Report from workshop in Utrecht 16-17 Feb 2014: Making robotic autonomy through science and law?

Authors: Kjetil Rommetveit, Kristrún Gunnarsdóttir, Niels van Dijk, Martijntje Smits

Introduction

This document reports on the Epinet workshop on the making of robot autonomy, held in Utrecht 16-17 February 2014. The workshop was part of a case study focused on developments in this area, in particular, autonomy for assistive robots in care and companionship roles. Our participants were of relevant expertise and professional experience: law and ethics, academic and industry robotics, vision assessment and science and technology studies (STS). The workshop was intended to explore the *expectations of robot autonomy* amongst our participants, against a backdrop of recent policy views and research trends that are openly pushing an agenda of "smarter", more dynamic and more autonomous systems (e.g. European Commission, 2008[1]; EUROP, 2009[2]; Robot Companions for Citizens, 2012[3]). In short, future robots should be designed for action in the home, in the streets, in care, at work and any other setting in which they can operate alongside humans and on behalf of them in assistive roles. Future robots should help address the grand societal challenges in Europe, in particular, those of an ageing population, sustainable healthcare and welfare.

Robotics development is intimately connected with visions of robot autonomy, however, as a practical achievement, robot autonomy remains till this day part real, part promise. Visions of robot autonomy are in part based in fantastical depictions, in particular, any human-like intelligent appearances like those found in mythology, folklore and the science fiction genre (e.g. Hephaestus' golden assistants; the Golem; Asimov, 2004;[4] Pixar Animation Studios, 2008 [5]). Ideas of robot autonomy are nevertheless powerful societally and culturally-specific visions, even if the very notion of "autonomy" is vague and inconsistent in recent accounts of future robots. These accounts still come together with considerable force in directing the efforts of researchers and experimenters, for example, in establishing funding priorities. They have a function in strategic planning for future developments and, thereby, for attending to particular scientific and technological problem domains. Accounts of future robots are also informing and shaping the efforts of legislators, ethicists and lawyers. To that effect, one can say that there is an *official vision* of future robots, a yardstick with which everyone implicated in robotics development has to measure their expectations.

---

1 European Commission (2008).*ICT Research. The Policy Perspective: European robots, getting smarter, safer, and more sensitive*. European Commission.

2 EUROP (2009).*Robotics Visions to 2020 and Beyond - The Strategic Research Agenda for Robotics in Europe, 07/2009*. Coordination Action for Robotics in Europe (CARE).

3 Robot Companions for Citizens (2012).*MANIFESTO. More than machines*. The RCC consortium. http://www.robotcompanions.eu

4 Asimov, I. (2004). *Robot Visions (short stories and essays, 1940-1989)*. New York: Roc.

5 Pixar Animation Studios (2008). *WALL · E: Computer-animated romance science fiction film*. Walt Disney Pictures.

With these considerations in mind, our workshop proceeded by addressing thematically defined areas of interest to our case study:

- Robot autonomy: which perspectives are endorsed to shape and develop the [*official*] vision?
- Robot autonomy: which perspectives shape ethical and legal frameworks?
- Future robotics: who should participate, and which knowledge-domains should be included to shape policy visions?

In the following, we report and expand on our analysis of presentations and discussions during this two-day event. Among other things, we observe how the *official vision* is challenged, what the concrete ethical and legal issues are and what lies ahead for interdisciplinary collaborations in this field.[6]

First day – Robotics

Which perspectives are endorsed to shape and develop the [*official*] **vision?**

Three questions were circulated beforehand to a group of roboticists who were invited to address them one at a time in brief presentations, followed by rounds of questioning and discussion.

- How do you perceive of the vision to create autonomy in robotics systems used for care and companionship?
- What purposes will be served by pursuing this vision?
- Which are the main professional challenges faced by yourself in engaging with the vision to create more autonomous systems for care and companionship?

The three questions were intended to hone in on the ways in which roboticists, in their everyday practice, relate to, reason about and operationalise robot autonomy. The main findings presented here are grouped thematically into three sections, each addressing a set of concerns about the relationship between robotics and wider

[6] The report is based on written notes from the workshop, and so is paraphrasing most of the arguments and discussions. Direct quotes are used in a few instances where text has been taken from presentation documents. This is indicated accordingly.

society, including ethics and law. Previous assessments of the making of robot autonomy – by Epinet partners – also feature into the accounts we give of the events on the first day, i.e., in our own reflections on the relationship between the two.

1. Roboticists relate strategically to robot autonomy in their daily work, be that in teaching, research or clinical care work.

