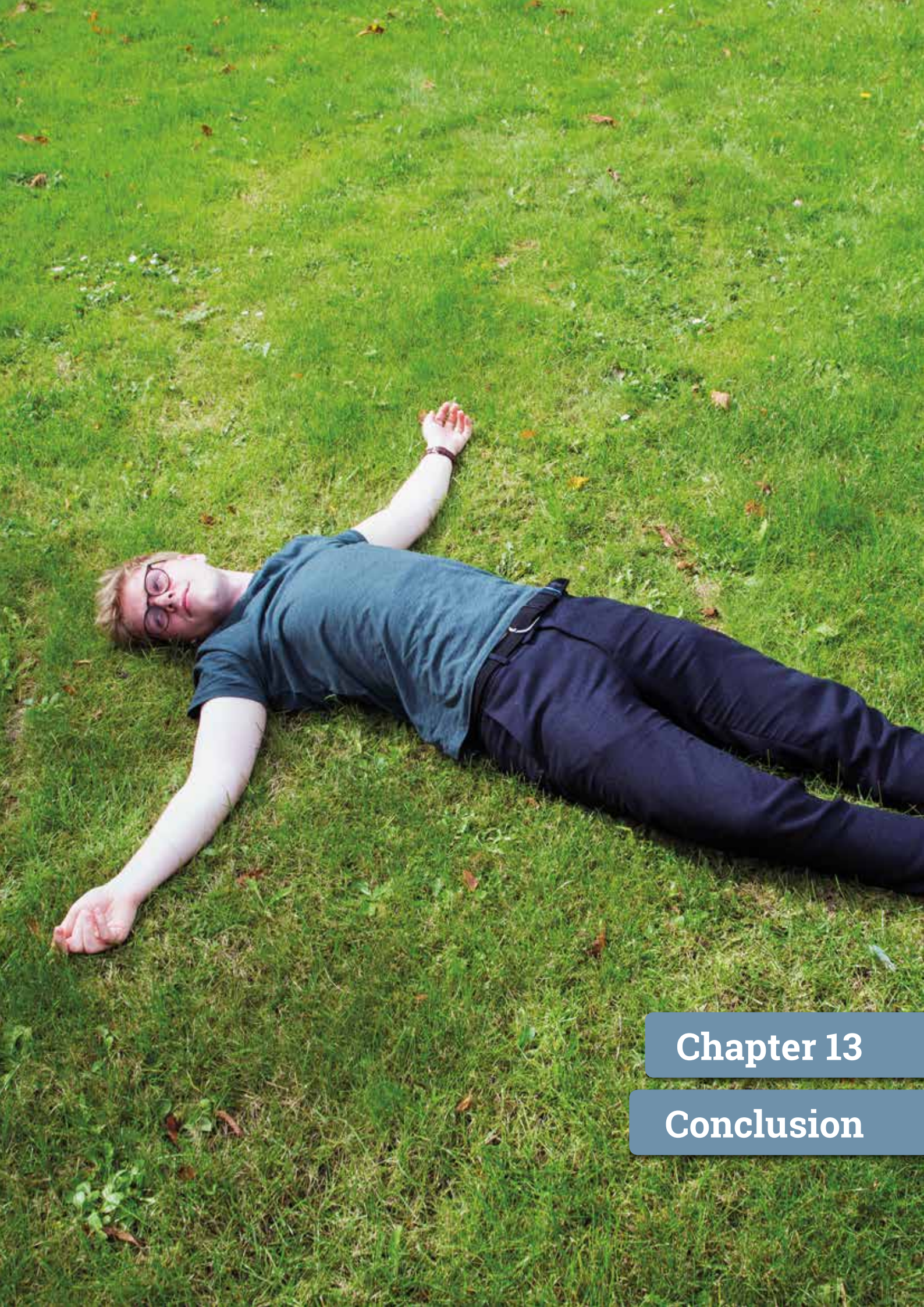

Chapter 13

Conclusion

I think a very, very important issue clearly is to have, in every large project, people from an interdisciplinary background because that's the way for problems [to get solved] which necessarily involve a human part, and a technological part. To get at least some idea of what to expect and foresee, you need people from really different areas.

(Jorge, head of research lab, robot maker, BUDDY)

13. Conclusion

What kind of future do we want to create with robots and artificial intelligence?

