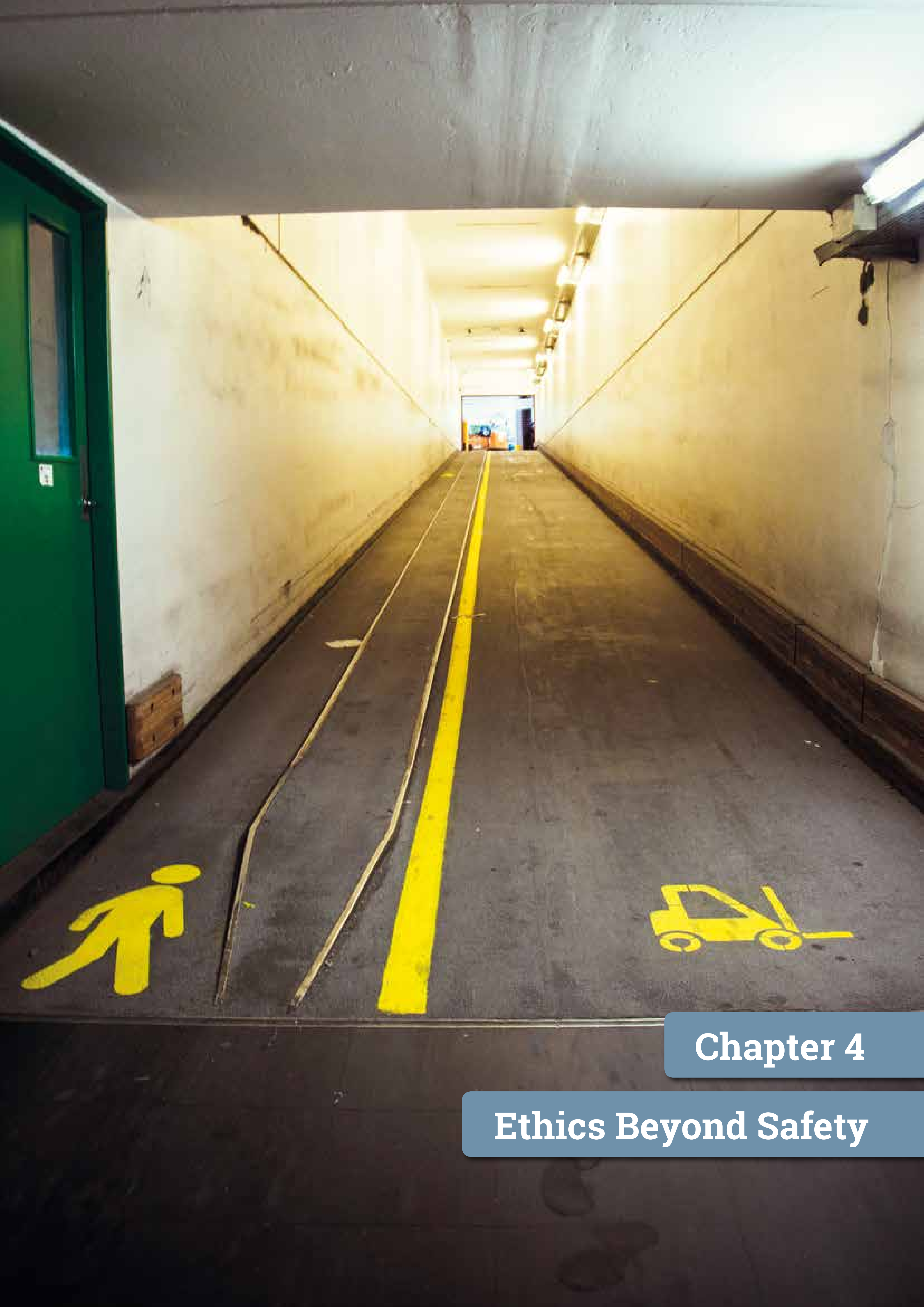


Chapter 4

Ethics Beyond Safety

If you think for example of elder care, I mean, we don't want the elderly people to only have contact with robots. Is that a good thing? Or is that a bad thing?

(Oswaldo, industrial designer, robot developer, SPECTRUS)

◀ Some robot makers have a tendency to separate humans from machines when thinking about ethics. (Photo by Kate Davis)

4. Ethics Beyond Safety

You will find here

- Empirical examples of how the affected stakeholders and robot makers define ethics
- Examples of ethical challenges that may emerge in the processes of robot design, development, and implementation
- Overview of the main academic ethical frameworks in robotics
- A discussion of how REELER's approach to ethics differs from other ethical guidelines

You will acquire

- Awareness of how to expand safety-oriented approaches towards a holistic, socially distributed ethics
- Awareness of the need for a socially distributed ethics, perspective taking, collaborative learning, and relational agency
- Awareness of how to identify wider benefits and problems related to ethics in robotics, rather than just focusing on safety

In the REELER project, a key emphasis has been on ethics as an umbrella term covering a variety of aspects of responsible robotics and ethical robot design. Recent years have seen robots 'uncaged' from their secure environments to increasingly affect everyday lives of humans. This has ethical implications – as noted by this robot developer:

> *Whenever the robot interfaces [with] the human being, there are ethical issues coming out.*
>
> (Arturo, engineer, robot developer, REGAIN)

For that reason, ethics plays a central role in the REELER Roadmap and this publication. All of the subsequent chapters in the publication deal with issues that are relevant for the design of more responsible and ethical robots. Many of these issues overlap only to some extent with existing guidelines found and collected by members of the *Ethics in Action* workgroup, a global initiative by IEEE (Institute of Electrical and Electronics Engineers), the world's largest, international organization for engineers and technical developments.[1]

In this chapter, we take a closer look at the *concept* of ethics. What do robot makers mean by ethics, if they refer to ethics at all? REELER has found that the term is not part of daily concerns in most engineering practices (*see also Sorenson 2018 and Hansen 2018*). Many other robot makers, especially educators in the engineering sciences, funding agencies, policymakers and philosophers, think lot about what is covered by the term ethics. Engineers are often practitioners, practical people working on solutions to specific problems, and from this perspective robot developers tend to connect ethics with one main area in their everyday work lives: safety. This is a pattern across the REELER cases. When asked to reflect on ethics in our interviews, the developers and other robot makers mostly consider ethics to be about safety and avoiding small or major catastrophes. Many also consider how their products may help people and do good for mankind, but these considerations are not always connected to reflections on ethics. Despite the empha-

1 IEEE Ethically Aligned Design (EAD) workspace (https://ieee-sa.imeetcentral.com/ead)

Ethics in robot design: *Personal and collective awareness of ethical issues as well as the ability to actively engage with both ethical reflection and practices with the goal to pursue value-sensitive design and responsible research and development in robotics. The key premise is the orientation towards 'others', which includes the practice of taking other people's perspectives and understanding their motives.*

sis on ethics as safety and the notion of doing good, some robot developers do also regard humans a nuisance. They seem to have a hard time understanding end-users and (as pointed to in the Introduction) overlook the wider circle of affected stakeholders. REELER also finds a number of reasons why it may be difficult for robot developers to make ethical and responsible robots; one is that robot parts are scattered in time and space in the various design phases. Furthermore, existing ethical guidelines are often made by other robot makers within the 'inner circle', as defined in the introduction. Consequently, these guidelines unfortunately say little about the real-life problems the affected stakeholders can help to identify.

As discussed throughout the publication, the main problems with ethics in engineering is the gap between robot makers and affected stakeholders. Thus, we end this chapter by calling for a more relational responsibility and discuss how ethics in the approach REELER proposes differs from existing ethical guidelines.

4.1 Robot makers' views on ethics

Some robot developers have knowledge about ethics as a philosophical topic, but in our data material those mainly concerned with ethics in a philosophical sense are people educated to do so. Developers will, as Gunnar in the opening quotation, ponder about ethical issues when asked to do so, but in their everyday work, ethics is not their first concern. And it is often considered to be dealt with by 'ethical people' hired to deal with these matters. Thus, ethics as a general concept is generally severed off from the daily work in robot development, which is mainly technical in nature.

"When you are trying to solve this problem, you're not thinking about the ethics – you're just trying to figure [out] a solution to something technical. How do I get the robot to move this part from here, to here? "

(Robotics engineer quoted in Sorenson 2018, 18)

During REELER interviews, we have asked our interviewees for their associations with the word ethics.[2] This is done by asking them to mention up to five words they associate with ethics. They often find it is a hard question, but try to find answers that tend to move beyond their work on robots.

Interviewer: "If you were to define ethics, what is ethics? The association again, like with a robot, now with ethics."

Monika: "It's Hard. Ethics. It's like a behavior. Generally speaking, behavior that does not affect another person in the wrong [bad] way. Hard to define. Well, in my opinion, ethics is a behavior that does not have a negative effect on another person or on another being [entity]."

(Monika, scenario developer at robotics start-up, robot maker, ATOM)

Though some of the robot makers (especially the facilitators who make policies) are very aware of the concept of ethics and have given it much thought, the robot developers are often in doubt what is meant by the word – something we deal with in the section on isolated ethics.

Interviewer: "If I say ethics, how would you define ethics? From your own perspective."

Ernesto: (Laughs) "Too philosophical."

(Alph, robotics start-up founder & CEO, robot developer, WAREHOUSE)

This does not mean robot developers do not care about ethical issues, only that in their daily work these issues are considered practical and technical problems – most often tied to safety issues. As noted by REELER researcher Jessica Sorenson, this points to a "discrepancy between the way the engineers approach ethics, as a theoretical moral orientation, and the way they approach design, as a practical problem-solving activity" (Sorensen 2018, 18).

2 See Annex 1 Methods and Methodology (responsiblerobotics.eu/annex-1) for a copy of our interview guide, and Annex 2 (responsiblerobotics.eu/annex-2) for more interview quotations dealing with ethics. These annexes can be found in the digital version of Perspectives on Robots: responsiblerobotics.eu/research/perspectives-on-robots).

However, as we shall see, we have also talked to robot developers and robot makers who think beyond theoretical ethics and have a more holistic and societal view on how their robotic technologies should be ethical.

In brief, REELER identifies two general ways of relating to ethics among the robot developers:

A. Problem-solving views on ethics:
Robot developers primarily consider ethics a matter of safety. When they take care of safety issues, or find their robots do good, they believe their robots are ethical, and there is not much more to be said about ethics. Some robot developers view ethics as a 'problematic' term which is outside of their problem-solving realm. Ethics is not to be foreseen, but can appear if something goes terribly wrong. Though it can be considered unethical to cause replacement of people, or violate privacy, or have machines making wrong decisions, many robot developers do not see this as their ethical responsibility because the technology is considered to be neutral. Humans, however, may be seen as a nuisance and a problem in themselves as they slow down efficiency in robots, especially in workspaces.

B. Holistic views on ethics:
This area covers conceptions of humans in Human Robot Interaction (HRI) relations. Some robot developers take an active stance against delivering robots for military purposes, for example. Other robot developers and robot makers with a holistic view on ethics see affected stakeholders as a rich source of ideas for better design, and they express curiosity about what matter in the lives of humans.

4.1.1 Ethics as safety and problem solving

In this section we deal with the predominant association among the robot developers when asked to reflect on 'ethics'; that ethics mainly has to do with safety problems. Here, being ethical is to focus, from a technical point of view, on what can go wrong in a robot project with respect to the materiality of the robot and the presence of humans. This finding is no surprise, as general discussions about ethics in robotics have typically been associated with safety. REELER's data strongly confirms this aspect; safety is an inherently technical and system-oriented concept, often connected to regulations, rules, and standards.