It should come as no surprise that presentations reflecting on what to expect of robot autonomy were quite different one from the other. Roboticists ground their approach in their respective research speciality, professional background and experience, research aims and ongoing projects. They all spoke of 'intelligence', 'adaptability', 'learning' and other manifestations of machine 'smartness' and potentially greater autonomy, but each had different concerns to share, emphasis, and so on.

We provide an account of these reflections, ordered under three headings: *clinical work*, *basic research*, and *experimentation*, each of which is chosen pragmatically and informed by self-identification and the focus of the work of these roboticists.

a. Clinical work.

Two of the roboticists work in clinical and therapeutic settings, and this background serves to structure the way they imagine and describe autonomy.

Roboticist 1 works with therapeutic robots, especially for children with autism, but also elderly sufferers of dementia. He described autonomy as relational, and as a bit hybrid. It is partly about control, partly like magic or make-believe, like the *Wizard of Oz*. Hence, the user experiences the robot as autonomous, but the autonomy is only partial and always a function of relations: The researcher is never really out of the loop but remains ‘hidden in the closet’ during the treatment sessions.

The robot’s appearance is important to the user’s experience of the robot as autonomous.[7] The user must also like the robot, and for this purpose the system itself must be adaptable. *Roboticist 1* emphasised different degrees of adaptability, referring to both humans and robots. This is expressed in ways in which the patient-robot relation may dynamically evolve: the human adapts and/or the robot adapts, thus they enter into a process of mediation with uncertain outcomes. This process touches upon core value judgements, such as relations with other humans, religion, righteousness,

[7] What Epinet researchers have termed the *make-believe*.

and more, some of which is illustrated, for example, in the movie *Robot & Frank* (2012), where a robot is turned into a thief by its owner.

Roboticist 4 (electrical engineer, roboticist) works with children recovering from traumatic brain disease and injury. To him, autonomy appears as what needs to be administered to the kids. Similar to *Roboticist 1*, the concept of autonomy is described by *Roboticist 4* as thoroughly relational: The autonomy of a robot is described as the capacity to evolve and adapt to the needs of individual users in order to increase its effectiveness and improve the engagements with it. For instance, the robot may remain active or calm, depending on the need of a child (patient). This, however, presupposes that robotic devices have access to data on the children (patients) they interact with and built-in capabilities to act upon such data in pre-defined ways. In this way, the autonomy of the system hinges on the amount of data there is, on processing capability and processing desirability.

Roboticist 4 specifically described his work as 'applied science' where the main challenge was not so much the development of new robotics technologies but rather the use of existing ones: the technology needs to be affordable and working, rather than novel and innovative. They use Google and smart phones, integrating standard elements into the larger systems that have been compiled for therapeutic purposes. Hence, the definition of autonomy is not only a function of the relation with the patient but also of the state of the art in robotics (and related technologies) working together in real-life settings.

b. Basic research.

Two roboticists could be described as working on basic research, although, they engage with real-world robotics and real-world problems. Rather than attending to the specific needs of some user category, they orient themselves toward general research problems and principles, the making of 'enabling technologies', potentially useful in all types of robotic applications. Their notion of autonomy is intertwined with these orientations to robotics development.

Roboticist 2 situated the questions of autonomy within the philosophy guiding his research group. Seeing robotics as an integration of science, engineering and society, they define autonomy in terms of rational, social and intelligent behaviour. Central to this is the question of how the robot performs affect, emotion and reasoning: how can the robot be capable of reading the actions of humans, and translate into actuation of one action among possible options? If so, can robots then be *intentional* agents?

One example he took was how a robot develops a set of rules based on body scheme generalisations, i.e., abstract knowledge about the behaviours of the human/user. Robot rationality could be derived from learning from a mapping of successes and failures, involving the capacity of the robot to generate from that new paths of behaviour. Robot sociability could be understood as the ability to perform (autonomously) culturally-specific tasks, such as the etiquette of handing over a knife, with the handle facing the recipient. *Roboticist 2* spoke in this respect of control (low) level, signal level and a higher signal processing level, the key complications emerging when translating between them.

