neither anthropomorph nor zoomorph, but requires us to develop an entirely new mental model to understand it?*

### **interdisciplinary nature of the field of HRI**

- using various concepts: typical of social & humanities + associated with technology
  - term 'empathy'
    - describe & explain interaction between robots & humans
- raises questions about the possibility of applying this term in situations in which only one of the participants of the interaction is a traditionally understood social subject
- requires answers to questions about such problematic concepts as values and culture





# Matthew Rueben, Eitan Rothberg, and Maja Mataric

## Applying the Theory of Make-believe to Human-Robot Interaction

### **APPARENT PARADOX :**

### **PEOPLE OFTEN MAKE ASCRIPTIONS THAT THEY KNOW TO BE LITERALLY FALSE**

- robots may be treated as if it were a dog, or as if it had certain intentions, emotions, or personality traits.

How can one do this while also believing that robots cannot really have such traits?

Kendall Walton: theory of make-believe

→ extensions to Walton's theory, some implications for how we make attributions and use mental models

→ informal account of human-robot interaction from the human's perspective



## **TWO APPROACHES TO ANTHROPOMORPHISM**

1. discourage to take human likeness as a way to integrate robots in socio-technical environments - rational
2. novel and original approach to social robots – take them as social - visionary
  - transcends purely material status and affirms the special kinds of relations constituted in human-robot interaction

HERE:

mediating interpretation of the issue along the lines of phenomenological theory of image-perception



# Chris Chesher and David Silvera Tawil

## Robot Handshake in Space

### **ROBONAUT R2 & DAN BURBANK PERFORMED THE FIRST HUMAN-HUMANOID HANDSHAKE IN SPACE (FEB 2012)**

- welcomed R2 as a crew member, engaging with the robot as a social agent rather than just a thing to be examined or controlled
  - Heidegger's terms: experienced neither theoretically as present-at-hand nor practically as ready-to-hand
- experienced live-to-hand, given respect as an other
  - R2 is capable not only of executing programs, but also playing its part in socially choreographed rituals and everyday performances within a social gestalt



*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

# ETHICS I

## Normative Issues of Human-Robot Interaction: Autonomy, Moral Agency/Patiency, Empathy

Shuhong Li, Aimee van Wynsberghe, and Sabine Roeser

- [The Complexity of Autonomy](#)

Anna Strasser

- [Social Norms for Artificial Systems](#)

Dina Babushkina

- [Robot to Blame](#)

Oliver Quick

- [Sympathetic Robots](#)

*Q2: To what extent are your recommendations or observations context-dependent, and how does this affect the possibility of doing robot ethics?*

*Q3: If you could act as ethical advisor for the developer team of a social robotics application, how would you conceive of your collaborative role and how would you structure your interaction with the team?*



# Shuhong Li, Aimee van Wynsberghe, Sabine Roeser:

## The Complexity of Autonomy

ELDERLY CARE RECEIVERS ARE EXTENSIVELY AND PROFOUNDLY AFFECTED BY INTERACTING WITH CARE ROBOTS

**autonomy** = a complex core value in elderly care

- few studies have been able to address this complexity in elderly care in the robot era

to illustrate the complexity

→ taxonomy of autonomy

→ employed as a tool to evaluate ethical aspects of the effects of care robots on autonomy in elderly care

this taxonomy is instrumental for impact assessments of care robots on care receivers' autonomy both retrospectively and prospectively



# Anna Strasser

## Social Norms for Artificial Systems

### REASONS TO ARGUE FOR SOCIAL NORMS REGULATING OUR BEHAVIOR TOWARDS ARTIFICIAL AGENTS

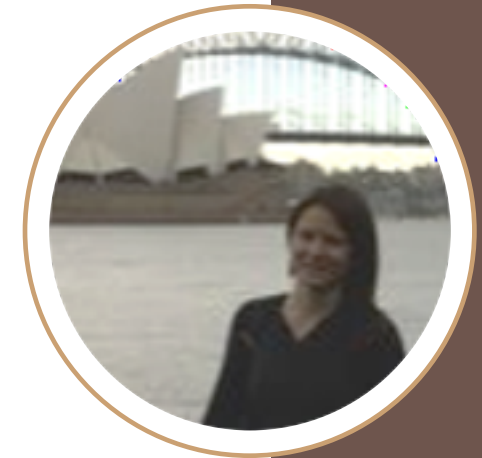
- problematizing the assertion that moral agency is a necessary prerequisite for moral patiency
    - reasons which are independent of attributing moral agency
- consequentialist strategy:
- analyzing potential negative impacts of human-machine interactions
  - focusing on factors that support a transfer of behavioral patterns from human-machine interactions to human-human interactions
  - introduce social norms to avoid transfers which lead to cruelty among humans



# Dina Babushkina: [Robot to Blame](#)

## PRESERVE RATIONALLY JUSTIFIABLE MORAL PRACTICES

- which face challenges because of the increasing integration of social robots into roles previously occupied exclusively by persons
  - focus on the attribution of blaming as an example of such practices
- blaming robots
- (a) does not satisfy the rational constraints on the reactive attitude of blame and other related reactive attitudes and practices such as resentment, forgiving, and punishment
  - (b) is by itself morally wrong



# Oliver Quick: [Sympathetic Robots](#)

- GENERAL CHALLENGES FOR THE DESIGN OF SYMPATHETIC ROBOTS
  - lack of conceptual clarity
  - difficulty of quantifying an instance of sympathy

*Candace Clark's sociological account of sympathy*

- a solid theoretical basis for advancing the design and acceptance of sympathetic robots
- some potential challenges in its implementation
  - e.g. alarming potential for sympathetic robots to act ideologically imperialistic, or to 'nudge' users across cultural boundaries





*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

## ETHICS II

### Normative Issues of Human-Robot Interaction: Nudging and Coaching

- John Sullins and Sean Dougherty: [Ethical Nudging of Users While They Interact with Robots](#)
- Raffaele Rodogno: [Who Is Afraid of Robotic Nudges?](#)
- Markus Westberg and Monika Jingar: [Preserving Personal Perspectives in Coaching Technology](#)