Human safety: *Bodily and mental integrity of people.*

The majority of the interviewed robot developers do not see it as their task to extend ethical thinking beyond how their robots work mechanically. Here are some examples of how robot developers see responsible robotics as a matter of making safe robots.[3]

I think that it [ethics of the robot] stops at the safety level. Again, we are back at the risk analysis. Because the robot has to be designed and developed to be safe, to avoid, to make [problems]. In this case, we have function of electrical stimulation. (...) And one of the problems is that they have to affix electrodes in some kind of garment, because a person with a pacemaker may decide to put the electrodes on his chest. So, we have some bigger problem. So, you know, safety has to be within the robot. Safety of the managing of the machine.

(Albinus, robot developer, REGAIN)

It is not only the robot developers (i.e., technical engineers) who mainly connect ethics with safety. The robot makers in general also make this connection. A pattern across cases is that the first association when robot makers are asked about ethics is safety tied to engineering norms, regulations, and standards. Below are examples from robot makers (a policy-maker and a business manager), who also see safety as a key ethical issue.

Interviewer: "Human workers work with robots or this type of machines. Do you think there will be any issues, ethical or other issues related to working with robots?"

Yves: "Yeah, yeah, yeah. There is a safety issue because so far autonomously acting machines were kept in cages."

(Yves, policy advisor, robot maker, COOP)

Interviewer: "We are talking about [the robot], could you think of any type of risk, of ethical problems related to the development?"

Simone: "Not necessarily. It is clear that if we are going to take a measure that concerns the safety of passengers, ethically and responsibly we need to always carry out the highest quality standards."

(Simone, sales manager at a robotics company, robot maker, OTTO)

3 More quotations can be found in the online *Annex 2* at *responsiblerobotics.eu/annex-2*

Problem-solving thinking is sometimes applied to ethics.

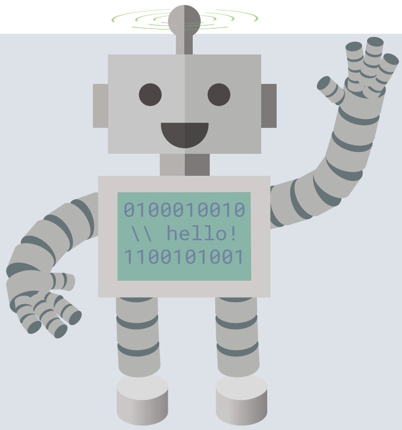
In the above preceding quotation, the robot maker sees ethics as connected to the safety of the passengers and less to as issues tied to the robot in itself, as it is not going near the passengers. However, it *is* going to be operated by the transport personnel who will be in close proximity to the robot, but Simone does not point to ethical issues in that regard: i.e. the human workers operating the robot are not in focus here.

In more recent examples of advancements in robotics and its new related fields such as Human-Robot Interaction, safety, continues to be a key focus. Even when addressing safety from new perspectives, for example how people have to work in direct proximity of robots rather than in separate environments (Bicchi et al. 2008), safety remains a dominant concern in robotics, but also one of the main advantages of implementing robotics technologies. This is not just tied to how robots are safe machines in themselves, but also pertains to arguments for why we should choose robotic solutions in the first place: because they are more safe. From an ethical point of view, this is one of the reasons some robot developers are convinced they 'do good', because their robots will be safer and more dependable than the humans. For example, some of the start-ups in that are designing self-driving cars, point to security and safe driving as one of the main selling points (in the HERBIE case). The cars are controlled by AI software that allows detecting objects and avoiding collisions while following a predetermined path. In this way, self-driving vehicles will have fewer accidents than when driven by humans. In the aerospace industry, the introduction of increasingly sophisticated automated systems in airplanes is also often justified by the need to improve safety (Mindell 2015). A similar logic guides the design, implementation and use of robots in many of the REELER cases; society demands increased efficiency without compromising safety. Here affected stakeholders often share motives with the robot developers:

> *Passengers' safety was there before and is there now. However, like all things: the speed increases, the number of trains increases, and this is a next step. But security is the foundation. So, in any case, either done with traditional tools or done with a robot, it must be safe.*
>
> (Kian, operator at the metro company, affected stakeholder, OTTO)

No matter what robot type we have looked into, whether the robots were argued to increase safety of a work task or not, safety of the actual machine developed is also an issue. This approach has often been reflected in the engineering codes of ethics that emphasizes safety in the first place (e.g. in the IEEE prescriptions – see below sections on regulations). It has been argued that in robotics, the safety issue is generally within robot software and design (Lin 2011). Given a long tradition of industrial robotics, safety has typically been associated with the system design and properties and its impact on the physical surroundings. Industrial robot applications have also been subject to numerous safety regulations and measurements that have been successfully put into practice.

STORY FROM THE FIELD:
Protect the fingers



A Northern robot company, which we here name Cobotics, has been approached by a large manufacturer which makes parts for cars and tractors. They want a robot that can work together with their human employees; they would like their workers to increase the work pace because the production has increased. The problem is the workers rivet metal plates and they cannot speed up without the work becoming unsafe. When REELER visits the robot company they are developing a new type of robot, the NITTER. For safety reasons, the robot was at first supposed to replace the human workers. But because they have run into a number of problems, they are now considering turning NITTER into a cobot that will work together with the humans. The robot could not adjust to the slightly diverse riveting tasks, so the humans still have to place and adjust the materials and remove them after the riveting.

If there were no human workers around, safety would simply be about fencing the robot and making procedures for when and where humans can enter the fence. In normal factory riveting, a robot is an autonomous tool and humans are not involved in the process. But this time, the robot is put to a collaborative task. Michael, one of the researchers in the COBOT case, explains to us that with the collaborative robot NITTER, they *"want the worker to put in the parts, it's like two sheet metal parts, and the robotic machine to rivet them together, because this riveting is all over the plates, and you need to rivet the two parts together. The riveting is the monotonous parts, but because we have a lot of work to rivet, there is just more work in getting it in and off, and making inspection from the shop floor. So, this is like a typical task. We say that if we get 2 to 3 of these robotic machines, one worker can cope with that [work of getting the metal parts in and off, and making inspection] and the worker doesn't have to lift that heavy riveting machine anymore, which is an ergonomic help, also."*

Michael also explains that the worker can qualify the process, because *"He can take a look if the NITTER is running out of rivets, something like that"*. Michael shows us that the safety issue of this collaboration lies in the fact that *"the worker can put his finger here [Michael points at a clamp], somehow, and cut his finger off. Which we solved with a pneumatic solution, so we just take care that this is closing safely and then we put the pressure on it. Because we need a certain pressure to make sure that the bolts align. Otherwise we don't get the holes riveted correctly. And next thing [concerning safety] is that the rivets have got a pointy head. And this could hurt. So, we also make sure there's no chance to interact with that tool tip. (…) We put our fingers inside and say it hurts a lot. [Laughter]. Ronald did. I didn't. I was afraid to do it. [Laughter]. I just did it after he said it was safe [laughs]."*

(Based on an interview with Michael, researcher, robot developer, COBOT)

One could argue that giving great importance to safety is a good starting point for incorporating other aspects of ethical reflection and conduct needed to ensure both safe and ethical robot design. But our analyses show that the concept of safety is viewed mainly as a technical challenge and in terms of the system performance.

In the subsequent sections, we point to other ethical issues, some of which are connected to safety while others move beyond safety. The latter concern seeing humans as problems, machines as decision makers, and machines as hackable.

4.1.2 Ethics of human machine interaction

As mentioned, ethics as safety is very prevalent when robot developers work on cobots, like NITTER. Such collaborative robots are part of frequent and close physical as well as social interaction with and among human beings, which requires even stricter safety measures put on the machinery.

> *So, from the moment the thing is a robot, it gets a whole different object. So, if I put myself in a dentist's chair today, and he's driving around, he has six, seven axles, I think he has no emergency stop, no safe positioning or anything else. If you have such a little robot that moves a bit, then there's an incredible safety story around it.*
>
> (Kai, mechanical engineer and cluster leader, robot maker, COBOT)

This may lead to frustration as some robot makers find safety requirements overly restrictive concerning the machines – instead of demanding humans behave differently. Humans should, for instance, learn to adjust their movements around the robots to follow their speed.

> *If I now move my head or get up from below and move my head up, I'm much, much faster, than the robot could ever move. Now I brake, the robot, moves only very slowly, but what we do not consider currently is that, yes, humans are able to move very fast. And what would I do as a worker if I hit my head on the robot? Would I say, "Yes, I stood there"? "He hit me!" So, how do I prove that I [the robot] did not hit him now, but he ran to me [the robot]?*
>
> (Kai, mechanical engineer and cluster leader, robot maker, COBOT)

Because of restrictive safety measures, some robot developers like Kai, see ethics as something that prevent them from making their systems work and deliver what they have promised. It could be a problem that humans move faster than the robot (something we also see in the case of construction robots) but it can also be the opposite, that humans slow things down. For example, whenever the operator appears in close proximity to the robot, the robot needs to slow down, and hence *"it will never finish the work"*' (Emilia, director of research and innovation, robot maker, COOP).

Thus, in the concern for safety, humans are both an object of safety and a problem; a problem because they may be more difficult to control than the robot which may prevent the most efficient solutions, as seen from an engineering perspective. In these situations, the robot developers' ethical considerations appear to center more around negative occurrences and wrongdoing when humans are involved than principles and shared values with affected stakeholders, which could guide ethical thinking and conduct.

Across cases, we find a view on human robot interaction that suggests a human-robot dichotomy with ethical implications in the field. This is, for instance, expressed through the notion that introducing robots to different sectors often involves choosing either humans or robots, rather than combining the two.

> *If you want to have result, if you want to have performance from the robot, you have to separate people from robots.*
>
> (Alph, robotics start-up founder & CEO, robot developer, WAREHOUSE)

One of the major concerns is the situation where one literally makes a choice between humans and robots in the work context. Robot developers are sometimes aware and concerned of how their robots in general affect the work force (because their robots are considered more efficient than human workers). This concern is prevalent across all cases and groups involved, including policymakers and affected stakeholders. However, from the problem-based point of view taken by most robot developers, the objective is to make the most efficient and safe robots as possible. The ethics concerning e.g. humans put out of work, is beyond this task. From the problem-solving point of view, some of the robot developers see ethics as tied to humans – not to their technology.

> *Interviewer: "Can you think of any ethical challenges related to the use of robots in warehouses?"*
>
> *Danny: "Ethical challenges? What do you mean – you mean in terms of people? I don't know. I guess that's more a difficult one for me in the sense that the nature of what I do is to sell automation solutions to customers. So, my ethics, are really borne around helping organizations to improve, being more competitive, and to allow them to have an environment that stimulates growth and opportunity. That's what I'm focused on rather than perhaps the people aspects of robotics."*
>
> (Danny, sales manager, affected stakeholder, WAREHOUSE)

Ethics of humans is, in other words, separate from ethics of the machines.

> *Did you say ethics? Ethics, that is something about, yes, eh I'm thinking more about humans, in any case. What is ethically correct and that is mostly what I think of. It might not make me think of robots.*
>
> (Werner, operation and production technologist, robot developer, WIPER)

This separation of ethics into ethics of humans and ethics of machines is not holistic, but problem-oriented in the sense that the problem is how to increase productivity and efficiency.

