



# Collaborative learning for ethical robot design

**Karolina Zawieska**  
De Montfort University

**15 February 2018, Vienna**  
Robophilosophy Conference



"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731726"

# REELER: Ethnographic fieldwork in Roboethics

## Robots

- Different application areas
- Human-Proximity Levels
- Different European countries

## Participants

- Roboticists/Robot Makers
- Affected Stakeholders

Case: Educational social robot designed to teach programming and logical thinking in the early school education



# Roboethics: Learning challenge

- > Affected Stakeholders: Roboethics without robots
- > Roboticists: Roboethics without ethics
  - Novelty, for both robot designers and for the society
  - Misconceptions on both sides
  - Only limited encounters between the two

Ethics is difficult to grasp:

Not just as **reflection** upon particular matter but also  
'a particular way of **being** in relationships, in the world'  
(A. Peterson, 2001)



Question:

Roboethics is a new discipline that requires  
new learning approaches

<input type="radio"/> Strongly disagree	<input type="radio"/> Disagree	<input type="radio"/> Neutral	<input type="radio"/> Agree	<input type="radio"/> Strongly agree
---	-----------------------------------	----------------------------------	--------------------------------	---

# Roboticists' Perspective



# Dismissal of ethical concerns

- Ethics viewed as a separate field (J.P. Sullins, 2015)  
(especially if asked explicitly)

*‘Well, that [defining ethics] is problematic. With regards to technical issues, It’s quite easy’*

- Engineering perspective: short-term, simplifying and task-oriented

*‘It’s hard for me to imagine that a device that is supposed to replace a computer game and teach something new to children, to help them understand the things that your child sooner or later should learn, that is will somehow affect the social change.’*

*‘I think every human being always keeps in minds that when we do something it should not to have a negative impact on others’*



# Dismissal of ethical concerns

- Engineering mind-set: different view of what is ethical or not

*'Initially, we wanted to try to bring a computer game to the real world, so that instead of playing on a computer screen, it could be done on the floor. After a few modifications we came to the conclusion that we should try to teach something to children and we came up with the idea of an educational robot with a strong focus on entertainment. Our first concept was to create the robots that would fight with each other. The children would program robots so they [robots] can fight. But parents opposed.'*
- Attempts to develop ethical robot design have no guarantee of success

*'And now, our generation has been trying to come up with the assumptions for what will be wrong for these kids who have been already part of it, right? It can be perfectly normal for them'*



# Tacit assumptions and practices

- In principle: Robots are designed to bring benefit rather than harm people  
*‘Actually the robot constructor who starts building the robot knows what it [robot] is supposed to serve. And if he knows that it is supposed to serve unethical purposes, then ethics is not important to him’*
- Intuitive understanding of social and ethical issues and informal ways to verify it (one own’s kids, competitions, workshops etc.)  
*‘it is worth meeting extremely different people who will ask questions we would not even think of’*





# Distinction between 'Us' and 'Them'

- Limited roboticists' responsibility

*'we try to anticipate a lot of risks [e.g. limit duration of use of robots]... As a producer, we are sure that our product is safe, so if it creates any danger it means that in 90% it is because of misuse'*