*Q2: To what extent should a discussion of nudging pay particular attention to the special affordances of social robots who present themselves neither as things nor as persons but as social others? To what extent should a discussion of nudging pay particular attention to the special affordances of social robots who present themselves neither as things nor as persons but as social others?*

*Q3: Is the stance on nudging that you take dependent on a particular normative ethics? (Just taking the variety of Western ethics as example, would not deontologists, consequentialists, and virtue ethicists disagree about the admissibility of nudging by social others?) If so, could this send a problematic message to engineers?*

# John Sullins and Sean Dougherty

## Ethical Nudging of Users While They Interact with Robots

### ETHICAL STANDARDS FOR ROBOTIC NUDGING SYSTEMS

potential for moral harm of ill designed robotic nudging systems can cause

- ethical argument to justify certain kinds of permissible robot nudges
- actionable standards to ensure highest ethical standards of building robotic systems

→ <https://standards.ieee.org/project/7008.html>



Raffaele Rodogno

## Who Is Afraid of Robotic Nudges?

- looking on permissibility of nudging by robots from the perspective of liberalism
  - liberalism has less than univocal answers to this question
- politics of nudging by robots will in the end hinge on the way in which autonomy is understood and is valued as a political value



# Markus Westberg and Monika Jingar

## Preserving Personal Perspectives in Coaching Technology

ROBOTS AND DIGITAL AGENTS FIND THEIR WAY IN AN INCREASING NUMBER OF AREAS IN OUR EVERYDAY LIVES.

- history and impact of digital companions and coaching devices on our health and lifestyle
- exploring how a growing entanglement with devices that quantify our lives may reduce or invalidate the user's personal experiences and preferences.
- proposing that the kind of bond that is being formed is one
  - that feels nurturing
  - that the user can trust, in order to preserve personal values and autonomy



# Flourishing Isn't Free: Paying Down Our Moral and Social Debt in Robotic Systems

## TWO INADEQUATE POLES OF SOCIAL ROBOTS & ETHICS:

- (1) superficially utilitarian analyses of ethical 'risks' *bots (fail to question the underlying sociotechnical systems & values driving robotics development)*
- (2) speculative, empirically unfounded fears of robot-pocalypses *(leave those underlying systems and values unexamined and unchallenged)*

## NOW: NORMATIVELY RICHER & MORE EMPIRICALLY GROUNDED

- from risk-mitigation to deeper thinking about how to design different worlds in which we might flourish with social robots
- BUT we have to reckon the moral & social debt accumulated in existing robotics systems & our broader culture of sociotechnical innovation
  - explores legacies & accumulated debts, and what it will take to liberate social robotics from them





## DAY 3: Thursday, 20 August

- Plenary 4: Robert Sparrow
- Moral Robots I
- Moral Robots II
- Trust in Robots and AI
- Methods
- Cases
- Plenary 5: Aimee Van Wynsberghe
- Workshop: Think & Perform Tank

# "Why Did You Just Do That?" Explainability and Artificial Theory of Mind for Social Robots

## ASPECT OF TRANSPARENCY: EXPLAINABILITY

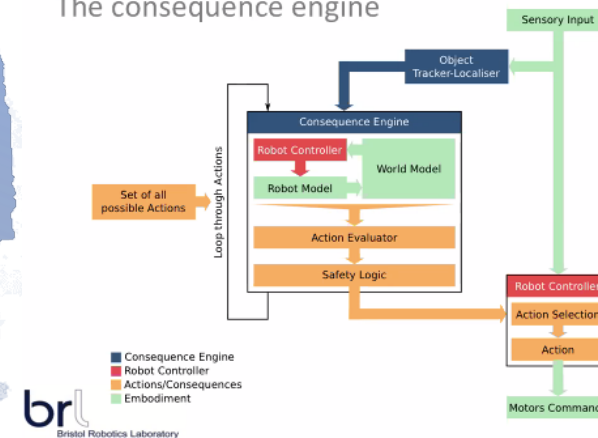
apply artificial ToM to the challenge of providing social robots with the ability to explain themselves

**HYPOTHESIS:** simulation-based internal models offer a powerful and realizable, theory-driven basis for artificial theory of mind



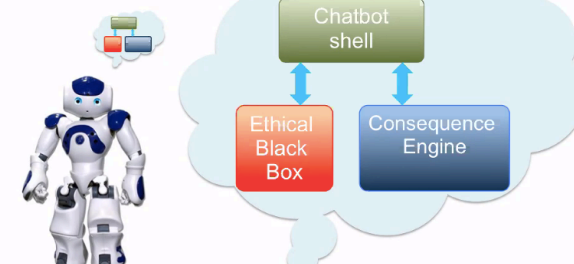
Winfield AFT (2018) Experiments in Artificial Theory of Mind: From Safety to Story-Telling. *Front. Robot. AI* 5:75. doi: 10.3389/frobt.2018.00075

The consequence engine



## Putting it all together

A proposed architecture for robots that can explain themselves...