> *The need of higher productivity is a reality for different sectors. So, this increase of productivity and the cost of the human operator is higher, higher in particular in Europe. So, there is not the choice of the robot versus the operator: It's no work in Europe versus having the work in Europe. Or when working with the robot and an operator, not to increase the number of operators. I think it's not an [option]. We have to be able to understand this. So, the option is not to [keep humans in the loop] and not lose all the jobs, because, otherwise, in Europe, we will just not have no production.*
>
> (Emilia, director of research and innovation, robot maker, COOP)

Though some robot developers tell REELER researchers their aim is not to earn money, others emphasize that, at the bottom line, earning money in companies is what counts. Therefore, robots should not be oversold as better solutions for the employees, if they in fact are mainly beneficiary for the company owners.

> *And generally, it is ethically questionable [to place robots in all kinds of human environments], I think. I did many industrial applications before. We say that the robot supports the human but actually it revolves around a ROI [Return on Investment]. There we should not delude ourselves. Especially the automotive manufacturers. They use and calculate the return of investment at the moment. That means that the robot will definitely replace humans. And it is not the way that the robot supports humans, most rarely. The COBOT would like to do that but in the end the robot will be bought if it pays back for the client.*
>
> (Nathan, mechatronics engineer, robot developer, COBOT)

Here we see alignment of motives in the inner circle between some robot developers and those robot facilitators who have money to buy robots in order to make their production more efficient, which basically implies replacing humans with robots.

> *[I]f a robot could do twice the amount of work as a human, then I ask: [the human] needs a stop for a break, and then needs to stop working. If they [the robots] can keep going, I suppose tactically for a company or for business, that's only a good thing. For people, it's probably not a great thing because essentially it means that there's less work.*
>
> (Conor, recruitment agency general manager, affected stakeholder, WAREHOUSE)

Due to high productivity goals and the need to stay competitive, (*see 9.0 Economics of Robotization*) such a situation often seems to be more of a necessity rather than an actual choice, presumably for the benefit of the entire labor market. When it comes to the benefit of safety and efficient problem-solving, the humans are in many of our cases no longer an attractive workforce. Although robot developers do create collaborative robots, or cobots, with the aim of working together with human workers, the presence of humans is still seen as a problem in some situations.

> *The person [engaged] in the production is not used to having a robot like this around. So, in order to allow us to put a robot in a production line, they are asking [to] put this safety fence [around it]. Now, it's a laser as soon as the operator is arriving, the robot slows down. This means that the robot will never finish the work. Then the manager of the production will say the robot is not working, is always stopping, we don't need [it].*
>
> (Emilia, director of research and innovation, robot maker, COOP)

This also connects to some robot developers' view on affected stakeholders as 'simple people'.

> *And now you see more components of the system because, to work with the system, you need the elements where the people will be interacting with the robots. And by people, I mean not the robot designers but the pickers, the simple people who work in the warehouse.*
>
> (Felix, CEO advisor, robot maker, WAREHOUSE)

Rather than cobots truly working together with humans, human-robot collaboration is often limited to performing tasks simultaneously in a shared space. Here, humans are increasingly considered a problem because they have different rhythms and are perceived as unreliable. If they have to stay, production must change, and humans become a safety problem.

> *And there the idea is that the human is not fully trusted anymore but the robot is doing the qualitative tasks, because you can verify that. The robot does everything calmly and the human is doing the work where he cannot mess up a lot. I am already suspecting that the positions [for humans] will not become more interesting because of that, because of the take-over of the final installation by robotics.*
>
> (Nathan, mechatronics engineer, robot developer, COBOT)

Some developers therefore stress that in order not to be a problem for productivity and efficiency, humans have to change as well.

> *When we are speaking about collaborative robots, I have put this laser scanner [in the robot] that is detecting when the operator is coming close to it. But I always say, why is it the robot, who has to detect the human operator? Why should a human operator, knowing that there is a robot working, [not] stay far and not come [closer]?*
>
> (Emilia, director of research and innovation, robot maker, COOP)

These concerns about humans are mainly found in our cases tied to 'big scale' robotics working in production or big organizations (such as COBOT, WAREHOUSE, COOP); but there could be a tendency for this approach to humans to spill over to other sectors, as when robots are placed close to people's everyday lives in homes or public institutions, for instance. This could for instance mean changing human routines or environments (*see 5.0 Inclusive Design*). Or robot developers in the inner circle could begin to question their own capability to deal with the ethical problems arising from humans and robots collaborating.

> *Interviewer: "What do you associate with ethics questions?"*
>
> *Nathan: "In robotics?"*
>
> *Interviewer: "Yes."*
>
> *Nathan: "Hm. Well. As the robot is leaving the safety zone more and more, thank God, and comes together with the people. There you have to ask the question what makes sense and what burden can we put on the human as well. Where the robot is interacting. The best example: I was at a workshop in the conference where [they told about how] the robot was put in a kindergarten. And that is kind of disturbing, if you let disturbed children, I think it was even autistic children, if you let them play with the robot. That is questionable, yes."*
>
> (Nathan, mechatronic engineer, robot developer, COBOT)

4.1.3 Ethics as a 'problem' out of our hands

Among several reasons why robot makers tend to avoid to systematically engaging with ethical reflection in robotics is the tendency to associate ethical considerations with negative occurrences related to robots and affected stakeholders (both end-users and other affected stakeholders). As already noted, in general, when asked about their theoretical understanding of the notion of ethics, people outside the academic field of philosophy, including robot developers, often find it difficult to define ethics. When applied to robotics and real-life scenarios, ethics often appears to be more of problem and a matter of wrongdoing rather than of what is actually a right thing to do. In other words, ethics has often been seen in terms of 'ethical traps' and 'dilemmas' that apply to both robot makers and robotic systems. The relation between the two is rather straightforward: to be an ethical trap, the circumstance must present an ethical dilemma. However, we should not think robots solve ethical dilemmas; that is for the developers (Miller et al. 2017). This applies not least when robot developers relegate decision-making to their machines. For instance, the dilemma of who to save if a robot accidentally threatens the bodies of two persons and only one can be saved. In this situation, the robot may be caught in a trap where neither are saved. Much of the empirical data points to the fact that when robot developers identify dilemmas and recognize that 'something bad happens', that is when they become aware and see the importance of ethical concerns. This view is shared by robot developers and affected stakeholders.

> *When you said ethical to start with, I was thinking kind of like the worse-case scenarios, like things that could go really wrong.*
>
> (Mathias, system integrator, robot maker, SPECTRUS)

> *I think a lot of the time what happens with technology is that people don't become aware of the ethical issues before something bad happens.*
>
> (Nils, university lecturer, affected stakeholder, WAREHOUSE)

Perhaps this is also why ethics has sometimes been discussed in terms of limits, i.e. staying within a certain framework that defines what robots should not, rather than should do (as we note in a later section of this chapter, this is related to the science fiction writer Isaac Asimov's proposed 'laws of robotics'). Though research should be free to explore anything, robot developments must rely on regulations, which will prevent harmful things from spreading.

> *I think all the research needs to be done – at least for the knowledge. And there are limits in implementing things, and in the fields of application, of course. This is what ethics is about.*
>
> (Arturo, engineer, robot developer, REGAIN)

For some affected stakeholders these limits should be 'in' the designers not the machines.

> *In the context of these ethical conditions, it would be every designer or programmer simply to set limits which they cannot cross and must be aware of what the values are there and how it should look.*
>
> (Bruno, city sport facilities manager, affected stakeholder, ATOM)

In any case, ethical reflection is viewed as making limits rather than an added value to robot developers' work.

With decision-making comes the question of control. By delegating more and more tasks and decisions to artificial systems (both robots and computers), humans have less and less control over such systems and ultimately the ability to take over the tasks that have now been assigned to the machine. While it may seem that we improve work conditions by assigning to humans mainly quality control and supervision tasks, in practice the key decisions may be determined by the machine and not human logic (e.g. when flying a plane or following a specific work plan). Some interviewees (in the below cases, policymakers) do not see this as a risk of technology dependence, because they see the human qualities as seriously flawed – wherefore humans should not be in control.

> *The main goal was to replace the control that is done manually with the tools managed by man, with an absolutely objective control that the machine can give. The control carried out by man, unfortunately like all things done by man, is also subject to errors.*
>
> (Giovanni, metro company, head of unit and application expert, robot maker, OTTO)

> *You need to justify yourself. You may say: 'Well, my feeling was that it would be better to do it this way', but the point is that the machine told you exactly the opposite and you decided to go against the machine and ultimately, which is probably the most probable outcome, the machine was right Then you need to explain why you didn't follow the machine?*
>
> (Yves, policy advisor, robot maker, COOP)

Some affected stakeholders seem to have more confidence in the humans.

> *I think whatever has to do with weapons, like police or military, it must not be a robot [who makes decisions] in any case. And other things - which is about the life of the human beings and the safety of human beings, it shouldn't be a robot making decisions, or at least has the final saying or the final word. For the medical sector I can imagine that sometimes a robot is more precise to do a surgery or whatever than a human being. But it should be the doctor deciding what the robot is doing and not the robot.*
>
> (Michael, traffic controller, affected stakeholder, COOP)

The robot developers are very much aware of the problem that also some of the technologies they make may be flawed, but it is viewed as something that is beyond their own line of work because the problem exists everywhere.

> *Interviewer: "Are you ethically responsible for ensuring that this robot should not be able to be hacked."*
>
> *Alonso: "That's a huge domain. We say 'thieves and policemen' in Italy [constantly challenging each other]. Because there are always people who try to hack systems, people who develop systems that will build anti-hacker protection. There are directives on this about safety with tests, but it's always difficult to keep up, but this is a very transversal issue. It goes from Windows operating systems down to other operating systems; it goes up to applications. So indeed, whenever you develop an application that is really connected to the network, you are subject to hacking. But that is a really transversal, huge domain. So, in this, the producer has to develop – this is a software issue more than hardware – develop functionalities in a way that they obey the current safety directives. And then there is always someone who discovers how to enter and you have to deliver some patches."*
>
> (Alonso, participant in robot expert panel, robot developer, REGAIN)

4.1.4 Ethics of legislation

From our research it has become clear that many robot devices are technologies capable of recording information about use and users. If robots are wirelessly tethered and connected to the internet, they are not just devices that are used in particular situated locations, but can become recording and surveillance devices. Therefore, data (in the form of use, images, text, audio) can be captured by the device. Privacy and surveillance issues are found across cases in the robot developers' reflections. They also reflect on the question of AI built into robots for data harvesting. From the robots in agriculture, to industry, to healthcare, we find this problem.

> *I think there are ethical issues with the introduction of the robots in the robot-human interaction. Let's say you have a robot for elderly people, right? It's super amazing, but then you charge a huge price for it. Those are also ethical considerations, right? So apart from that there is also of course privacy, because I mean we are being recorded every second now. I mean if you have a phone with you, you're being recorded. There is no way around it. Yeah, and I think there is also this component of [ethics].*
>
> (Oswaldo, industrial designer, robot developer, SPECTRUS)

Robot developers often consider the issues of data protection and privacy to be dealt with by legislative measures. The policymakers have been trying to be proactive around this ethical issue. There are several directorates of the European Union that protect individual privacy and data protection, putting restrictions on how corporations (and theoretically governments) can access and use personal data of users of computer systems including the much debated GDPR regulations. [4]

Also, privacy is a core principle of the European Union. Communication and information technologies have reshaped many crucial principles and issues of privacy for citizens of Europe. Prior to the internet, robots acted as digitally connected devices and rarely linked to any wider system. Now robotic systems are often wirelessly tethered, and companies that sell their products to other companies or consumers update software systems via the internet. This means that data gathered from robots are now a currency. For ethics, this means ethics moves out of its embeddedness in a robot system. However, the responsibility of those developing robots does not seem to be likewise enhanced by anything beyond new regulations. This may be because the policymakers believe in strict and explicit rules.