Roboticist 3 started out by saying that he has been trying to explain to his students what autonomy is, but it turns out to be very difficult. He tried to ground the question of autonomy in the (broader) question of 'what is a robot', therefore, proposing to define autonomous robots as intelligent machines capable of certain degrees of learning and interaction. Insofar as this is the case, robots may not be completely different from humans. According to him humans can be seen as intelligent machines and robots can be seen as electro-mechanical machines. This he described with historical examples of automatons (e.g., the automated rooster, flute player, the Karakuri doll), and electromechanical automatons through to self-parking cars. Referring to his work at Waseda, *Roboticist 3* used the example of robot-assisted medical training systems possessing a kind of intelligence. Other examples he took, include different kinds of assistive robotics with cognitive and perceptual capabilities, or bodily-kinaesthetic capabilities. A fundamental problem with all these systems, according to *Roboticist 3*, is a persistent conflict. When machine intelligence goes up user friendliness goes down. Autonomous robots may approach one or another form of 'intelligence' but this comes at the cost of that intelligence not being well suited for human-machine interactions. On the one hand, robot intelligence is seen as similar in principle to human intelligence. On the other hand, there is a problem in building such intelligence into functional robotic applications. The question therefore, is whether it is only a matter of time before science and technology (such as AI) can provide better solutions to overcome this problem or whether the conflation of human and machine agency, as essentially of the same nature, is a blind alley.

c. Experimentation.

Roboticist 5 was the only participant to explicitly address the question of autonomous robots as a vision for policy, which he saw as politically over-determined and flawed. He also addressed an alternative vision of autonomy, attempting to take into account both the technical and ethical aspects. For this reason we have termed his contribution *experimental*, since it did not fit easily into a specialised robotics area. Seeing the

political *official* vision as erroneous, he advocated scaled-down expectations of autonomy. According to *Roboticist 5*, semi-autonomous robots could be hugely beneficial if very carefully and ethically developed. The minimal autonomy is defined as the ability to decide what to do next without human intervention. What we rather need is relative, controlled autonomy. Controlled autonomy is still commanded by humans. This notion he further expounded with reference to the concept of dynamic autonomy which is based upon Sheridan's (1980) scale of 10 gradual degrees of autonomy (from full input from a human to no input). According to this presenter, it is hard to conceive of applications with which we would want to relinquish all human control or intervention. A robot should at least be smart enough to figure out when it is in trouble, and so call for humans to respond. Different degrees of autonomy have also been coupled with notions of ethical responsibility to categorize what could possibly constitute ethical robotic agents.[8]

From these different accounts we learn that roboticists qualify autonomy in ways that can be rendered functional within their professional practices, experiences and projects. Their views do not resemble any grand (political) vision. Rather, they set out goals and purposes for robots that actually converge to some degree, as we will now turn to.

2. Envisioning *aims* and *purposes* and encountering problems.

It appears that the problem-domains encountered by roboticists vary significantly. Those working in clinical setting are immediately confronted with ethical problems and social-psychological considerations. Basic research is preoccupied with the particularities of technical obstacles and opportunities. And, the experimental approach to autonomy lacks sufficient institutional support to move forward. But, it should be underscored that we encountered considerable convergence of aims and purposes. This is exemplified by the ways in which the roboticists come up with similar lists of goals, including (in no particular order):

- To **enable and promote dignity, security and quality of life** (e.g. by robots taking care of certain hygiene tasks)
- To **enable elderly persons to live independently**, care for themselves and stay at home
- To **reduce costs and increase savings in the care services**

[8] See for instance https://philosophynow.org/issues/72/Four_Kinds_of_Ethical_Robots

- To **free up care and other professional capacity** by relieving personnel of routine tasks
- A kind of meta-distinction underscored by several participants was the question of whether robots in new situations and environments will cause isolation and loneliness, or whether they can support sociability and mobility. This is a valid concern that can be taken into account in assessing robotic applications in practice, and it points to the value of **supporting sociability and connecting people.**

Pursuing these purposes gives rise to particular practical, technical and ethical problems. Participants discussed scientific and technological problems. They spoke of robotics as an 'enabling technology', however, in many cases also linked with ethically-relevant considerations. According to the distinction between clinical and basic research, roboticists working in basic research emphasised the technical and scientific difficulties, while clinical practitioners were more sensitive and outspoken of ethically-relevant problems. Relations with publics on such issues were seen as important, yet difficult. As one participant put it, robot autonomy scares people. Yet, participants in the workshop shared perceptions of what the problematic issues are, even if they see and articulate them differently. All were in favour of more ethical investigation and debate relating to robot autonomy, and all were, in principle, in favour of improved communication with publics. Most (or all) also seemed positive about the Responsible Research and Innovation (RRI) programme of the European Commission but were not necessarily clear on what it entails in practice (most of the roboticists only heard about RRI from Epinet researchers or the ethicists/legal scholars who were present on the day).

a. Technical problems.