- Users' feedback (& collaboration with social scientists) to improve a product

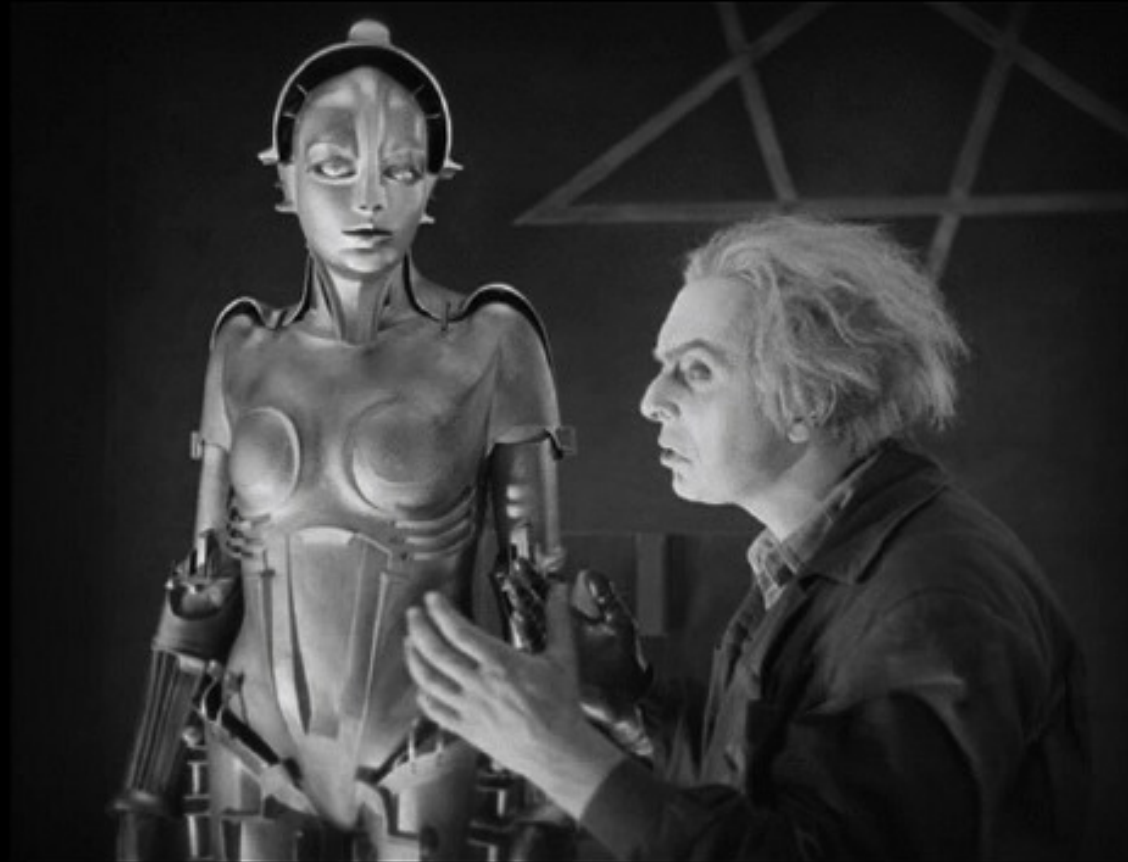
*'in addition, we have carried out so-called pilots [with participation of kids]. We are a company, so there is actually a commercial aim for us in all of that. And our mission is to sell as many robots as possible'*

- Freedom of choice

*'we do not know what effect it [robotics] will have and whether in 10 years, when you will be asking them how they feel about it, they will tell you that they would not want the human being to replace their relationship with the robot'*



# Learning



# Need for Collaborative Learning

‘a **situation** in which **two or more people** learn or attempt to **learn** something together’ (P. Dillenbourg, 1999)

REELER goals:

- Close the gap between Roboticists and Affected Stakeholders
- Facilitate ethical robot design

Roboticists’ goal: Why learn about ethics? Need for practical applications

\* Collaborative robots: Robots that are able to share space with human workers and execute tasks in parallel



# What learning situation?

Design process

- Formal/informal learning contexts
- Temporary/permanent collaborations
- With and without end-users and HRI
- Laboratory/real life settings
- If learning by doing, including by mistakes – what limits?  
(best & worst practices)

*‘at the initial stage of the robot development, we used to explain to children what we wanted and how we wanted it. Later we decided just give it [robots] to kids and see what was going to happen’*



# Which Learners?

**Robot makers:** Persons who contribute to the process of creation of robots

- The difference between Roboticists & Affected Stakeholders is sometimes blurred. For example, do policy makers co-create robots?
- The impact of material and socio-cultural contexts – more than individual/collective learners (e.g. teaching programming at school to educate future programmers)
- Learning together with whom?  
(other roboticists, social scientists, end-users, policy-makers etc.)



# What learning subject?

What is specific about robotics and what is not?

What are the constraints for roboticists and their creations?

What degree of new knowledge?