We are at the climax of the great grand narrative of technology. This narrative tells a story of constant progress and of robot technology relieving people of hard work, giving us free time to develop ourselves through meaningful tasks and new interesting work. As a whole, the development of innovative robot technology is not only necessary, but can also be a blessing for a society. However in order to create the future we want, we also need a story of pitfalls and realistic scenarios of how we can and ought to deal with robots and AI (artificial intelligence). In other words, we need a reality check on the narrative, on the storytellers - the robot makers, as well as the listeners - the affected stakeholders.

Human culture has always been defined by material tools. What may be different this time is that these tools, robots and AI, are developed in somewhat closed environments far from the realities where these technologies are going to be put to use. Furthermore the new type of innovative robots developed today, and studied by REELER, have moved from factories into the lives of people in hospitals, schools, construction sites, public streets and homes. This move calls for a new awareness of the ethical responsibilities that follow from robots engaging and entangling with people in their everyday life settings.

The protagonists of the grand narrative of technology can no longer just be developers, funding agencies and other robot makers but must include the end-users, as well as the overlooked directly and distantly affected stakeholders. *Perspectives on Robots* introduces new voices and serves as a reality check on imagined futures. Our research has focused on what we can do better to create ethical robots and AI that fit the different life-worlds of affected stakeholders.

In REELER we have undertaken one of the most comprehensive ethnographic studies of robotics in Europe ever. Our analysis runs across 11 cases, each representing a different type of robot and covering many sectors. We set out to explore general gaps between robot makers and affected stakeholders, scrutinize consequences of these, and develop new approaches to bridge them. We have presented these explorations in the previous chapters of *Perspectives on Robots*, as well as on the REELER Roadmap homepage (*www.responsiblerobotics.eu*), where it is also possible to access this publication's supplementary material [1] and the Online Interactive Toolbox (*www.responsiblerobotics.eu/toolbox*).

The fact that there are gaps between the realities of robot makers and affected stakeholders has been documented across the 11 cases in REELER.

The initial aim of our research work was to develop research-based tools to make robot developments more ethical and robot makers (i.e. robot developers/engineers, application experts, spokespersons and facilitators) more aware of affected stakeholders' needs. To accomplish this aim, the REELER Roadmap presents a two-pronged strategy for the future, which we propose to the European Commission.

1. Develop and disseminate tools that enhance *robot developers'* (engineers, mostly) awareness of what is to be gained from collaborating with and taking end-users and affected stakeholders' perspectives into account early on in the development phase.

2. Develop alignment experts as a new profession, where people are educated in methods of aligning view and visions of robot makers and, often unheard, affected stakeholders. Alignment experts can also give voice to distantly affected stakeholders, when relevant.[2]

This conclusion will summarize all the chapters in the full, online, version of *Perspectives on Robots*, including the three chapters *2.0 Robot Beginnings, 3.0 Collaboration in the Inner Circle* and *11.0 Gender Matters*, which are not included in the body of text in the printed version of the publication.

Perspectives on Robots consists of three parts:

1 Annex 1: REELER methodology. Annex 2: Excerpts from REELER's ethnographic data.

2 If alignment experts are to have real impact, the inner circle of robotics have to be convinced of their utility.

13.1 Summarizing Part One: Introducing the inner circle of robotics

1.0 Introduction, 2.0 Robot Beginnings, and *3.0 Collaboration in the Inner Circle*

In the early phases of our ethnographic fieldwork, REELER find that robot development is often distributed across different actors and organizations, and the person buying the robot may not be the same person using, encountering, and being affected by it. In response to this fact, REELER has developed a new vocabulary presented with the Human Proximity Model in *1.0 Introduction* as well as in *2.0 Robot Beginnings* and *3.0 Collaboration in the Inner Circle*.[3] The Human Proximity Model identifies different interest groups that either collaborate with, and learn from, each other or do not collaborate and learn from each other. Collaborations, or lack thereof, ethical consequences unfolded in PART TWO, which addresses robot developer's relational responsibilities. The HPM and its accompanying vocabulary opens the door to analytical discussions of the gaps REELER explore, and to new discussions of ethics regarding relational responsibility that will help ensure that robot makers conceive of and create more ethical robots.