> *I do believe in moral rules, which are defined by customs but also very much by the penal code and which, in a way, embodies what is considered as good or bad in the society.*
>
> (Yves, policy advisor, robot maker, COOP)

Many robot developers expect and rely on the recommendations given by the ethics committee and relevant bodies or specific individuals appointed by the company in the role of consultants. These are steps towards an effort to address ethical challenges related to information-gathering robots that even persons viewed as experts in ethics may find 'difficult' and 'disturbing' (Sparrow, 2007). Another area of legislation concerns the ethical responsibility for how and where robots replace workers. As with the surveillance and hacking of data, the choice not to employ workers (rather than replacing them with robots) is considered a problem beyond the robot developers' control. Some industries face a shortage of a suitable workforce, in particular with regards to 'unskilled' workers. Thus, rather than invest in new facilities and human resources, a company, in for example manufacturing or agriculture, may choose to implement robots that are capable of performing manual tasks in a much faster and precise manner, while working 24-hour shifts. This does not necessarily mean the company fires workers, they just do not hire new ones. Such an approach can be also described as a labor avoidance strategy.

> *What organizations may do over three to five years now is look and see what type of facilities do we require and how do we want to operate them. And one of the options they can now consider today is the use of robotic technology within those new facilities to absorb that growth of the organization. So, it doesn't necessarily have to be about labor reduction; it can be a means of labor avoidance.*
>
> (Danny, sales manager, affected stakeholder, WAREHOUSE)

This may seem mainly as a matter of an economic or political decisions, and robot developers do not express that they too have a say in creating analytical and implementation frameworks for robotic technologies, and for the way society keeps humans in the loop. Their concern is about safety and doing good with their individual robots – and following rules and regulations.

4 With regard to large corporations, they face different issues with protecting data from hackers and thieves and potentially other corporations. Here are the directorates:

Data Protection Directive (95/46/EC) of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. This directive specifies a number of confidentiality and security safeguards for this and other interactive on-line services.

Council Directive 83/570/EEC of 26 October 1983 amending Directives 65/65/EEC75/318/EEC and 75/319/EEC on the approximation laid down by law, regulation or administrative action relating to proprietary medicinal products.

The Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects (art I.3 and I.4 related to the careful assessment of risks to the subject), and all articles of section III Nontherapeutic clinical research, related to the obligation for patients' informed consent and right to withdraw as well as to the safeguard of patient's dignity and personal integrity.

The Data Protection Act (1988) and Data Protection Amendment (2003), Directive 2002/58/EC on Privacy and Electronic Communications (amending Directive 97/66/EC), regulating personal information protection across the telecommunications sector; ISO 13482:2014 for Robots and robotic devices - Safety requirements for personal care robots.

With the entry into force of the Treaty of Lisbon in December 2009, the Charter of Fundamental Rights of the EU became legally binding, and with this the right to the protection of personal data was elevated to the status of a separate fundamental right. A better understanding of Council of Europe Convention 108 and EU instruments, which paved the way for data protection in Europe, as well as of the CJEU and ECtHR case law, is crucial for the protection of this fundamental right (Publication on European data protection law p. 3)

Though the problem-solving aspects of the robot developers' conceptions of ethics, as well as their reliance on legislation, seem both reasonable and sensible, the perspective may have the consequence that the robot developers (as well as the robot makers and company owners in general who share their views) do not think of ethics from the perspective of the generally affected stakeholders in a moral and holistic view. However, we do see that ethics in engineering education is beginning to move beyond issues of safety and legislation.

4.1.5 Difficulties defining ethics

This, however, bring us back to the issue of who actually care about and engage in discussions of ethics. As noted, we find a huge group, both robot makers and affected stakeholders, who have a very hard time defining 'ethics' or associating anything with ethics at all. A cross-case finding in REELER is that for many of the people in the closest proximity to either developing, buying, using or otherwise being affected by robots, the word 'ethics' is not something they are familiar with – as these examples with two robot developers and two affected stakeholders show:

> *You enter this area without us necessarily sharing an understanding of what we mean by the word ethics.*
>
> (Elias, university researcher, robot developer, WIPER)

> *Yeah. Well, it's not so easy for myself to say something about ethics. Um, I should say five words*
>
> (Theo, university researcher, robot developer, SANDY)

> *Interviewer: "Five words that you associate with the word ethics?"*
>
> *Mette: [Gasping] "Help. No, that was not one of the words I associate with ethics, was it? [Laughs] Ethics. That is dignity, it is well contemplated. It is the proper. It is, it is hard... And it is exciting."*
>
> (Mette, affected stakeholder, COBOT)

> *Yeah, it's not something I know loads about, ethics. So, I'd say that the ethical questions aren't really what I work with most, but of course, in relation to this project, I, well, how can you say, I don't know, it's a question about the robot not hurting humans, that could be one thing.*
>
> (Alexander, university robotics researcher, robot maker, WIPER)

The affected stakeholders in general also seem to have a more varied approach to ethics which can range from having a good time at work to moral and not harming human values. The robot makers, whom we mainly focus on in this chapter, also connect a variety of words with ethics like honesty, thoughtfulness and behaving properly, but also, when asked directly, see ethics as something tied to the human sphere. As noted, they do seem to recognize it for the demands put on their work – but often have just as hard time defining it as the affected stakeholders. However, as a recurring 'absence' (*see Annex 1 on Methods and Methodology*[5]), 'ethics' does not emerge as something about taking care of end-user defined needs or affected stakeholders' concerns. It is either about safety, keeping up with regulations, making robots that seem evidently good or pleasing a customer.

4.2 Towards more holistic views on ethics while 'doing good'

Robot developers and policymakers, in particular, have been concerned with serving the public good and delivering work that may potentially benefit humanity (Khatib & Christensen 2010; Downey et al. 2007; Vesilind & Gunn 1998; Davis 1991). This view can be shared by affected stakeholders, who see no ethical problems because they believe robots are simply good themselves – not least when the affected stakeholders directly stand to benefit from the robots. In this case a manager:

5 see responsiblerobotics.eu/annex-1

Interviewer: "Do you see any ethical dangers or any other dangers related to the use of educational robots such as ATOM?"

Tadeusz: "No, I see the mere benefits of it. Please note, even a child who cannot write, cannot read, she or he takes it and learns programming, learns logical thinking, learns algorithms."

(Tad, science festival organizer, affected stakeholder, ATOM)

In general, our affected stakeholders believe that robots will do good. In so far they have something to say about ethics, it is not unusual to observe discussions of ethics in positive terms, i.e. in relation to guiding principles and norms we all aspire to, to a varying degree, that facilitate individual and societal well-being. However, some affected stakeholders express a different view and do not believe robots as such are able to do good. Though the believe some robots may help people (for instance cleaning and lifting) they are also skeptical whether too many robots in society will be good.

I think is not so good with too many robots. For some things it's ok but with the evolution there will be more and more and then it will not be good for the people.

(Anita, cleaning staff, affected stakeholder, SPECTRUS)

In contrast to the notion of ethics as safety and to avoid evil or unwanted interactions with robots, some robot developers simply find *"a robot is a machine that helps a person"* (Edgar, system architect, robot developer, SPECTRUS).

And some robot developers see themselves as people who simply do good, for mankind, simply because they develop robots.

I am still positive. And not because I'm a robotics person. I'm not telling you this from a business perspective or because this way I will become richer. It's not for that. It's because I strongly believe that the robot can help mankind.

(Alessio, robotics start-up founder, robot developer, COOP)

Some of the robot developers are, however, aware that a wider ethics cannot be separated out from technology – and that the technology cannot 'do good' by itself. Even if the technology may seem 'neutral', these robot developers are aware that they share a responsibility with users for how a robot is used to create unsafe situations for other humans:

We are responsible for the technology we develop, of course. It can be used in one way or another, it is always the same. I mean, the military uses of robotics. I am part of the robotics and the AI-research community, and we make many manifestos, many documents asking politicians to regulate.

(Carla, robot developer, BUDDY)

Some robot developers can even see themselves as affected stakeholders, when it comes to ethics, even if they do not connect these considerations with their own robots.

Interviewer: "Then we also have questions regarding the topic of ethics. What do you think of when you think of ethics?"

Valerie: "I kind of imagine safety for humans, but, well, physical, but also: 'What is happening with my job? What is [happening] with my life?' Like that."

(Valerie, mechanical engineer, robot developer, COBOT)

Robot developers do try to put themselves in their users' place when they point to safety as not only physical but also psychological human well-being:

> *A big challenge is to make sure that the robot is safe enough to interact with a human. Make sure that you will never involuntarily harm him or scare [him] – not even harm, but just scare the person that's in front of it. When the robot starts moving, suddenly it's quite normal that people get a little bit surprised and sometimes frightened and don't want to interact anymore with the robot.*
>
> (Daniel, software developer, robot developer, BUDDY)

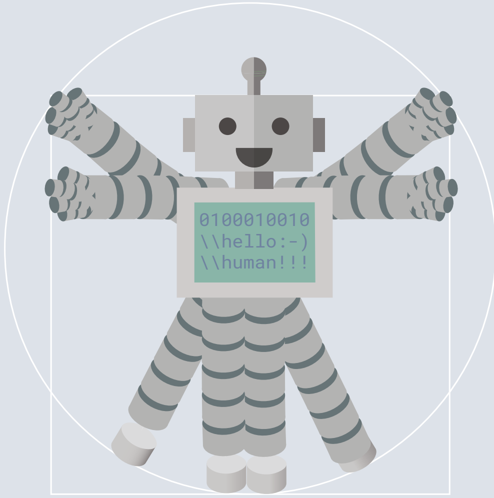
In our case material it seems that the developers' ethical considerations are also sometimes tied to the size of the company. Big companies have special people dealing with ethics whereas in smaller companies they try to develop an all-round approach to ethics which includes the engineers. For example, some companies are small enough to allow the employees to spend time discussing a given ethical concern.

> *Usually we get together, we sit, we speak and then we take a decision as a group. It is a young, small company, so if I have some ethical doubt about something I will just go directly to [a name] or to [a name], which is our CTO, or to whoever is relevant at the moment, grab a few of them, let's talk about this, let's take a decision on that. And then we will sit together, we will take a decision and we will decide where we go. We do not really have a specific process; we just get together and decide things.*
>
> (Daniel, software developer, robot developer, BUDDY)

We find the greatest awareness of ethics in some particularly concerned engineering educations, where ethical considerations also comprise trying to include more women in engineering.

A holistic ethics takes into consideration the whole person and their physical and social environments.

STORY FROM THE FIELD:
How to teach ethics



Many engineers across cases tell the REELER researchers that they did not learn much about ethics during their education, but this may change. In fact, one of the universities taking part in REELER's research has, in the past 15 years, radically transformed their engineering education to teach ethics in a more holistic way. REELER meets Elias who is a professor at the university in Northern Europe and tells us about this development.

"I think, most engineers think about [ethics], when you start working with real projects, which we also do alongside our students. [We want] to have a more holistic point of view. One thing is if you just need to create a gadget in this phone that will be able to do this and that, but when you look at how I'm using a phone as a whole, how the user uses it and what does the user require and stuff like that, then you actually begin thinking about these things."