... and help the user to develop an appropriate ToM towards the robot

*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

# MORAL ROBOTS I

## Generating Moral Behavior in Social Robots

*Q2: Why should we try to build moral robots?*

- Tomi Kokkonen: [Protomoral Machines](#)
- Aleksandra Kornienko: [Moral Machines](#)
- Felix Lindner: [Permissibility-Under-a-DescriptionReasoning for Deontological Robots](#)
- Selmer Bringsjord, Michael Giancola, Naveen Sundar Govindarajulu: [Toward Defeasible Multi-Operator Argumentation Systems for Culturally Aware Social Robots](#)
- David Miller: [Considering Human-Computer Moral Conflicts](#)

*Q3: The question of whether 'artificial morality' is possible depends on one's conception of what moral reasoning consists in and one's moral theory. How do we get around the problem that different moral theories yield different answers to moral questions, with the effect that 'artificial morality' is not a well-defined target?*

# Tomi Kokkonen

## Protomoral Machines

TWO CONNECTED QUESTIONS ABOUT MORAL AGENCY & ROBOTS:

- *How can we ensure that robots behave in accordance to relevant ethical considerations? Is it possible to have genuinely moral machines?*
- evolutionary perspective
  - importance of a middle-range perspective to morality of machines
- not restrict to
  - present-day perspective of current ethical concerns
  - far future theoretical issues concerning the possibility of genuine morality neither
- we should reflect on what it would mean to create protomoral machines



# Aleksandra Kornienko

## Moral Machines

AUTONOMOUS MACHINES ENTERING OUR EVERYDAY LIFE SHOULD NOT MAKE IT WORSE

- only kind of morality we have known is natural.

compare

biological and evolutionary aspects  
of human morality

main approaches to artificial morality

- two bringing up important aspects of natural morality that are of crucial consideration for the field of artificial morality if we want to make future human-robot societies flourishing and sustainable.





Felix Lindner

## Permissibility-Under-a-Description Reasoning for Deontological Robots

Deontological ethics: actions bear moral value which is not determined by the action's consequences but rather by intrinsic moral value of the action itself

→ action theory: actions are always performed under a description

- → integrated system that accomplishes inferences of action descriptions from concrete action sequences as executed by a robot

Using this system, specification of deontological values becomes very flexible



# Selmer Bringsjord, Michael Giancola, Naveen Sundar Govindarajulu

## Toward Defeasible Multi-Operator Argumentation Systems for Culturally Aware Social Robots

CONCEPTUAL FACT: ROBOTS (LIKE AI-INFUSED CARS, JETS, SPACESHIPS) WILL CARRY HUMANS INSIDE THEM



- case study: inconsistent attitude measurements → tragic crash of a jet & death of both pilots
- desiderata for an **automated defeasible inductive reasoner** able to suitably prevent such tragedies
- 1. first-order defeasible reasoner - OSCAR : can generate a partial solution to the dilemma the pilots couldn't conquer
- 2. address shortcomings of OSCAR relative to the desiderata  
→ more expressive reasoner based on an inductive defeasible multi-operator cognitive calculus (IDCEC) inspired by a merely deductive (monotonic) precursor (DCEC)



**solution exploits both the social & cultural aspects of the jet/robot  
we suggest be engineered in the future**



David Miller

## Considering Human-Computer Moral Conflicts

COMPUTER SYSTEM CAPABILITIES INCREASE → HUMAN-ROBOT  
CONFLICTS LIKEWISE GROW

systems

- must be designed with conflicts in mind
  - taking into account the type of conflicts expected, human and machine capabilities, and value tensions
- Moral conflicts present a special type of conflict, demanding research in to both human moral and ethical judgments and actions, and how human-computer moral conflicts will play out in various situations.



*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

# MORAL ROBOTS II

## Ethical Issues of Autonomous Driving

Manuel Dietrich

- [Understanding Autonomous Driving as Institutional Activity](#)

Judith Szalai

- [Sustaining the Higher-Level Principle of Equal Treatment in Autonomous Driving](#)

Martim Brandao

- [Discrimination Issues in Usage-based Insurance for Traditional and Autonomous Vehicles](#)

*Q2: If we try to resolve the problem of ego-centric decision making or structural discrimination in autonomous driving by appealing to principles of equity—will the discrimination problem not recur?*

*Q3: Are the ethical issues raised by autonomous driving new, and if so in which ways?*

## Understanding Autonomous Driving as Institutional Activity



### STRUCTURAL DISCRIMINATION INDUCED BY FUTURE AUTONOMOUS VEHICLES

- novel conceptual perspective on how to frame autonomous driving
  - to understand discrimination effects and potential strategies to reduce them
- autonomous driving = institutional activity
- enabling connections to theories about institutional justice
- open new ways to react to discriminatory concerns
  - e.g. sketch how to apply institutional justice to a system for autonomous highway driving



Judith Szalai

## Sustaining the Higher-Level Principle of Equal Treatment in Autonomous Driving



CULTURAL SUSTAINABILITY OF ARTIFICIAL INTELLIGENCE USE THROUGH  
AUTONOMOUS DRIVING

novel moral situation

- requiring specific, advance, reflectively endorsed, forced, and iterated choices, with yet uncharted forms of risk imposition

→ necessity and possibility of maintaining one of our most fundamental moral-cultural principles - equal treatment of persons

→ on a governmental level not individual



## USAGE-BASED POLICIES OF VEHICLE INSURANCE COMPANIES WHICH TRACK USERS TO ESTIMATE PREMIUMS

- can lead to indirect discrimination of sensitive personal characteristics of users
- negative impact in multiple personal freedoms
- contribute to reinforcing existing socio-economic inequalities

CLAIM: there is an incentive for autonomous vehicles

- to use similar insurance policies
- to anticipate new sources of indirect and structural discrimination
- analyzing the advantages & disadvantages of alternative insurance policies for AVs: no-fault compensation schemes, technical explainability and fairness, and national funds

# TRUST IN ROBOTS AND AI

## HOW TRUST IN AI CAN BE STUDIED & MAINTAINED

- multidimensional aspects of trust in technology
- encourage interaction and collaboration among early stage researchers.
- different dimensions of trust
  - algorithmic authority, normative practices, governance, interpretability and relatedness
- bringing together humanities, social sciences and HR  
→ establishment of sustainable socio-cultural values around robotics & AI





Laura Crompton  
University of Vienna (AT)

## A Critical Analysis of the Trust Human Agents Have in Computational and Embodied AI

TRUST - NEBULOUS BUZZWORD CENTERING AROUND AI

- human influenceability through AI  
→ trust in AI is to be seen as problematic
- notion of socio-technical epistemic systems
- trust in AI is strongly related to what could be understood as algorithmic authority
- trust should be seen critical!