‘sensitivity’ to the ethical implications of work  
(J.P. Sullins, 2015)

- Ethics and other disciplines (typically psychology)
- Existing and/or new knowledge & practices (robot as a tool or a companion)
- People’s preferences, concerns and needs (conflicts?)
- Socio-cultural and individual differences
- Communication

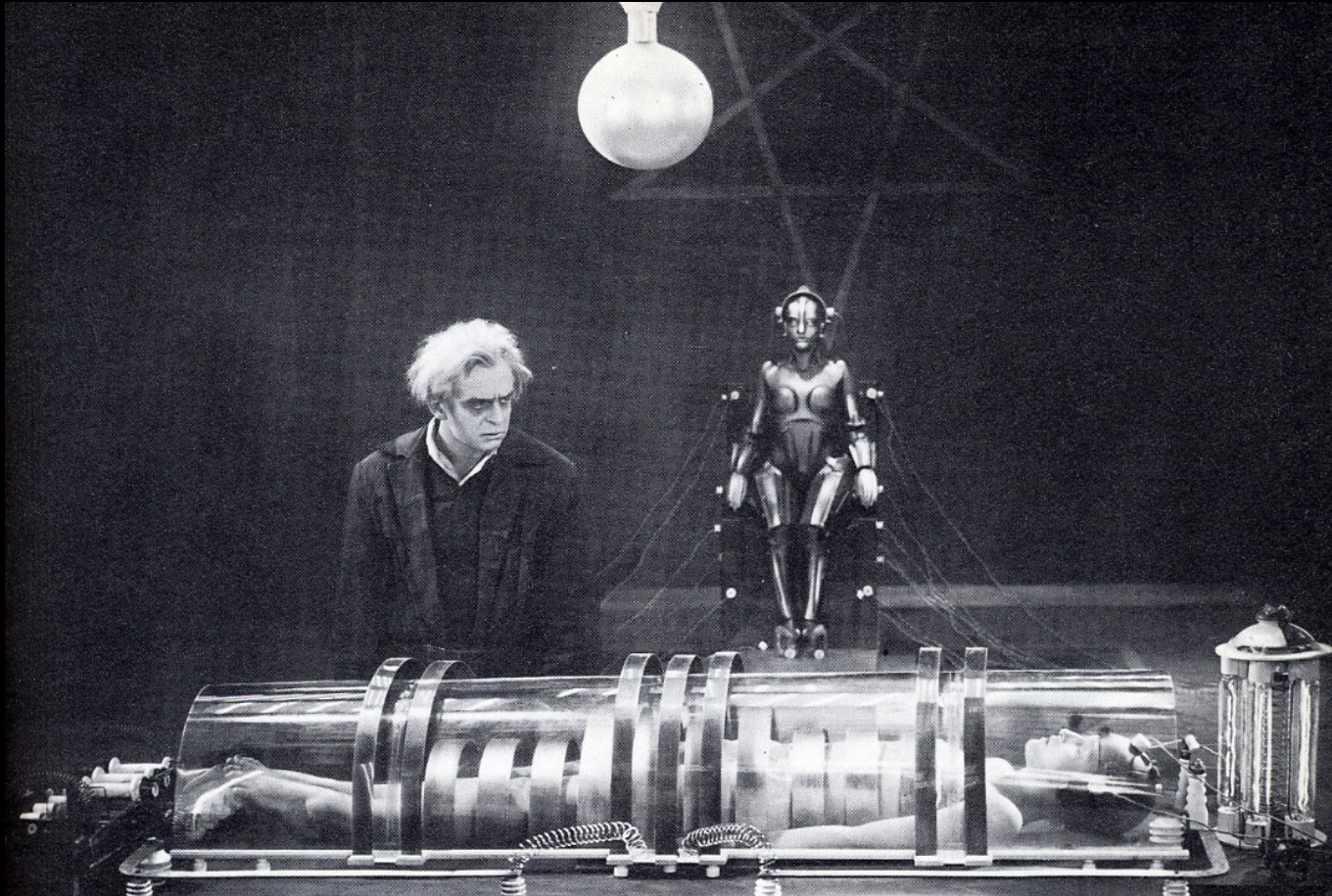


# Together – how?

- Shared goals -> Case-specific
- Shared understanding -> Customised
- Shared language
- Shared practices
- Shared time perspective
- ...



# Conclusions





# Collaborative Learning For Ethical Robot Design

## Attitude

- Design approaches not only about robots but also about human beings
- Ethical challenges as puzzles to solve
- Long-term perspective

## Motivation

- Why roboethics at all?  
Implicit and explicit motives
- Added value – to whom?
- Clarity of goals for robots  
as a must



Question:

Roboethics is a new discipline that requires  
new learning approaches

<input type="radio"/> Strongly disagree	<input type="radio"/> Disagree	<input type="radio"/> Neutral	<input type="radio"/> Agree	<input type="radio"/> Strongly agree
---	-----------------------------------	----------------------------------	--------------------------------	---