13.2 Summarizing Part Two: Enhancing robot developer's awareness of affected stakeholders

4.0 Ethics Beyond Safety, 5.0 Inclusive Design, 6.0 Innovation Economics, and *7.0 Learning in Practice*.

Throughout our analysis we find that robot developers, who have generously shared with us their work and concerns for the past three years, care deeply about the quality of their robots, and are genuinely concerned with developing the best possible robot solutions. While many have little formal knowledge of ethics (*see 4.0 Ethics Beyond Safety*), they are often both interested in and care for the users of their robots. The majority of developers either work directly with end-users or listen to spokespersons and application experts speaking for end-users. In our general conversations and fieldworks, we see that developers are a very diverse group. In this diverse group, we also see some developers being less preoccupied with concerns for humans in their work. Here, humans and robots can be seen as dichotomies which involves choosing robots over humans, rather than combining the two.

Some robot developers also perceive robot buyers (who may never use the robot themselves) as end-users (who do use the robot) and many do not consider the potential added value of including directly affected stakeholders in their design work. Further, we see that directly affected stakeholders, as well as distantly affected stakeholders, are rarely given a voice in the activities and decisions taken in the inner circle of robotics. Thus, we conclude there is a need for this new vocabulary addressing ethics in terms of *relational responsibility* between robot makers and affected stakeholders. It is important to underline that this responsibility is not just a relation between end-users and robot developers, but involves all of the persons in the inner circle who are responsible for legislation and funding as well as all affected stakeholders. However, as there is also a clear power imbalance between these groups, we see the need for alignment experts to help giving voice to the affected stakeholders and translate their views into useful inputs in the debate.

5.0 Inclusive Design exemplifies some of the ethical issues arising from the closed collaborations, when affected stakeholders are not part of the group of collaborators and wider development decisions, and it suggests the need for new, grounded ways of thinking about users in relation to robots. This may, for instance, mean including consideration for not just end-users (like patients) but also directly affected stakeholders (like staff in a hospital) in decision-making processes. *6.0 Innovation Economics* discusses the importance of collaboration for innovation economics systems, which comprises multiple actors engaging in situated everyday practices to bring technological breakthroughs from the research laboratory to the market. Modern innovation economics distances itself from any linear, hierarchical, deterministic view. Rather, it frames technology development as taking place by knowledge-based collaborations of heterogeneous networks of entrepreneurs, research institutes, government, pressure groups, and other types of economic actors. Such innovation networks evolve endogenously over time, with autonomous actors entering, refocusing, and exiting, hereby also driven by emergence, maturation, transformation, and dissolution of their industries, etc. This understanding situates technological development within the social relations and activities of persons and organizations.

Likewise, *7.0 Learning in Practice* argues that by developing new ways of thinking and pursuing different, more *situated*, ways of knowing through education and through learning in situ (about users and robots in context), robot developers and affected stakeholders can achieve closer mutual proximity, and become much more aware of each other's sociomaterial worlds.

In all, Part Two points to how existing familiar collaborations, and the lack of stakeholder collaborations, can lead to exclusionary development processes – which may also hamper innovation as developers do not reap the full potential of including other perspectives in their design processes. In order to overcome this gap, we argue that robot developers may benefit from a relational expertise; learning what matters to others in collaboration toward a shared goal.

REELER has developed a number of experimental tools for exercising these perspective-taking skills. To help develop relational expertise in identified end-user/robot developer relations, we suggest both educational tools and the help from alignment experts. These may also help ensure that

3 Due to lack of space, we have not included these two chapters in the printed version of this text, but they can be found in the internet version.

Robot makers have the opportunity to craft a richer, more inclusive chapter on robotics in the grand narrative of technological progress. (Photo by Kate Davis)

the proper end-users and directly affected stakeholders are identified in the design process.

Thus, in part two we conclude that robot developers have a lot to gain from learning from end-users and affected stakeholders (possibly with the help of alignment experts). This awareness may be a road to more ethical and responsible learning in robotics that will hopefully lead to new and more productive innovation processes.

Some of the steps REELER argue are needed to better bridge the gap between affected stakeholders and robot makers fall, in some respects, outside the scope and responsibility of robot developers. This is most certainly the case when we look at the consequences for the broader group of distantly affected stakeholders, which part three zooms in on.