We hear that previously (maybe 10-15 years ago), an engineer was considered a nerd, who was detached from society, and thus might create something harmful because they were 'out of touch' with people. Consequently, they used to teach ethics to engineers by emphasizing how things could go really, really wrong. Today, both the image and the practice of engineering have changed, and it is not uncommon to *"have projects with the students which help people who have limited mobility or something like that. Then you go out and talk to these people and get an idea of their world."* Contrary to previous teaching of ethics, this university no longer emphasizes the problems and the dilemmas.

"I actually think we did that, in old days. Just when [the students] started, we had these ethical dilemmas we put in front of them. They did that with a lot of commitment. But I don't know how much they learn from it in the long run. [Today] with the broad approaches we ask the students critical questions. We say: 'Hey, listen, what does this mean for the user, this thing you're doing?' Or whoever is going to use it, society, or the company, or whatever."

Elias emphasizes that the word 'ethics' has come to mean studying and working together with users – and it is *their* ethics that matters.

"Well, I don't know, what we mean with the word 'ethics' but now we are coming and looking and working with users. You get to be a part of their ethics, like, their world. No matter if it is something that matters to us, right, and can it cost more. Yeah, I don't know, I think there are a lot of aspects related to ethics."

Contrary to the more problem – and safety-oriented engineers, Elias and his colleagues find it important to take responsibility for how their engineering work also affects society and wider contexts.

"Today I'd say that we have a clear view for them [students] to go out and work in a context, where both society and people are an important part of the success criteria and framing conditions. So, you can't create anything technologically isolated without having to decide your position in relation to it. Especially not if you have [a] more holistic education, which I think we have a lot of, then you can't avoid being a part of those conversations."

Elias explains further that in the past 15 years a number of engineer courses has popped up - not just at his university, but in various places - where design and the wider context are also included in ethics teaching. Here, particularly welfare technology, is an example of a domain *"where you go in and teach something else than purely techniques and technology"*.

(Based on an interview with Elias, university researcher, robot developer, WIPER)

Very few across the REELER cases seem to share the impression mentioned by Elias that ethics is about *"coming and looking and working with users...You get to be a part of their ethic, like, their world."* (Elias, university researcher, robot developer, WIPER).

This may be because the idea of thinking of ethics holistically and beyond the machines is still new. Most robot developers still feel they need ethics to be something they can put into a formula – which is not possible with a distributed, holistic ethics.

4.3 Beyond the robot makers' views on ethics

In addition to the robot makers' own views on ethics and our analysis of ethical responsibilities tied to the robot developers' own associations, we will explore one more theme in this chapter. This is the issue of how ethical concerns are distributed. This theme is *not* addressed by the interviewees themselves. Nonetheless, our analysis across cases identifies two problems tied to ethics that are rarely dealt with in the extensive literature on ethical guidelines for engineers.

The first concern is the inherently *distributed* character of robots. A robot consists of many different technical hardware and software parts and thus we should expect that the ethical responsibility for making responsible robots is distributed accordingly. In none of our cases is the robot developed in one place only. It is always developed in many different places, by different people. In a typical case, some robot developers take care of, for instance, software, while others handle the hardware, a third group integrates the two and others take care of marketing. Sometimes a particular person is appointed to take care of ethics. Acknowledging this fragmented distribution of responsibility for robot parts, raises the question of where to place ethical responsibility: Is it in the software, the hardware, the assembly, the management, the marketing, or the appointed 'ethics persons'? Furthermore, contrary to the expected, we do not find ethics distributed across all of these aspects of robot development. Ethics, as a professional concern, is rather *isolated* in special departments, or groups of professionals, where it is identified and debated in a closed set of people. In other words, the parts of the robots are distributed, while the ethical debates are isolated (or 'undistributed').

A. Robot's parts are, as material objects, geographically distributed, which risks to dilute ethical concerns for the robot as a finished product.

And at the same time,

B. Ethical concerns are fairly isolated i.e. only concerns a smaller group of people in ethics committees and philosophy.

4.3.1 Distributed technology = distributed ethics?

Design of robots is distributed across time and space all over the globe. It follows that ethics should be likewise distributed, but this issue has not been dealt with in the robot maker's reflections.

In our analysis of our general data collection, which included many visits to robot laboratories all over Europe, we found a pattern of ethics not being brought up by any of our interviewees and which is also hardly touched upon by philosophical or academic papers in general. However, as we define ethics in robotics we must take into account that technology is *distributed*.

In all the 11 case studies of different robot types we find a similar way of working: robots are never built in one place (from scratch). This will not come as a surprise to those who build robots, but it does raise some questions about how to make ethical and responsible learning in robotics, when the people and technologies are often only connected occasionally – and robots are tested as assembled in the later phases of a project. Furthermore, in all of our cases, the robot developers never build a robot from scratch but include 'off the shelf' components, which they themselves do not feel responsible for. Within each case, the technical parts are developed in different places by different people, even in the biggest companies. This is a finding across all cases regardless of robot type.

In one case, SANDY, a university in Spain is coordinator and responsible for software parts, a university in Belgium is responsible for making obstacle detection and viewpoint analysis, a company in Sweden works on grippers, a company in Turkey on autonomous mobile platform, a company in Netherland takes care of testing and a Swedish university takes care of dissemination and analysis. This may look like a typical EU project with cross-country collaboration, and we have a couple of these in REELER. However, this pattern applies, to some extent, to all the cases – whether distributed geographically within Europe or disciplinarily within a single company. The point is that with a very distributed technology, ethics *should* be both distributed and centrally coordinated to ensure the assembled robot is ethical. Yet, we see no indication of this in our data. On the contrary, ethics is often delegated to special people with special functions severed from other parts of robot development.

> **Distributed ethics:** *Making robots is distributed – but that also means ethical responsibility for the whole project should be distributed.*

If no one in a robot developing group takes responsibility for the ethics of all the robot parts, a robot, with software that turns out to be hackable, can become unethical, even if all the other hardware parts have been carefully evaluated as ethical. Put differently, the ethical responsibility for a given robot lies not in its separate parts but also in the way they are combined, and eventually the way they are used and misused. In this way, ethics in robot technologies is inherently a matter of relational and distributed responsibility.

STORY FROM THE FIELD:

Distributed ethics – the EULA robot

On K-BOT's premises in a Northern European country we find several factory halls and warehouses. K-BOT robots are combined with other K-BOT robot parts and several robot cells are showcased. EULA is one of K-BOT's newly developed robots. The innovative development of novel robots takes place in two offices: the Innovation Office (INO) and the Company Investigation Office (CIO). It also involves collaboration with many suppliers across many processes.

In INO the employees try to think 10-20 years ahead. They work with *"futurology"*, as Peter, Innovation Manager, explains in an interview. Johanne, also Innovation Manager at INO, explains they are involved in *"blue sky"* research, and several interviewees mention that INO carries out *"the evangelism of robotics"*, a term Peter uses himself and which spreads the gospel of robots' blessings. Both Johanne and Peter often represent and promote K-BOT at fairs and conferences accompanied by the CEO. They are spokespersons for K-BOT at conferences, in newspapers, and in robot and business associations.

From a REELER perspective, ethical considerations should be part of the futurology efforts just as the ways K-BOT present robot futures can evoke ethical issues.

CIO conducts research in novel robot developments. Some of the researchers also call this the pre-development department because it comes before product development. Here, the robot developers work on projects that are still 5-10 years away from being a marketable product. As one robot developer quoted himself saying to K-BOT's management: *"It's research! You cannot sell it!"* CIO has approx. 30 employees (not counting unsalaried students and salaried student assistants) to develop ideas for future productions, and write grant applications, reports, etc. CIO has stable funding from the company's own sources and is, according to Kai (cluster leader in CIO), not dependent on research grants.

This type of work also calls for ethical considerations as the ethical responsibility, to some extent, can be seen as distributed among the application writers and the grant providers.

Robot developers and innovation economists generally advocate for researchers to follow their own research interests in research and development (R&D) phases, across publicly funded research projects and independent of grants. K-BOT's robot EULA is made in this way. It is the result of a technology first developed at the State Aerospace Centre, then moved to the K-BOT company which developed it to its present TRL9. Today, the robot is in mass production at a K-BOT factory.

To develop and build EULA, engineers have been working in different teams. Some teams are responsible for the mechanical constructions, other teams make the control system and the control cabinet. Others develop the software. A design process of a new robot usually takes a year and involves several meetings between a chief designer and different groups of engineers, those who do the cabling and those who insert the motors, etc., where the designer refines and adapts his design.

Different companies and subcontractors deliver the parts for EULA. The transmission equipment is, for instance, from *Smooth Drive*. The motors come from *PS Systems*, and the sensors from *ReadyDrive*. Both *PS Systems* and *ReadyDrive* are spin-offs from the State Aerospace Centre. The rolling bearings come from a Dutch company (*The Dutch Ball Bearing Company*), a French company (*TXT*), and other big bearings companies. In the R&D phase, the prototypes are typically 3D-printed, but ultimately, they will be produced either by K-BOT itself or by a range of other companies.

In another case, one of EULA's mechanical parts is combined with biopsy equipment, so that it can make biopsies on cancer patients. This has many applications and will affect many people (male and female) and their work environments. Consequently, all of the various applications involve ethical aspects.

EULA can also be seen as just one component, e.g. in an assembly line where it is integrated by system integrators such as the system engineers from K-BOT's own Application Engineering Team. It is always necessary to integrate, coordinate and match EULA, and other robots, with the other business customer's – and operators' – needs. Here a number of ethical aspects arise.

Above all these different departments and processes is K-BOT's management and board. Parallel, and not part of this chart, are divisions in other parts of the world, not least USA and Asia.

The K-BOT company has its own ethical officers who invite company staff to take, for instance, electronic courses in ethical behavior. Their work is, however, severed from the other departments, and ethics consequently becomes rather detached from the actual development and sale of EULA.

(Based on interviews with Peter, Johanne and Kai, robot developers, COBOT)

4.3.2 Isolated ethics

As in the EULA case above, the people knowledgeable about ethical issues are often experts who actually are interested in and debate ethics. Across cases in REELER, they seem to constitute a small group of people tied to certain functions – whereas the main bulk of interviewees are not engaged in these debates. This leave their voices out of the debates.

From a REELER perspective, this raises the question who is given voice when it comes to ethics in robotics? We have tried to give voice to a new group of interviewees, the affected stakeholders, but it has not always been easy. One of the ethical challenges identified within the REELER fieldwork emerged in relation to the process of participant recruitment for our research. In general, in order to involve an individual employee in the REELER study, an approval needs to be obtained from a relevant supervisor. In several cases, an employee expresses interest in participating in the interview, but for different reasons he or she could not obtain the needed approval *(see Annex 1 Methods and Methodology)*.[6]

To give voice to this group of people is definitely beyond the responsibility of robot developers, but it is connected to the issue of who is given a voice within the 'inner circle of robotics' *(see 1.0 Introduction)*. Customers and clients (people who buy robots) definitely have a say, but the same does not go for the people eventually working in close proximity with the robots or otherwise affected by them.

Furthermore, it is important to note that while, in theory, many of the committees in the area of Robotics/AI Ethics are open to anyone who is willing to contribute to these areas, in practice, they usually consist of the individuals who come from academia, industry, or public institutions, and hold a certain degree of power or a relevant position. In other words, there are only limited possibilities for the average end-user, such as manufacturing workers or middle-level managers, to participate in research or initiatives that would influence guidelines for responsible robotics and ethical robot design.