**EPISTEMIC TRUST**



Michael Funk  
University of Vienna (AT)

## Gamification of Trust in HRI?

- *concept of credibility*  
specific focus on the implementation of ethical rules in robotic safety systems
  - cultural issues play a crucial role / cannot be controlled in a top-down approach
- focus on process-oriented bottom-up understanding of trust
- different social situations of normative practices  
a model for “gamifying trust” can combine in a transdisciplinary way philosophical and engineering points of view

**process-oriented  
bottom-up TRUST**





Jesse de Pagter,  
TU Wien (AT)

## Conceptualizing Trust in Objects of Speculation: A Narrative Approach to Robot Governance

- context of technology governance:
  1. robotics' speculative character → emerging technology is developed
    - pointing at the different expectations regarding its societal impact
  2. robots as speculative objects → explained as important to engage with
    - arguing for a narrative approach towards robot trajectories
  3. → analysis of trust building through technology governance
    - can engage with the speculative character of emerging robotics on a societal level



Guglielmo Papagni  
TU Wien (AT)

## Interpretable artificial agents and trust: Supporting a non-expert users perspective

ANALYSIS OF THE CONNECTIONS BETWEEN DIFFERENT FORMS OF  
INTERPRETABILITY OF ARTIFICIAL AGENTS

- influence of forms of interpretability to the development of trust on the side of non-expert users
  - semantic issues such as transparency & explainability
- implementation of forms of interpretability
  - highlighting their limits and advantages, in order to maximize understandability for non-expert users and artificial agents' trustworthiness



Isabel Schwaninger  
TU Wien (AT)

## On the Interplay of Relatedness and Trust in Situated Human-Robot Interaction in Older People's Living Spaces

ISSUES OF TRUST REGARDING ROBOTS TO SUPPORT OLDER PEOPLE

- trust: multidimensional & ephemeral  
→ no systematic understanding of trust in real-world contexts
  - empirical studies:
    - various interwoven forms of relatedness (among people / among people & institutions, places & technology)  
→ holistic understanding of trust
- provide insights on how to design for trust in situated human-robot interaction in older people's living spaces

*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

## METHODS

### How to Develop Culturally Sustainable Applications of Social Robotics

*Q2: How do you see the relationship between interactions and values, and how do you think we should approach value conflicts, for example, conflicting assessments of external evaluation and user experience?*

- Patrick Grüneberg: [Culturally Sustainable Social Robotics as Empowerment Technology](#)
- Mikkel Mørk, Christoffer Madsen, Sebastian Madsen: [From Values to Design Requirements](#)
- Gary Smith & Mark Ornelas: [Mind Reading in the Production of Culturally Sensitive Robot Behavior](#)
- Masoumeh Mansouri: [Can Current Methods in Knowledge Representation and Reasoning Make Robots Culturally Robust?](#)
- Anna Dobrosovestnova, Glenda Hannibal: [Human-Robot Co-Working and the Implications for Workplace Identity Performance](#)

*Q3: Where do you see the main obstacles for a concrete promotion or realization of your methodological recommendations?*

# Culturally Sustainable Social Robotics as Empowerment Technology

## IMPLEMENTATION OF CULTURALLY SUSTAINABLE SOCIAL ROBOTICS

→ HIGH REQUIREMENTS ON THE DESIGN OF SOCIAL HUMAN-ROBOT INTERACTION

- concept of empowerment technology (ET)
  - a value-driven framework to advance the interlocking of human value & computational modeling
- capability-based model of the interactive unity of human & machine
- case study: robotic childcare system

**culturally sustainable social robotics in terms of ET is possible if  
human empowerment values are addressed held by local  
stakeholders**



# Mikkel Mørk, Christoffer Madsen, & Sebastian Madsen

## From Values to Design Requirements

VALUE-DRIVEN DESIGN PROCESS BY USING THE METHODOLOGY OF ISR

- → preliminary structure for a design process consisting of an upstream and a downstream movement that translates values into design requirements.

two form a design loop ...

- iterations of this loop → satisfies the guiding-requirement
- suggest alternative strategies that might satisfy the guiding requirement and ensure a value-driven design





# Mind Reading in the Production of Culturally Sensitive Robot Behavior

PROBLEM OF HOW TO GET A ROBOT TO ACTUALLY BEHAVE IN A CULTURALLY SUSTAINABLE WAY

- to conform to cultural values robots must be equipped with the capability to mindread
- conditions of applying cultural norms
  - refer to the internal states of the agents taking part in the interaction
- → for an artificial agent to correctly apply a cultural norm, it must infer the internal states of other agents

**mindreading is essential for the production of behavior that respects human cultural expectations**



# Can Current Methods in Knowledge Representation and Reasoning Make Robots Culturally Robust?

CULTURALLY DIVERSE ROBOTICS RESEARCH IS OVERWHELMED BY CULTURAL STEREOTYPES

concept of Culturally Robust Robots accounting for the dynamic & flexible nature of culture

- grounded on an implicit assumption:
  - current AI methods are epistemologically adequate to represent & reason about “culture”
- questioning this by looking at two knowledge representation and reasoning (KR&R) methods used in intelligent robotics
- current methods are inadequate

**call for a critical revision of the use of KR&R in social robotics**



# Anna Dobrosovestnova, Glenda Hannibal

## Human-Robot Co-Working and the Implications for Workplace Identity Performance

CRITICALLY REFLECT ON THE PROMISES OF CO-WORKING WITH SOCIAL ROBOTS IN SERVICE INDUSTRIES

- important difference between human & robotic workers:
  - human capacity to perform an individual identity through work
- concepts of naturalization & normalization from feminist studies of workplace identity
- and discuss how social robots, "imprisoned" in one role: that of a friendly and consistent helper, may present psychological and political challenges to how human workers relate to and perform their workplace identity.