13.3 Summarizing Part Three: Expanding beyond the inner circle

8.0 Imaginaries, 9.0 Robotization of Work, 10.0 Meaningful Work, 11.0 Gender Matters, 12.0 Human Proximity, and *13.0 Conclusion.*

Not only robot developers can benefit from increased awareness of their relational responsibility. The REELER research show how robots will, and already do, affect society as a whole. In *8.0 Imaginaries*, we see that the way robots are represented by robot makers and application experts taps into a wider issue of how representations of robots in popular and news media affect the public, including policymakers. Here REELER calls for a reality check. Imagery of human-like 'intelligent' and 'autonomous' robots has ethical implications, as this imagery affects how European societies and politicians envision their robotic future. Chapter *9.0 Economics of Robotization* presents a large-scale discussion of the future of work, specifically addressing the expected economic impact of robotization including broad sectoral changes in employment. These impacts move far beyond the individual robot developer's or even the robot companies' ethical responsibility. Generally speaking, alignment between how robots are imagined in society and what robots can actually do is needed.

In *10.0 Meaningful Work*, we engage in a close-up discussion of the many qualitative transformations of work that robotization entails, sometimes with acknowledged benefits, other times resulting in an overall degradation of meaningful work. This chapter calls up contrasts between the values held by workers and the values inherent to the robotization of human labor. It points to a cultural gap that extends beyond robot developers to other robot makers and affected stakeholders, including employers, policymakers, labor unions, and educational institutions. Another matter which reaches beyond the responsibility and ability of the individual robot developer is the insular environments of technological developments. The fact that technological developments, like robots, are mainly driven by men with particular backgrounds and experiences, while the effects of these developments are felt by all is taken up in *11.0 Gender Matters*. In this chapter, we also present

issues of gender in design and robotics/engineering culture which, if left unchecked, may contribute to an inequitably gendered society.

REELER suggests, in Chapter *12.0 Human Proximity*, a new education of alignment experts who can help confront the above-mentioned challenges. Alignments experts could supplement the relational expertise of robot makers to ensure that affected stakeholders also take responsibility for their role in the situated implementation of robots.

As argued in these chapters, the robot makers, including many robot developers, write scripts without having a clear idea of who the end-users will be who are in the closest proximity to the robots in everyday work. Furthermore, the directly affected stakeholders are often not considered in these stories. They are, for instance, the nurses, the physiotherapists, the car mechanics or the school teachers, who will not be users of robots helping patients, driving cars or teaching math, but they will still need to accommodate and help implement the robots. The robot developers have a hard time understanding the messy social and material environments where their robots are to work, as their stories are often (if at all considered) written for neat and clean 'mock-ups' far from the reality on the shop-floor. And though it is not the responsibility of robot developers to ensure that distantly affected stakeholders get a new meaningful job or education (like fruit-pickers losing their job to a robot, or secretaries in need of reskilling), it could be seen as an overlooked responsibility of other robot makers (such as funding agencies and policymakers). All of these new ethical responsibilities also come with a need for someone to consider the long term potential positive and negative effects of the expensive robots developed: whether the robot is welcomed after a while of scepticism, or the innovation investment is lost because the robot is mothballed or sabotaged, or because, in a long term perspective, the robot changes environments in undesirable ways.

All story-telling is normative. As emphasised by innovation economy we need heterogeneity to ensure innovation. We need new voices in the narrative – and at the same time a more comprehensive and holistic view on why we develop robots and AI and for whom. This is not just to be more ethical, but also because a surprising number of the robots in REELER are to some extent founded on public funding (nine of eleven cases), and thus seem to a have a direct public responsibility. However, REELER research has also shown that collaboration in heterogeneous groups, and the alignment of different motives, can be a very difficult process. Even the identification of who to collaborate with (end-users, directly or distantly affected stakeholders) does not seem to be something that can be left to the robot makers to decide. Thus, we have developed a two-pronged strategy meant to steer the technological progress narrative toward a new, richer and more inclusive chapter on robotics.

13.4 The two-pronged strategy

The wider effects of robots explored in part three cannot be solved by robot makers or affected stakeholders alone. Across cases, from educational consumer robots to commercial service robots or industrial robots, we have found that the developed robots simultaneously in- and excludes people, put new demands on users and directly affected stakeholders, and change environments, habits and work routines. In economic terms, REELER has shown that collaborations, rather than linear models of innovation, lie at the heart of developmental processes. We have also shown that affected stakeholders are rarely included in these collaborations. Even if robot developers exercise relational expertise and engage in collaboration with affected stakeholders, a built-in asymmetry remains as these collaborations are likely to be initiated by the robot makers with focus on the robot developers' chief activity: robot development.