This poses serious questions about the validity of work on robots and ethics, where only a small group of individuals decides on the fundamental issues that affect society as a whole. Also, serious ethical concerns emerge with regards to the transparency of projects and practices in robotics R&D, with particular regard to projects that are often partially or fully publicly funded. Following the assumption that 'whoever understands what the robot can and can't do, has responsibilities assigned' (to paraphrase one of the REELER participants), one could argue the main way to create an inclusive framework for robots and ethics, and to increase transparency, is through education that would apply to all.

It is interesting to note that several study participants argued that ethics should be 'a must' in robotics. This includes imposing the approaches that would actually force robot makers (both at the individual and institution/company level) to incorporate ethics into education as well as professional practice. The latter includes imposing top-down approaches that potentially could come from the European Commission, something that the Commission itself has suggested in the document *Artificial Intelligence for Europe*: 'the importance of ethics in the development and use of new technologies should also be featured in programmes and courses' [7]. The way ethics could be imposed in robotics research is of course through legislation as well as different types of guidelines that may also become part of legislation. In any case, however, in order to be effective, education and regulations related to ethics should eventually become a part of the robotics culture in a holistic approach to ethics.

> *If you think of the technical side, you could implement regulations of what robots can do and what robots cannot do. That's one thing, but I think if you teach this ethics for the people that are going to design, then you create this culture.*
>
> (Oswaldo, industrial designer, robot developer, SPECTRUS)

Nonetheless, the question is whether it is enough to change the culture of engineering – and whether the responsibility for ethical robots should be placed solely on the robot makers (the developers, application experts, and the facilitators who fund and make policy regulations)? This question seems to be a new one in the existing discussions about ethics in the robotics community.

4.4 Ongoing theoretical discussions of ethics

In order to identify whether REELER can bring something new to the ongoing discussions of ethics, we have made a review of ethics in relation to robotics. As already mentioned, the present discussions of ethics, relevant as they are, do not seem to have great effect in the community of robot developers. Another point we want to make is that while the discussions identified in our review on ethics are all relevant, they overlook two important aspects. Discussions focus on areas like safety, robot rights, taxation, and robot autonomy, but largely seem to overlook the problem that a) affected stakeholders' motives and perspectives are missing when robots are conceived and created, and b) ethics is difficult to work with unless underlying motives are aligned and the distributed character of robot making is taken into account.

6 see responsiblerobotics.eu/annex-1

7 https://ec.europa.eu/transparency/regdoc/rep/1/2018/EN/COM-2018-237-F1-EN-MAIN-PART-1.PDF

Ethics of robots is still a fairly new field. Ethical inquiry in relation to the design and use of robotics goes back to 2002, when IEEE held its first ever workshop on roboethics, but ethics as a discipline goes back millennia. In broad terms, ethics can be understood as the inquiry into right and wrong – both in relation to the individual and society. More precisely, many debates refer to *normative* theories of ethics, i.e. theories dealing with how individuals *ought* to act (Driver 2007). Traditionally, such normative theories have figured in the design of robots under the heading of *engineering ethics*, a field that dates back to the 1970's (Weil 1984). Engineering ethics is concerned with identifying and grounding moral conduct of engineers in their practices, but has often been bogged down in theoretical discussions about the need for codes of conduct (Luegenbiehl & Puka 1983). Such discussions have engaged with the three, dominant ethical traditions: *deontology, utilitarianism,* and *virtue ethics* (e.g. Martin & Schinzinger 2005).

This chapter will not provide an in-depth account of each of these; however, we will provide a very rough outline. Deontology, the study of *deon* (meaning duty) is the study of moral obligations. On this account, right conduct is acting in accordance with some duty or obligation, e.g. don't kill innocents – no matter the consequences (Alexander and Moore 2016). Deontological consideration often underlies strict codes of conduct specifying the duties of engineers, e.g. from the American Society of Civil Engineers (ASCE):

"Engineers shall hold paramount the safety, health and welfare of the public and shall strive to comply with the principles of sustainable development in the performance of their professional duties." (ASCE 2017)

This highlights an often-voiced criticism of deontology, and of many such codes in robotics – they do not provide actionable guidance (e.g. Sorenson 2019). Turning such imperatives into actionable practice often requires serious interpretation work, which everyday duties leave little time for (Ross 2007).

Utilitarianism, often popularly surmised as "the largest possible good for the largest possible amount of people", claims that the good lies in the maximization of total utility. Utility can be cashed out in different ways, but it is traditionally done in terms of happiness. On the utilitarian account, an action is good to the extent that it results in an increase of the total amount of happiness in the world. Such approaches are rarely codified, but often serve as the default mode of reasoning about moral issues. REELER research problematizes this approach in robotics. If it is left to the 'inner circle' to define the 'happiness' brought by robots, it glosses over the many unheard voices of affected stakeholders.

Finally, virtue ethics centers on the cultivation of virtue, and does not provide a rough-and-ready recipe for action. Virtue is defined as a habituated, characteristic, proportional and reliable disposition to act in a certain way (Annas 2011). For instance, embodying the virtue of generosity is to be motivated to reliably act generously (more in times of abundance, less in times of need), and to feel good about doing so in a way, which is in accordance with the agent's character (Annas 2011). In relation to REELER findings, virtues are no guarantee for responsible and ethical robots, as even the most virtuous robot developer may overlook stakeholders' needs and concerns if not learning about them.

4.4.1 Theoretical safety perspectives

However, REELER's review of ethics reveals that for the most part, ethical concerns in robotics are on safety, autonomy, and robot rights. In early papers exploring robot ethics, the model of ethics from Issac Asimov's short story Runaround (1940) were cited regularly:

A robot may not injure a human being or, through inaction, allow a human being to come to harm.

A robot must obey orders given by human beings except where such orders would conflict with the First Law.

A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

It is surprising how frequently these 'laws' are mentioned in academic papers on robot ethics. They are also surprisingly often referred to by the robot makers in REELER's data (a fifth of our interviews with robot makers refer to Asimov's laws). In practice, these laws are considered too sophisticated to program into robotic systems that rarely can operate autonomously for significant lengths of time without some human supervision (e.g., driverless car, social robot).

Furthermore, Asimov built on a thought experiments, and had no idea of how complex a situated practice with the actual robots developed can be. Thus, through REELER's data we identify a gap between how some robot makers (e.g. EU and ethics experts) envision roboethics and robot regulations and the actual work on robots. The robot makers referring to Asimov in general seem fully aware of this difference, but also refer to it as an ethical guideline. This could indicate a need for a more comprehensive understanding of ethics that takes root in situated everyday lives (as attempted by REELER).

Do REELER results add other new perspectives on the debates on safety? Of course, robot developers have long been aware of different risks related to the implementation of robots in the proximity of human beings. However, efforts to address what could be seen as ethical concerns have been traditionally limited to the consideration of human safety (bodily and mental integrity of people) in close proximity to machines only.

In general, starting from the 1970s, the field of engineering ethics has emerged with a focus on safety. In line with the engineering principle of serving the public good, different institutions and organizations formalized engineering ethics into codes, canons, standards, etc. (for a detailed discussion, see Sorenson, 2018). Most of them emphasize human safety

Safety is the most common association made by robot developers, when asked about ethics. (Photo by Kate Davis)

in the first place.[8] Much of what is written here is in line with those robot makers, who also consider ethics to be first and foremost about safety. Such thinking applies to robotics-oriented design approaches as well as different engineering codes of conduct and legal regulations. A recent statement by the European Commission on AI, robotics, and 'autonomous systems' also points to safety and security as one of the key ethical concerns (European Group 2018).

Safety related to the robot's contact with a human being has also been viewed as an inherently technological challenge (Association 2013, 1631), as also found in REELER research. One could be asking, however, what makes ethics particularly important for the current developments in robotics. One of the main reasons is in an increasing and close integration of robots into our society: The moment robots are placed in the human physical and social spaces, these new ethical concerns emerge. Traditionally, to a large extent safety was ensured by keeping people away from robot 'caged' or 'enveloped' robot workspaces (Floridi 1999). However, the moment people and robots get to share physical (and social) spaces, the application of 'a segregation paradigm' (Bicchi et al. 2008) is no longer possible. With new types of robots that go far beyond industrial applications, these new challenges of human-robot interactions (HRI) emerge. New types of robots endowed with an increasing degree of intelligence and autonomy as well as situated not only in the physical but also human social spaces, have emerged with new ethical concerns. And yet, the ethics in robotics, as confirmed by REELER, continues to be seen mainly through the safety perspective without addressing these wider ethical issues through research.

However, the need to address such concerns in a systematic manner has resulted in various efforts made to develop research and legal frameworks for responsible robotics and ethical robot design. The strong emphasis on certificates and legal regulations that ensure that robot as a product is safe for its users fits with the applied and technical nature of robotics research. These include, for example, safety regulations for the machinery used for industrial applications.[9] Many new types of robotic systems like autonomous cars have yet to receive appropriate regulations. Others, such as personal care robots, have already been addressed in terms of safety

8 **Institute of Electrical and Electronics Engineers:** We, the members of the IEEE ... agree: to hold paramount the safety, health, and welfare of the public, to strive to comply with ethical design and sustainable development practices, and to disclose promptly factors that might endanger the public or the environment. (IEEE 2018)

American Society of Civil Engineers: Engineers shall hold paramount the safety, health and welfare of the public and shall strive to comply with the principles of sustainable development in the performance of their professional duties. (ASCE 2017)

National Society of Professional Engineers: Engineers shall hold paramount the safety, health, and welfare of the public. (NSPE 2007)

9 ISO/TR 20218-1:2018(en) Robotics — Safety design for industrial robot systems — Part 1: End-effectors; SO 12100:2010, Safety of machinery — General principles for design — Risk assessment and risk reduction

requirements , or otherwise fall under existing sectoral or industry regulations (health privacy laws, for example).[10]

Over time, the notion of ethics in robotics has begun to expand beyond safety addressed mainly in terms of physical injuries or damages to health and related direct human-robot interactions. For example, on the one hand, the British Standards Institution (BSI) standard *BS 8611: 2016 Robots and robotic devices*, builds upon the existing safety requirements for different robots; on the other hand, it makes it clear that the Standard in question views the question of physical hazards and safety design features as part of ethical design, but that these are covered by safety standards. In other words, 'Ethical hazards are broader than physical hazards'.[11]

A typical approach, however, is to address wider ethical frameworks in robotics but still link them back to the question of safety. *The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems* launched in 2016 by The Institute of Electrical and Electronics Engineers (IEEE) is a good example. The initiative brings together hundreds of experts and stakeholders with the goal to inform and influence debate and work on the autonomous systems in a way it is guided by human-centered values and priorities human well-being. While the initiative follows a broad perspective on ethics according to which 'ethical' is not synonymous with 'safe', safety is still a recurring theme across different areas covered in the report and it appears among the main indicators of the human well-being (Ethically Aligned Design, Version 2 (EADv2), 2017). Also, when discussing ethical principles in robotics and roboethics, the European Commission and related European Parliament's Committees tend to put an emphasis on the 'protection' of human persons, in particular human basic rights and freedoms (e.g. human dignity or liberty) (*see for example European Parliament 2017*). In this sense, the main focus remains on a person's bodily, mental, and social integrity with robotics technologies being seen as a threat.

Throughout this publication, REELER adds a number of perspectives to the ethics as safety approach through ethnographic research into people's everyday lives – notably that humans are whole persons and their engagement with robots should be viewed holistically not just as a matter of technical dependability and safety.

4.4.2 Autonomy and ethics

Ethics in philosophical debates, policymaking and general academic papers often discuss the ethical implications of robot autonomy. REELER's data does have some references to robot autonomy in relation to ethics – but not so much as found in the general academic literature on robots. This is probably because there is a huge awareness among robot developers (however not the policymakers) that robots are in the end no more autonomous than allowed for by their human controllers – and no more autonomous than the batteries or electrical cords allow.