*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

*Q2: Can we generalize the insights gained from the observations of the case you present or are there any limitations?*

# CASES

## Challenges for Culturally Sustainable Robotics: Case Studies

- Kirsikka Kaipainen, Salla Jarske, Jari Varsaluoma, and Kaisa Väänänen: [Persuading Youth in Civic Participation with Social Robots: What Is Appropriate?](#)
- Kathleen Belhassein, Victor Fernandez, and Amandine Mayima: [A Horizontal Approach to Communication for Human-Robot Joint Action](#)
- F. Y. Jansen: ["How Nice That I Could Love Someone"](#)
- Kirsten Brukamp: [The Material Re-Turn of the Avatar—Computational Commemoration of the Deceased via Social Robots](#)
- Salla Jarske, Sanna Raudaskoski, and Kirsikka Kaipainen: [The “Social” of the Socially Interactive Robot](#)

*Q3: As you compare your research with topically related studies in HRI or social robotics, how would you describe the main methodological differences, and what might be gained by using your methodologies?*

# Kirsikka Kaipainen, Salla Jarske, Jari Varsaluoma, and Kaisa Väänänen

## Persuading Youth in Civic Participation with Social Robots: What Is Appropriate?

- study addressing the question: *What do youth consider as appropriate behavior for a civic robot?*
- 3 scenarios were evaluated in 3 workshops with 51 fifteen-year-old participants

	Scenario 1: Robot on a Climate Strike	Scenario 2: Interview with a Robot	Scenario 3: Follow the Robot
Context	School, outside lessons	Event for students	Shopping mall
Role	Climate striker	Reporter	Attractor
Appearance	Humanoid, stationary	Humanoid, stationary	Non-humanoid, mobile
Interaction	Speech, tablet, gestures	Speech, gestures, microphone prop	Gaze, sound, movement, text display
Emotions and behaviour	A: Neutral, assertive B: Depressed, anxious	A: Confident B: Nervous	A: Cautious, non-intrusive B: Panicked, intrusive

**credibility of the robot emerged as a central theme: a civic robot should emphasize its purpose and avoid any pretence – also when expressing emotions.**



# Kathleen Belhassein, Victor Fernandez, and Amandine Mayima

## A Horizontal Approach to Communication for Human-Robot Joint Action

- horizontal approach to the design of communication for joint action in human-robot interaction
  - different parameters of the whole joint action
    - context, embedded situation, human psychological profile during the design and test process

→ complementing standard building-block model that represents the state of the art in robotic communication

provide some general ideas of how the model can facilitate the use of available communicative strategies for creating more efficient culturally sustainable robots in contexts of joint action





F. Y. Jansen

## "How Nice That I Could Love Someone"

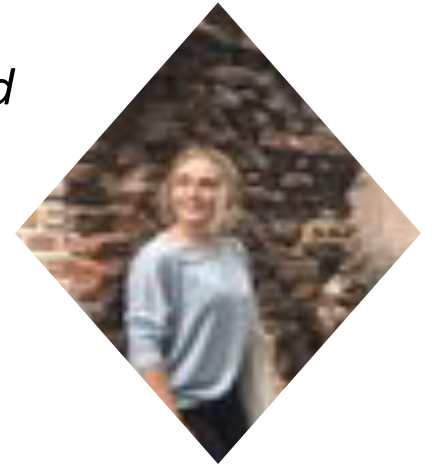
MEDIA HAVE A BIG IMPACT IN THE WAY HUMANS THINK ABOUT LIFE

- investigates the way film informs our ideas about a future where a loved one is replaced by something non-human which looks and acts human

*What happens to our understanding of life, ourselves and our relationships with others when we are able to build intimate relationships with an artificial intelligent human looking entity? - And what happens when this entity is an exact replica of a deceased partner?*

- film Marjorie Prime (2017) / episode Be Right Back (2013) from Black Mirror
- using a neoformalistic approach & rhetorical method of analysis to find out how these films are set up to convey a message and affect the spectator

**film can trigger critical thinking about human-machine relationships and can contribute to robophilosophy by analyzing film as virtual laboratories**



Kirsten Brukamp

## The Material Re-Turn of the Avatar–Computational Commemoration of the Deceased via Social Robots

### COMMEMORATION OF THE DEAD = CULTURAL PHENOMENON

- undergoes social & personal changes over time, involving changes that are mediated by technology

→ computational commemoration

- traits may be reenacted by a virtual or physical avatar
- construction of a material avatar yields a humanoid robot representing a deceased human.

case study regarding contemporary narratives aids in understanding the implications of future scenarios

- Serious challenges concern the realms of technology, sociology, psychology, law, and ethics.

# Salla Jarske, Sanna Raudaskoski, and Kirsikka Kaipainen

## The “Social” of the Socially Interactive Robot

- social robots project socially interactive skills including speech and gestures → normative practices that humans ordinarily rely upon in their everyday interactions with each other
- Social robots enable experiences
  - that are reducible to interaction as a normative practice, such as a sense of moral obligation to respond to a robot's greeting
  - consequences both for the user experience and the design of social robots that are currently overlooked

**theoretical-methodological tools from ethnomethodology should be applied to evaluate & investigate the experiences related to social interaction with social robots**



# Social Robots through the lens of care ethics

## ANALYSIS OF SOCIAL ROBOTS FROM A CARE ETHICS PERSPECTIVE



### relationality & reciprocity

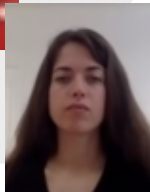
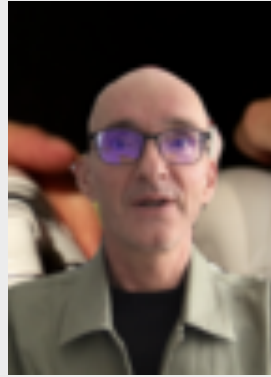
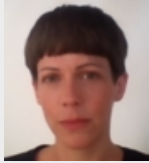
- impact of social robots on relational reciprocity
- humans are deceived into believing that the robot is deserving of reciprocity by the robot's appearance of responsiveness

### RISKS:

- re-direction of resources from humans towards robots
- social robots may threaten the ability to reciprocate to
- may weaken the incentive to give back to care workers

- 1 Designing for Reciprocity in Social Robotics
- 2 Care Ethics to give meaning to Reciprocity
- 3 Reciprocity in robots; what could go right?
- 4 What could go wrong? From micro to macro?
- 5 Re-thinking reciprocity in social robotics...