The original goal of the REELER project was to align robot makers' visions of a future with robots with empirically-based knowledge of human needs and societal concerns, through a new proximity-based human-machine ethics. We expected that by giving voice to those affected by robots, the project could propose ways to close the gap between robot makers and these affected stakeholders. To that end, we have developed the Human Proximity Model, written research publications, and produced a collection of tools for collaborative learning, including the board game BuildBot, the interactive serious puzzle game Brickster, and other tools for robot developers available in the online interactive toolbox. These tools constitute one pillar in our two-pronged strategy.

Though these tools are likely to raise awareness, they may not be able to proactively change existing circles of collaboration. Thus, if the sometimes diverging motives of affected stakeholders (sometimes, but not always, conflating with societal needs) and robot makers are to be aligned, we also need experts with a core expertise in aligning different motives across groups with different cultures, values, understandings, and (gendered, national, and economic) backgrounds.

Collaborative learning remains a key term, and as explored throughout this publication, the robot makers have many good motives for collaborating with each other. What we call for is collaboration between not only end-users and robot developers, but also collaborations with end-users and directly affected stakeholders, which the ethnographic research point to as potentially advantageous for robot developers. Yet, collaboration with end-users and directly affected stakeholders has also proven to be a minefield of time and money challenges for robot developers.

This is why we suggest a novel education/profession, *alignment experts*, who will take on the role as 'go-betweens' aligning the public and political expectations of a robotic future with the robots being developed. Since alignment experts are to fill the gap between the existing collaborations in the inner circle (with their spokespersons), and affected stakeholders,

they would need insights into both technological developmental processes and affected stakeholders' life worlds.

These experts will have the basic task of aligning the motives of the robot makers (including engineers, politicians, robot buyers, and funding agencies) with the real-life needs of end-users, and directly and distantly affected stakeholders. Alignment experts should be able to enhance relational expertise by ensuring more proximity between robot-makers and affected stakeholders, making all parties involved more aware of their own relational responsibility. This definition places the responsibility for learning about each other on both robot developers/makers and affected stakeholders/end-users, with the alignment experts acting as intermediaries ensuring human proximity and alignment of robot functions and applications with human needs and societal concerns.

One the one hand, our research recommend that alignment experts are independent of the inner circle as they must be free of various interests in funding schemes and regulations (contrary to the application experts studied in REELER). On the other hand, the strong culture of engineering within the inner circle of robotics suggests that for alignment experts to be successful, their competences and work methods must be accepted by the (powerful) people engaged in robotic business. This point is partly tied to the ways robot developers typically get and develop their ideas. The catalyzing ideas that initiate projects often come from environments familiar to the robot developers; from technological developments in the field of robotics, or from answer to demands from customers/companies, policymakers, and funding agencies. That is, the design and development of a product is likely to primarily be initiated by the company deciding to develop it. Alignment experts are, however, expected to be able to point to societal needs, to which the robot developers present a technical solution Alignment experts can then be useful in exploring whether the technical solution matches the needs of the affected stakeholders. In that respect, the REELER research also points to a potential for new robot ideas increasingly coming from end-users or other affected stakeholders through alignment experts.

In addition to translating (societal) needs into potential robot ideas, alignment experts should also feed into the debates on ethics to ensure that the political and academic discussions are relevant to and take affected stakeholders' perspectives into account. The profession would entail in-depth studies of what matters to those affected by robots, how to avoid pitfalls stemming from normative thinking, and which types of situated knowledge could be the basis of the new educations needed in a robotic society.

Our main conclusion is therefore:

In order to ensure ethical and responsible robot design, it is essential to work on a two-pronged strategy which entails:

a) enhancing robot developer's awareness of the group of affected stakeholders

b) aligning robot makers' and affected stakeholders' motives by increasing human proximity through the involvement of alignment experts, for effective collaborative learning.

This will ensure a reality check on both robot markers perceptions of stakeholders everyday lives and stakeholders perceptions robots – and thus a reality check on our shared future.

Thus, we do not see a future where enhanced ethical awareness is the sole responsibility of robot developers – but a future of relational responsibility that involve all stakeholders helped by, among others, alignment experts. In this way we can begin a new chapter in the great narrative of how technology, like robot and AI, can shape a brighter future for us all.