Furthermore, experts in the field agree there is no commonly accepted definition of autonomy in the AI or cognitive sciences (Frose et al. 2007, 455; Vernon et al. 2015). Autonomy of a robot implies some degree of freedom from its human controller (Frose et al. 2007). Autonomy may be set on a continuum with autonomy at one end, and heteronomy (its antonym) on the other. Or as a spectrum that includes different kinds of self-determination of a system: autonomy, supervised autonomy, or behavioral autonomy, operating in the same system. One such definition of autonomy is given here.

Autonomy can be defined as:

"the degree of self-determination of a system, i.e., the degree to which a system's behavior is not determined by the environment and, thus, the degree to which a system determines its own goals". (Vernon et al. 2015)

Behavioral autonomy represents a form of autonomy that is behavior led. Behavioral autonomy can be characterized by at least two distinct attributes:

(a) the degree of autonomy (i.e. the extent to which a system is assisted by a human in the achievement of its goals and the execution of its behavior), and

(b) the strength of autonomy (i.e. the extent to which a system can deal with uncertainty or unpredictability in any aspect of achieving its goals). There is a continuous spectrum of both degree and strength.

However, even in AI, the human controllers decide what is the ultimate goal of the 'autonomy' displayed by robots.

It seems that a lot of attention in ethical debates is paid to an issue which is not yet as relevant as many of the aspects on ethics unfolded in this publication. After all, machines are built by human robot makers and roboethics is 'human ethics applied to robotics' (Veruggio et al. 2011). Thus, from a REELER perspective, the concern is about human decisions on autonomy, not the autonomy of machines.

When REELER began its studies, we expected to find many ethical issues tied to machine built-in autonomy, but in practice this seems to be an overrated ethical concern tied to the idea of 'autonomous, intelligent robots' prevalent in media representations *(see 8.0 Imaginaries)*. The problem with this misconception is that this *also* may be seen as a way to relieve the robot developers of their responsibility for ethical thinking and agency. If the robot is autonomous, so, it could be argued, is its ethics.

10 ISO 13482:2014: Robots and robotic devices -- Safety requirements for personal care robots

11 BS 8611:2016 Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems

4.4.3 Robot rights

Closely tied to these discussions is the debate on robot rights – which seems to stem from a conception of robots which is completely out of line with what REELER researchers have found. Even the most social robots in our sample (ATOM and BUDDY) are so much machines that the idea of granting them robot rights is misleading. Not least in terms of ethics.

The strong focus in Academia and policy making on robot rights does not seem to be grounded in any real close proximity to actual robotic devices *(see 1.0 Introduction)* – and the consistent talk of future intelligent beings is out of line with the real robots created by robot developers (e.g. one of our participants has talked about a Hollywood version of robots).

In January 2017, the European Parliament began to draft a new set of regulations as part of the Commission on Civil Law Rules on Robotics. The document proposes to regulate the development of robots, so that the technology is developed safely, considering the ethical and social effects of the new technology. Moreover, the Fourth Machine Age (robots and AI) is predicted to radically alter work practices across the world, as robots and AI replace human jobs. The impact of these developments could increase unemployment while simultaneously reducing social security payments in the form of taxation or national insurance contributions to nation states. As a way of addressing this potential issue, the Committee on Legal Affairs propose the introduction of the term 'electronic personhood' for robots, to make companies and corporations liable for potential harms of the technology and reduced funding for state welfare provisions.

Personhood is a legal category, designed to indicate rights, responsibilities and obligations. Personhood is a controversial category, as it has historically been applied to both persons and things. The concept of 'corporate personhood' for instance developed in parallel with the concept of 'the person' in Western liberal democracy. The person was a legal construct that evolved out of the Enlightenment humanism (Davies and Naffine 2001). The use of the term 'person' is not without its problems. The person became a legal category at the onset of the liberal western democracies. The term 'person' was used in the US constitution's Fourteenth Amendment which included the rights of free slaves to be recognized as 'persons' and was later taken up as a term for corporations to be able to access the same rights as human beings. This is termed 'corporate personhood'. The extension of the political franchise to include those other than wealthy white men, was gradually extended over 400 years to include Black people, former slaves, working men and all eventually all women. In extending the franchise, inherent in the legal personality was a new way to represent humans as forms of property.

The attribution of the category of 'personhood' to robots opens up a minefield of issues from a human point of view. If a robot becomes a person, for example, does that mean it will need (is entitled to) the same treatment as a human being? Will the robot need holidays and breaks? Will it be 'cruel' to use a robot as an instrument?

From REELER's perspective, these questions seem superfluous compared to the real problems encountered by affected stakeholders, who in some cases are not really considered persons by robot developers, but nuisances and threats to efficiency and productivity. These real ethical problems, covered by REELER, seem to be more or less unnoticed by the ethics communities – and in any case overshadowed by the (granted!) much more spectacular debates on robot rights.

4.4.4 Roboethics

> **Roboethics:** *A field of ethical inquiry, which deals with how humans design, interact and relate to robots. In particular, how to ensure that the spread of robotic technology benefits rather than harms humanity.*

Since at least 2002, another field of ethics has sprung up, which deals specifically with robotics. This field, called roboethics, discusses the ways in which we design, use, and relate to robots (Sullins 2011) and it has continuously been a source of inspiration for REELER.[12] This new way to address ethical challenges in robotics places a responsibility on the engineering culture rather than just relying on regulations. Given its highly interdisciplinary character and a relatively high participation of social scientists and philosophers, roboethics has the potential to actually widen the scope of ethical reflection in robotics and bring it beyond narrow safety-oriented considerations. From a research point of view, roboethics covers a large variety of perspectives and disciplines that may significantly vary in their focus and approach (Tzafestas 2018; Crnkovic & Çürüklü 2012). While some approaches involve assigning ethical and moral capabilities to robotic systems (e.g. (Wallach & Allen 2009; Arkin 2009), a dominant approach proposes that 'roboethics is not the ethics of robots, nor any artificial ethics, but it is the human ethics of robots' designers, manufacturers, and users' (Veruggio & Operto 2008, 1504).

On the one hand, since its foundation in 2004, the field of roboethics has been growing. Since the early 2000s, a number of conferences, symposia, and workshops on roboethics took place, from the events associated with technical conferences such as for example ICRA (IEEE International Conference on Robotics and Automation) to the conferences fully dedicated to ethical standards in robotics or 'robophilosophy' (for a detailed review see for example (Tzafestas 2018). Other examples include providing guidelines in the form of a Roboethics Roadmap (Veruggio 2006) or a taxonomy for roboethics (Steinert 2014).

Roboethics also deals explicitly with the relation between media representations and the 'uncaged robots'. Here robots seem to be more controversial tools than other technologies in for instance healthcare (telecare, ICT, social networking, not

12 This is similar but distinct from a related subfield, machine ethics, which deals with the possibilities of moral machines and their behavior.

to say these technologies are free of controversy) because of their prior and parallel status in popular culture *(see 8.0 Imaginaries)*. Robots are not merely objects of the laboratory, but also of screen fictions and literary tales (Richardson 2015; Reilly 2011; Reichardt 1978; Breazeal 2002). For Europeans and North Americans, robots are not neutral cultural objects but are presented in popular culture as threatening and disturbing (Richardson 2015). These disturbing perceptions of robots are not helped by recent surveys that suggest robots and automation could put half of the world's population out of work (Yuhas 2016) which is also a concern in roboethics.

On the other hand, however, efforts to systematically engage with ethical reflection and roboethics research within a robotic community continue to remain limited. This is due to a number of factors, for example the very distributed nature of robotics research as well as developers' formal education that often leaves ethics unaddressed (except for safety concerns). It has been argued that until recently, social and ethical implications of robots have been "largely ignored" (Bekey, 2012) and the main ethical position within the robotics community was that of 'not interested in ethics' (Veruggio and Operto 2008).

Both of these claims are mirrored in the findings of REELER. However, REELER differs from most roboethical discussions by its ethnographic approach – diving deeper into the reasons why there is a gap between a community of ethics specialist and the community of robot developers.

We are not alone in emphasizing a closer collaboration with end-users and other affected stakeholders, through social scientists as intermediaries.

In recent decades, discussions on roboethics have shifted more towards the process of design, and are less interested in the individual engineer. According to Wynsberghe & Robbins (2014) two (or three, if one counts their own contribution) schools within this field can be distinguished. The first school:

"believes that ethics ought to be incorporated into research and design practices and holds a pragmatic view of ethics – that ethics in this arena must facilitate the design process rather than hinder it." (ibid., 948).

Furthermore, the advocates of this school argues that ethical reflections of this sort can and should be carried out by the practitioners themselves (Ibid.). The dominating approach within this school is *Value Sensitive Design* (VSD) (Friedman, Kahn & Borning 2002; 2006). VSD is: "a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner" (Friedman et al. 2002, 1). This is done through three different 'investigations', conceptual, empirical and technological. The conceptual investigation deals with questions such as: what do we mean by a particular value (e.g. transparency)? "Who are the direct and indirect stakeholders affected by the design at hand? … what values are implicated? … Should moral values … trump non-moral values?" (Friedman et al. 2006, 351). The empirical investigations complements the conceptual by going beyond armchair speculation and asking questions such as: "How do stakeholders apprehend individual values in the interactive context? … Are there difference between espoused practice (what people say) compared to actual practice (what people do)?" (ibid., 352) These questions are answered by utilizing the full range of empirical methods, quantitative and qualitative. Finally, technical investigations evaluates, which technology best serve to realize the value identified in the prior investigations (ibid., 353).

Other approaches, although less prevalent in the literature, within this school are *participatory design* (sometimes called co-design) (Muller, Wildman & White 1993, Spinuzzi 2005, Oswal 2014), centering on the inclusion of stakeholders into the design process, stressing the importance of exposing system designers to everyday realities of end-users. Similar in scope and purpose are the *Human-centered design* methodologies (e.g. Rosenbrock 1989, Giacomin, 2014), which emphasize the understanding of end-user needs. This involves determining the design requirements and defining design concepts based on what is known about the people involved, and what is known about the environment in which the interaction takes place (Giacomin 2014).

The second school believes that reflection on ethics in the context of design should be broader, and deeper, than engineers' training and time allow for (van Wynsberghe & Robbins 2014). For instance, through the moral philosophy that "helps engineers to interpret their responsibility and think more critically about it." (van der Burg & van Gorp 2005, 235). By the critics' accounts, such approaches are heavy on theoretical thinking, but contribute little in the way of actionable guidelines, van Wynsberghe and Robbins claim (van Wynsberghe & Robbins 2014).

Finally, the third school attempts to merge the two schools in demanding rigorous, deep and actionable ethical reflections, carried out by social scientists or humanities scholars (often ethicists, but not exclusively) – in cooperation with robot makers. These outside experts are to be part of the design process from the beginning, working closely together with engineers and contributing to the end product (van Wynsberghe & Robbins 2014). This is the youngest of the three, and currently van Wynsberghe and Robbins' 'ethicist as designer'-approach has gained the most attention, though other approaches exist (e.g. Seibt, Damholdt & Westergaard 2018).

It is with this 'school', that REELER has the most affinity. However, we also acknowledge the many issues arising when robots are developed and implemented, where experiments with robots are often ongoing. Therefore, we may need a new kind of applied ethics that relies on a new role for social scientists in development, which we call alignment experts. These alignment experts will have the sole function of enacting ethics as relational responsibility.