# Think & Perform Tank



EXPERIMENTAL, TRANSDISCIPLINARY RESEARCH ON CO-CREATION OF SOCIOCULTURAL PLACES AND SITUATIONS BY INTERACTION OF DANCER, MUSICIAN, ROBOTICISTS, ROBOT, AND THE AUDIENCE

- transformations triggered by robotic technologies → lead to robot cultures
- by their human-like shape, motions and behaviors social humanoid robots have the largest potential of cultural implications for an individual
- workshop invites to explore/test behaviors of social and physical distance considering proxemics, kinesics, cloud-computing, situatedness, intimacy, privacy, and audience topics of human-robot relations



# DAY 4: Friday, 21 August

- Design
- Perspectives
- Workshop V
- Plenary 6: John Danaher
- Outlook to RP 2022



*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

*Q2: How would ethically justify your design recommendations - i.e. what are the ethical pros and cons?*

# DESIGN

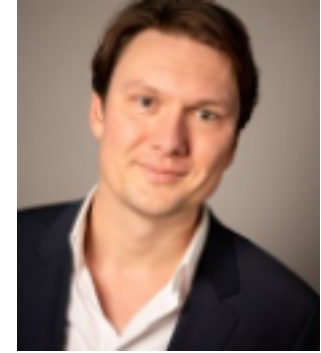
## How to Design Culturally Sustainable Social Robots

*Q3: Should robots have gender at all—should we not aim for designs that do not suggest race (ethnicity), age, gender, or other potentially discriminatory features? What kind of empirical results would be necessary to help us decide these questions?*

- Merle Wessel, Niklas Ellerich-Groppe, Mark Schweda: [Stereotyping of Social Robots in Eldercare](#)
- Malene Damholdt, Christina Vestergaard, Johanna Seibt: [Ascribing Gender to Social Robots](#)
- Oliver Bendel: [The Morality Menu Project](#)
- Claire Boine, Céline Castets-Renard, Aurélie Clodic, and Rachid Alami: [In Love with a Corporation without Knowing It](#)
- Mikkel Mørk, Sebastian Madsen, Christoffer Madsen: [Design for a Sustainable Robot?](#)

Merle Wessel, Niklas  
Ellerich-Groppe, Mark  
Schweda

## Stereotyping of Social Robots in Eldercare



### RELEVANCE OF SOCIAL ROBOTS IN ELDERCARE

- gender stereotypes play an important role & could be used strategically to optimize care
- gender roles among older people become more diverse and societal recognition of different needs and preferences grows  
→ stereotypical notions appear increasingly problematic

explorative ethical analysis of gender stereotyping in social robotics for eldercare

→ map potential problems and conflicts and discuss possible solutions for culturally sustainable social robots for eldercare in late-modern pluralistic societies

Malene Damholdt, Christina  
Vestergaard, Johanna Seibt  
Ascribing Gender to  
Social Robots



GENDER ASCRIPTION TO ROBOTS → REPEATING GENDER  
STEREOTYPES

- reducing this risk by delineating how gender is spontaneously assigned to robots

study (N=63 / interaction with the robot for 45-50 minutes  
- spontaneous ascription of gender)

- majority (n=36) ascribed gender to the robot, mainly based on voice
- remaining participants still assigned mental capacities to the robot

Oliver Bendel  
[The Morality Menu Project](#)



MACHINE MORALITY IS USUALLY FIXED

NEW: morality menu (MOME)

- users transfer their own morality onto the machine
  - machine acts in the same way as they would act
- case: implemented MOME for the MOBO chatbot
  - discussion of advantages & disadvantages
  - morality menu can be a valuable extension for certain moral machines

Claire Boine, Céline Castets-  
Renard, Aurélie Clodic, and  
Rachid Alami

In Love with a Corporation  
without Knowing It



MOST AREAS OF THE LAW RELY ON THE ASSUMPTION ON FREE WILL, WHICH  
MANIFESTS IN THE EXPRESSION OF CONSENT

- nature of human emotions toward social robots
  - question the concept of consent in the context of unreciprocated fictional relationships
- policies need to regulate the use of social robots in order to protect consumers
- especially vulnerable ones, asymmetry of power between them & robotic companies
- propose different statutory and design-based solutions depending on the purpose of the robots and the type of users

Mikkel Mørk, Sebastian Madsen,  
Christoffer Madsen

## Design for a Sustainable Robot?



“INTEGRATIVE SOCIAL ROBOTICS” TO SOLVE THE TRIPLE  
GRIDLOCK OF DESCRIPTION, EVALUATION AND REGULATION

concrete case: Silbot (Jeon et al. 2020)

- → how the 5 principles of ISR can mitigate the strained collaboration
- series of questions for research, design and development (RDD)



*Q1: What is the core message of your talk and how could your insights contribute to the development of culturally sustainable social robotics?*

# PERSPECTIVES Cultural Trajectories of Culturally Sustainable Social Robotics

- Michael Funk: [What Is Robot Ethics? ...And Can It Be Standardized?](#)
- Jordan Wales: [Empathy and Instrumentalization](#)
- Raya Jones: [On Human Freedom in a Posthuman Future](#)
- Gerrit Krueper: [Becoming Cyborg](#)
- Juho Rantala: [Anthropomorphism in Social Robotics: Simondon and the Human of Technology](#)

*Q2: Your contributions stand out in taking on a particularly wide scope—what are the concrete aims of your wide-scope contribution, who do you imagine will benefit from them?*

*Q3: Wide-scope cultural reflections and deconstructions of social narratives are the particular speciality of the Humanities—who can and who should be the audience for these observations and insights, and how do you envisage these observations to be translated into praxis?*



- language critical approach → 4 meanings of Robot Ethics
  - ways we practically talk about “Robot Ethics” / what we concretely mean by that
    - trademarks of the current English- and German speaking debate
- systematic approach
  - answer the question whether ethics can be standardized or not
    - a question that gained recent interest due to several international regulatory activities in the field of robotics and AI

Michael Funk

What Is Robot Ethics? ...And Can It Be Standardized?



Jordan Wales

## Empathy and Instrumentalization

### NEAR-FUTURE CONSUMER PRODUCTS WILL INCLUDE PERSUASIVE SOCIAL ROBOTS

- will attract us by their apparent subjectivity but that
- will invite us to instrumentalize that subjectivity as a tool of our desires

*How are we to “use” them without growing comfortable with slaveholding?*

- early fifth-century North African philosopher Augustine of Hippo
  - extend his reflections on human relationships with persons and artefacts to suggest how our empathy toward apparently personal possessions might be exercised in a manner that can upbuild rather than erode our capacity for authentic interpersonal intimacy



## CHALLENGES FOR CONCEPTUALISING 'CULTURALLY SUSTAINABLE SOCIAL ROBOTICS'

- approaching the relevant discourse from a critical distance  
discussion
- identifies dissonances among different positions and their associated moral claims
- points to some ambivalence inherent in the phrasing of the goal

Raya Jones

On Human Freedom in a Posthuman Future



Gerrit Krueper  
Becoming Cyborg

## THE MATERIAL IS ESSENTIAL TO THE ONTOLOGICAL QUESTION OF BEING HUMAN

cognitive capitalism & its rise of technology

- translates human body into literal instruments of labor  
→ transfer of skills & powers extending the body's capabilities, creating a cyber-body
  - ambivalence of the material reality of the cyber-body:
    - reality of exploitation & abstraction
      - designed to eventually create infinite capital accumulation
    - reality of liberation potentials
      - through re-appropriating one's cyborg body to connect to nature and the social
- this ambivalence recovers the real species-being



## ANTHROPOMORPHISM IS A COMPLEX PHENOMENON

- thought to be desirable in social robotics to enhance their functionality & sociality
- strict anthropomorphism can limit possible capabilities of robots

Gilbert Simondon's analysis: technology = inherently human  
→ theoretical description of anthropomorphism in social robotics

- foundation for more robust conception & understanding of advantages and disadvantages of social robots as well as significance of anthropomorphism for social robotics

Juho Rantala

Anthropomorphism in Social Robotics: Simondon and the Human of Technology



# Robots in religious contexts

Simon Balle, Aarhus University (DK)

Charles M. Ess, University of Oslo (NO)



In the wake of the robot revolution, social robots will eventually find their way into religious contexts.

Indeed, some have already done so. Recently 'Mindar', the android version of the Buddhist deity Kannon Bodhisattva, has been introduced in a Buddhist Temple in Kyoto; a humanoid designed as Ibn Sina has probed Islamic attitudes to robots in the UAE; and Catholic and Protestant contexts have seen such inventions as SanTO and Bless-U2 respectively.

As roboticists start to produce 'theomorphic robots' to represent and mediate the divine, there is an urgent need to include scholars of diverse religious traditions in the debate.

# Pioneering Religion in Robotics: A Historical Perspective



WORKSHOP 5

RELIGION ALSO HAS EVOLVED THANKS TO NEW TECHNOLOGIES, AND ROBOTS

- since ancient times robots are related to the sacred
- new creation of robots for the religious domain have to face the attitude of users of different cultures
  - trace back to the mindset present in the background religions, and their concept of human

West	Israel	Middle East	India	East Asia	Japan
Greek heritage (+) Robots in mythology (+) Automata	Monotheism (-) Dualistic view of soul (-) Body is a gift from God (-) Rejection of magic		Vedic heritage (+) Tales of war machines	Taoism (+/-) Harmony between man and nature (+) Goal of reaching immortality	
	Human-God distance (-) Golem		Hinduism (+) Widespread use of Murti	Confucianism (+) Animism (-) Little emphasis on science	
Christianity (-) Human mediation in faith (+) Proactive approach in science (+) Automata and actuated statues (-) Ora et labora	Judaism (+) Work as a punishment (+) Use of automation for the Sabbath	Islam (-) Conformity to religious law (-) Need of a moral for human dignity	(+) Use of puppets (+) Inclusive attitude towards other gods (+) Delegation of ritual	Buddhism (+/-) Use of shadow puppets	Zen Buddhism (+) Cohesion between body and mind
Anthropocentrism in philosophy (-) Dignity of man (+) Laws of robotics		Arabian science (+) Interest in practical applications	Multiculturalism (+) Inclusive attitude	Specificity of Chinese science (+/-) Moral regulation of innovations	Shinto (+) Sentient beings equal to humans (+) Concept of kami
Industrial Revolution (+/-) Shift of jobs					12th C. scrolls (+) Anthropomorphism of animals
					Collectivism (-) Work sharing ethics
World Wars (-) Self-destructive power				Pragmatism in modernisation (+) Labour innovation	Anime (+) Robots as heroes



Gabriele  
Trovato,  
Waseda  
University (JP)

# Does AI Have Buddha-Nature?

Reflections on the Metaphysical, Soteriological, and Ethical Dimensions of Including Humanoid Robots in Religious Rituals from One Mahāyāna Buddhist Perspective



WORKSHOP 5

TEMPLE KŌDAIJ MADE HISTORY WHEN ITS PRIEST ENSHRINED THE ROBOT MINDAR AS A PERSONIFICATION OF KANNON BODHISATTVA

Mahāyāna Buddhist:

- reject any form of dualism between the divine and the secular
- claim that “insentient beings are buddha-nature” & “insentient beings become buddhas”