4.5 Ethics as relational responsibility

Our cross-case analysis has opened for a new understanding of ethics than r is currently represented by the bulk of literature on ethics and also differs from the understandings we find in the robot developers' own reflections. While the chapters in Part Two of this publication focus on novel aspects of ethics that robot makers can address and act on (e.g. working more inclusively) *(see 5.0 Inclusive Design*) to generate a more user-oriented innovation economy (*see 6.0 Innovation Economics*), we have also found an aspect of ethics that may be more difficult to deal with: as design processes are distributed so should ethics be. Furthermore, we have found that situated practices pose many new problems that are not resolved with safety regulations or guidelines (*see 7.0 Learning in Practice*).

In this last section we introduce some of our own suggested solutions to the problems we have identified in REELER. We want to emphasize that in order to raise awareness of ethical issues, we need relational responsibility for making ethical solutions in robotics, which we develop here and in the conclusion of this publication.

In practice, what matters is not only how we define and study ethics in robotics, but also how we incorporate ethics into the actual design thinking and practices. This applies as much to individual robot developers as to the community of robot makers as a whole. Robots are socio-technical systems and the robot design is always distributed among different parties involved and situated in complex physical and social contexts. In this sense, in order to pursue responsible robotics and ethical robot design, it is of course much more complex than simply prescribing rules of conduct or delivering ethical guidelines for robotics as has already been done in rich measure. The key emphasis here is on the understanding of ethics as a form of personal and collective engagement. The collective responsibility points to human persons standing on both ends of the robot design process rather than on robotic systems or safety concerns or abstract considerations detached from the actual robot design and development process.

In general, when addressing responsibility in relation to robots, responsibility has been typically understood as a matter of individual accountability for the robot conduct and errors as well as resulting harms. Given the complexity of the robotic systems (in particular their increasing degree of autonomy), as well as a distributed nature of the robot design, attempts to delineate the corresponding responsibility often stop with pointing to the 'responsibility gap'. We argue here that to a large extent, the difficulty is due to the application of the individualist tradition to the notion of responsibility and limitations that come with it.

In fact, there has been a growing recognition of the need to go beyond a narrow understanding of ethical problems in the AI and robotic systems, in particular the 'Trolley Problem' approach [13] and there is a need to embrace the totality of the social and cultural contexts robots and robotics are part of. This includes development of such approaches as for example 'network responsibility' and 'distributed responsibility' (Ethics Task Force 2018; Crnkovic 2012) and the corresponding idea that responsibility and accountability should be shared among all actors involved in the design and use of a given robotic system. We propose here to bring such thinking further to include the notion of 'relational responsibility'.

The key concept in relational responsibility is, of course, that of relationship. A traditional ideology of individualism favors the conception of the human being as an individual endowed with subjective agency and the capacity of rational deliberation independent of the surrounding social, cultural and historical context (McNamee & Gergen 1999). From this perspective, single individuals are thought to be fully responsible for their own conduct and the ability and willingness to take and attribute moral responsibility is viewed as an integral and fundamental part of the conception of the person (AI, Robotics and 'Autonomous' Systems 2018). Relational responsibility emphasizes the role of relations where "individuals are such only by virtue of their creation in relationship" (McNamee & Gergen 1999, xi). In other words, the notion of relational responsibility relies on the assumption that humans are 'relational beings' (Gergen 2011) where all the meanings and language we share, including our understanding of morality, are constructed in the course of human interchange. In this sense, relational responsibility is a result of interdependencies and connections between different actors and the entire focus shifts from individual selves to 'we' (McNamee & Gergen 1999). Since relational responsibility does not allow identifying a fixed locus of origin for what is the case (McNamee & Gergen 1999), one needs to address responsibility as a process and a particular type of engagement.

Preliminary attempts have already been made to apply the concept of relational responsibility to robotics. For example, it has been pointed out that responsibility should be understood not only in terms of responsibility *for* something but also *to* someone (Coeckelbergh 2016). In this sense, the emphasis is on the link between being responsible and being a person rather than only on what one does. Also, given a broad and inherently social frame of reference for relational responsibility, it has been argued that the concern about responsibility should inquire into the conditions that make responsible action and responsible practice possible (Coeckelbergh 2016).

13 Wikipedia: The trolley problem is a thought experiment in ethics. The general form of the problem is this:
You see a runaway trolley moving toward five tied-up (or otherwise incapacitated) people lying on the tracks. You are standing next to a lever that controls a switch. If you pull the lever, the trolley will be redirected onto a side track, and the five people on the main track will be saved. However, there is a single person lying on the side track. You have two options: Do nothing and allow the trolley to kill the five people on the main track. Pull the lever, diverting the trolley onto the side track where it will kill one person. Which is the more ethical option?

Others have emphasized the importance of taking responsibility in relation to new technologies 'as a society' (Ethics Task Force 2018) and as 'members of humanity' (Ethically Aligned Design, Version 2 (EADv2), 2017). Following the approach, according to which moral responsibility can never be allocated to autonomous technology (AI, Robotics and 'Autonomous' Systems, 2018), we propose here that relational responsibility frameworks apply to human actors only and they exclude robots as potential moral agents (this is also why the notion of 'relational responsibility' seems to be more adequate than 'networked responsibility'; the latter opens the door to inclusion of non-human agents).

One way to further develop frameworks for relational responsibility in robotics is to focus on the notion of 'dialogue' that would engage both robot makers and affected stakeholders. When addressed through the lens of relational responsibility, dialogue can be understood as a form of expression and engagement that may actually transform the actors involved in such a dialogue. Also, since the notion of relational responsibility here refers to moral responsibility in the first place, it implies the need for the development of the entire 'culture of responsibility' rather than only specific analytical methods and approaches.

In this final section we begin by developing our own solution as a suggestion to how robot makers can deal with the unresolved ethical issues around the inclusion of affected stakeholders found in REELER and presented in the remaining publication. We suggest that robot makers, as well as affected stakeholders, expand their understanding of ethics to include the relational character of how to act responsibly, and have dialogues, when developing distributed technologies. This means giving voice to people outside the 'inner circle' of robotics – and as already mentioned this is not an easy task. What is needed is people who are experts in facilitating alignment between stakeholders and robot developers to ensure relational responsibility. In REELER terms, *alignment* takes place through processes of *collaborative learning*.

In collaborative learning humans in a group cannot only learn from each other (collective learning), they can also make use of their diverse competences and thus divide the responsibilities among group members. Following Anne Edwards, different groups need to learn how to share underlying motives for pursuing a joint activity, but that does not mean they need to share their core expertises (Edwards 2010). Though collaborative learning requires all parties are equally engaged when working towards a common goal, they do not need to collectively share all the knowledge and skills needed as long as they can make use of each other's diverse expertises. The expertise of the robot developers includes their technical skills and disciplinary knowledge of design and development. The expertise of the affected stakeholders includes their situated knowledge of what matters in their everyday lives. In order to develop responsible and ethical robots together, they need the required relational agency to:

"Recognize the resources and motives that others bring to bear as they begin to interpret the common problem space (the object-motive) and 2) resourcefully participating in expanding the problem space by "aligning one's own responses to the newly enhanced interpretations with the responses being made by the other professionals while acting on the expanded object" (Edwards 2010, 14).

Following Edwards' definitions of relational agency, REELER proposes that collaborative learning can lead to relational agency, which builds on an evolving expertise where the engaged parties recognize "what others can offer a shared enterprise and why they offer it; and being able to work with what others offer while also making visible and accessible what matters for you" (Edwards 2012, 26).

However, we also recognize that this is easier said than done and thus call for a two-pronged strategy (*see 13.0 Conclusion*). On the one hand, the field of robot makers must develop a new roboethical culture that ensures robot developers take affected stakeholders into account. On the other hand, given the distributed character of robotics, the robot makers, and affected stakeholders' diverse understandings of ethics and the situated character of the expertises involved, REELER suggests a new profession of *Alignment Experts* (*see also 13.0 Conclusion*). Alignment experts must be able to help build relational responsibility. Such experts know alignment must entail a growing understandings of what matters for each group involved. The experts should be able to interpret challenges in alignment of practices and mediate a common knowledge made up of what matters for each collaborating group of people.

4.6 Concluding remarks on Ethics Beyond Safety

REELER's data indicates a need for an overall shift from safety-oriented ethics in robotics towards a more holistically oriented 'distributed ethics' and above all a 'relational responsibility' approach. A key premise is that ethical concerns arise together with, and not in addition to, engineering work and they evolve as robotics evolves.

By reading across the robot makers' statements about ethics in our 11 cases, some patterns emerge: 1) the robot makers mainly see ethics as tied to problem-solving and safety as well as keeping or making regulations and standards, and 2) many have high standards and want their robots to do good, but they do not have the affected stakeholder's perspective. The expressed understandings of ethics limit the robot makers' responsibilities to making inherently good robots, following regulations and making technically safe robots. Nonetheless, despite the European Union's emphasis on ethics, many robot developers still do not think ethics is tied to *their* technical work, and rather regard ethics as tied to human aspects, which are seen as separate from the technical aspects. When robot developers do not go beyond technical safety issues directly tied to their specific products, they tend not to consider, for instance, hacking or privacy issues as ethical aspects tied to their design decisions. The ability to develop a more holistic ethics is further challenged by the fact that robot components are developed by different people in different places, which makes is difficult to build a common ethical ground around the materiality and meaningful assembly of these parts. Moreover, ethical reflection and engagement with ethics are sometimes viewed as only optional in robotics R&D or 'placed' in special departments separate from the actual development process, typically in larger companies.

Another finding is that some robot developers, as well as company owners, view humans as an obstacle to the robots' productivity and efficiency – and that their emphasis is on how humans and environments have to change to accommodate robots. REELER's research entails identifying patterns across variation, while also highlighting particular isolated findings that have great impact or import in roboethics. So, even if only a fraction of robot makers view humans as an obstacle to the robot's efficiency, it is still an ethical challenge that has not been dealt with in the academic and political debates about robots and roboethics, unlike issues of autonomy and robot rights, for example.

Yet, it is also important to stress diversity in the robot makers' approaches to ethics. Some do go beyond the problem-solving approach into a more holistic way of perceiving robots; as a technology that affects people's lives.

As our research shows, today some humans are actively prevented from participating in these debates when, for instance, their employers forbid them to participate in REELER's research. We also see that a number of developers do not engage in this topic, because they do not care about these debates. Furthermore, public debates on how to deal with ethical issues may be misplaced partly due to the lack of ethnographic research into what actually matters for robot makers and affected stakeholders, how they learn in their everyday lives (*see 7.0 Learning in Practice*), but also on the prevalent misconception of robots as more than machines (*see 8.0 Imaginaries*). These gaps between the ethics needed in an everyday-life perspective and what is found at the table of ethical experts and robot makers may be closed by the type of 'holistic' education proposed in this chapter. But we also call for enhanced human proximity (*see 12.0 Human Proximity*). As a means to facilitate this proximity, and to help affected stakeholders' and robot makers' collective engagement in developing relational responsibility, we suggest alignment experts as a new type of intermediary.

Alignment experts could help affected stakeholders get access to the relatively small group of humans who today make ethical decisions and regulations, and thereby help to bridge affected stakeholders' perspectives on ethical issues of robotics with policymakers', philosophers', and engineers' ethical considerations. Another task of alignment experts could also be to work on the 'isolated' character of ethics. The overall purpose of alignment experts is to foster alignment of motives with the aim of building a new form of relational responsibility and agency in robotics.