*Can we think of humanoid robots as conscious?*

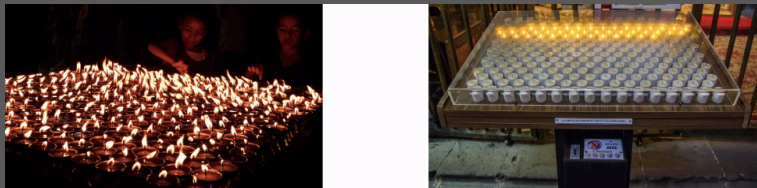
- deserving person rights, count as divine in a Buddhist context

*What are the practical & ethical implications of the possible Buddhist claim “all humanoids have Buddha-nature”?*



**Gereon Kopf**  
Luther College  
(US)

# Robots' Spiritual Superpowers



WORKSHOP 5

ROBOTS ARE PERCEIVED AS HYBRIDS  
BETWEEN 'THINGS' AND 'BEINGS' – AS  
TECHNOLOGICALLY 'OTHER'

- due to potential proactivity, reduced predictability, lack of transparency, appearance
- approach of robot 'superpowers'
  - social capabilities that are rooted in their 'thingness' rather than 'beingness'
- endless patience, non-judgmental, customizable
  - qualities difficult to realize by humans
  - new forms of social interaction in religious contexts

robots could be designed as novel and complementary, instead of substituting, 'species'

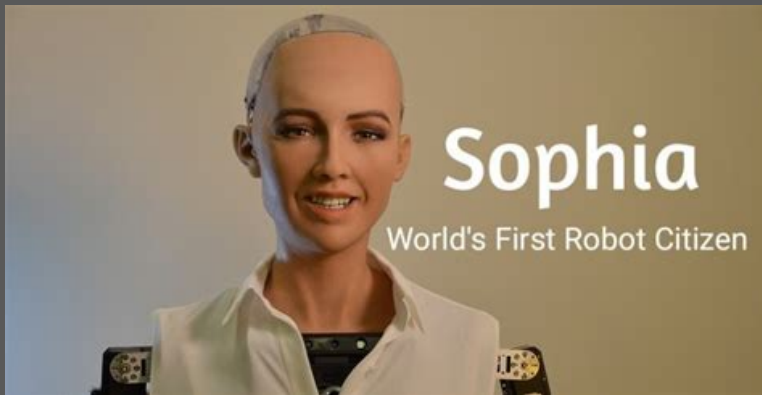


Diana Löffler &  
Marc Hassenzahl  
University of Siegen



# Concubine, slave, or wife?

## Religious reasons in the Debate About Marriage to Sophia the Robot



WORKSHOP 5

SOPHIA THE FIRST ROBOT CITIZEN IS AN INTERESTING STARTING POINT FOR PUBLIC DELIBERATION ON THE INCLUSION OF SOCIAL ROBOTS

question on Egyptian TV show (dedicated to taking Sunni Islamic advice on matters of life):

- *“Would Sophia count as one of four marriages or it would not count as marriage?”*
  - interrogate this episode of the show
- How religious reasons have been reflected upon elsewhere in public deliberation about social robots’ position in the Middle East?



**Nardine  
Alnemr**  
University of  
Canberra (AU)

# On the Magical Dimension of Religion. Theological Questions on Robots in Religious Contexts

WORKSHOP 5

## ROBOT PRIEST “BLESSU2”

→ important questions on key themes of religion(s) in digital societies:

*Are robots legitimized and authorized to pronounce blessings on humans – and why?*

- define the interrelationship of technology, religion & the human being
- Paul Tillich (1886-1965)
  - referred to the polarization of autonomy and heteronomy by raising the issue of theonomy
  - critical research on representing the divine in robotic technology



**Ilona Nord**

Würzburg University (DE)

**Thomas Schlag**

University of Zurich (CH)



# Between Luther and Buddhism: Scandinavian Creation Theology and Robo- philosophy

WORKSHOP 5

ROBOPHILOSOPHY USING ROBOTIC TECHNOLOGIES  
AS TESTBEDS FOR PHILOSOPHICAL QUESTIONS E.G.  
NATURE OF BEING HUMAN

- Robotheology: questions of the mind-body, Creator-creation, and faith-reason relationships  
agenda for non-dualistic approaches
- Scandinavian Creation Theology (SCT):
  - more optimistic conception of human nature
  - non-dualistic account resonating with Eastern approaches
  - fruitful for central issues such as distributed ethical agency and responsibility, love, sex, and trust



**Charles M. Ess**  
University of Oslo  
(NO)



# ROBOTS AND MORAL REVOLUTION

HUMAN SOCIETIES HAVE, HISTORICALLY, UNDERGONE A NUMBER OF MORAL REVOLUTIONS

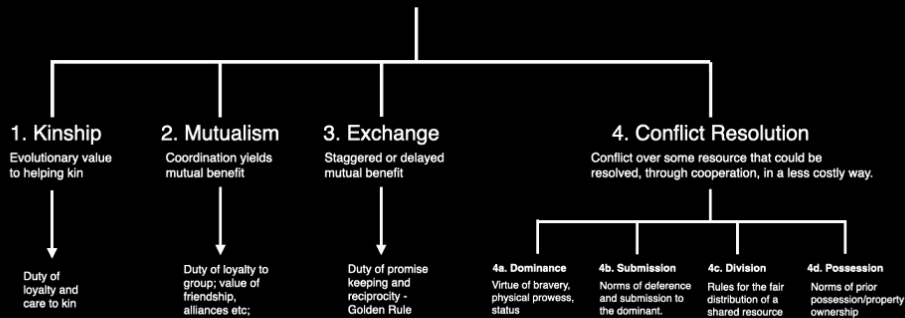
- some of these have been precipitated by technological changes  
*Will the integration of robots into our social lives precipitate a new moral revolution?*
- role of techno-social change in facilitating those revolutions
- structural properties of human moral systems and how those properties might be affected by social robots

**much of current social morality is agency-centric**

→ **social robots, as non-standard agents, will disrupt that model**

## Cooperative Problems

Any interaction in which cooperation would yield some non-zero sum benefit but there is an impediment to or barrier to cooperation



**Associated moral norms/values**