Roboticist 3 drew attention to the problem of building 'smarter', more autonomous systems which then are not user-friendly. Similarly, the requirement of affective and emotionally intelligent interaction, introduces endless difficulties. A possible solution, according to *Roboticist 3*, would be strategic research efforts across the human sciences and artificial intelligence (AI). He argued that research in the human sciences – here meaning AI – is trying to abstract human intelligence into encoded machine functions. But, this approach was critiqued (by *Epinet researcher 1* and *Ethicist 2*) for conflating human and machine agency, and for not being specific enough about how they each can and cannot be considered autonomous. There is no solution on the

horizon for overcoming the problem of how to build human-like agencies into machines.

On somewhat similar terms, *Roboticist 2* appealed to the higher functions of being rational, social and intelligent, as the ultimate source of purpose and aim. But, the project to construct a translation between the two in robotic systems has not been achieved so far, as pointed out by *Epinet researcher 1*. Even if we take it for granted that one or another higher-order purpose can be built into a robot, there is still the problem of translating purpose into lower-level tasks for execution in specific settings. According to *Roboticist 2*, in moving from the high level to the low, there are so many constraints, including also emotional constraints. According to him, this complexity has not been resolved in machine execution.

Finally, *Roboticist 5*, among others, underscored a central problem with more automated operations in human-centred environments, namely the unstructured and complex character of such environments as opposed to the structured confinements for which earlier generations of robots were built (and still are in manufacture). He emphasised how safety problems emerge in unpredictable environments, and how any human environment is per definition unpredictable.

b. Ethical issues.

Ethical problems pertain to *adaptability* and *learning*, and how human-robot communication is mediated. According to *Roboticist 1* this can be seen as the downside of adaptability: where do the dynamics and emergent properties take the user/the system? He specifically underscored how adaptability and adaptiveness may lead to conflicts with moral standards and belief systems. As an example, the question was posed whether, for instance, a robot could convert to [some religion] from whatever it has been taught. According to him, adaptiveness is a key issue with the highest ethical stakes.

Roboticist 2 highlighted how robots can destroy ecosystems of care. In this way the technology can create more problems than it solves. In connection with that, *Roboticist 5* argued that we should get rid of the fetish that technology is the only answer to human and social problems.

Several participants also mentioned the problem of dependency and the formation of emotional ties with robots. *Roboticist 1* mentioned the example of the little dinosaur (a robot) – and children learning to walk. His daughter developed a relationship with the dinosaur. He posed the question of whether such a relationship entails a loss of

relations with other children? He also mentioned how such questions become more pressing as systems become more advanced.

Problems were further raised about dependency on standards and already existing technologies such as Google and smart phones, as well as the connection with military research. Following the presentation by *Roboticist 4* in which autonomy was tightly connected to information gathered about the needs of patients, *Ethicist 2* asked whether, in the case of relying on Google, it would not be better to restrict the autonomy of the system. The implication was that Google has significant problems, especially relating to privacy, but there is also the more general issue of who is in control of any given system. The question then becomes what we mean by autonomy in systems whose functions depend on global conglomerates. Following up on this argument, *Roboticist 3* noted how Google bought Boston Dynamics, and how there are close relations between Google and military robotics. There are many problems relating to these associations. As for the use of 'standard elements', for example, Aldebran robotics (US) customize robots to look for information. *Roboticist 3* then asked what kind of information exactly these robots are looking for.

At a later stage in the workshop, *Roboticist 4* commented directly upon this problem of robots harvesting information, specifically as it pertains to the privacy of patients. Since all the robots in question are connected to the internet, that's a lot of medical information being aggregated and disseminated online. This raises serious privacy challenges for *Roboticist 4* and his team.

Overall, throughout these discussions, privacy and data protection kept coming up as a concern for most participants.

c. Coding ethics into robots.

During the discussions someone brought up the idea of encoding ethics into robots, akin to Arkin's proposal for military robots and Asimov's three laws. Although such ideas have been around for a long time, the view was expressed that this is now becoming a pressing need due to recent advances. But, unless such encoding is kept very simple, designing reliable ethical conduct into robot behaviour is likely to be even more complex than that of translating higher-order aims into lower-order commands for robot executions. Also, as asked by *Epinet researcher 5*: who should encode the norms into the robot? How will the norms apply in different domains of society, granted that different domains display very different logics and mores?

3. Assessing the *official* vision.

In posing our first question, we invited the roboticists to position themselves in relation to a vision of robot autonomy, e.g., the one promoted by the Horizon 2020 programme and the public-private partnership (PPP), RockEU. As we have seen, most roboticists did not do so directly. Rather, they relate to the question of robot autonomy with reference to their research environments and everyday practice.

Participants confirm that there is no single idea, nor a simple concept of 'autonomy' for robotic systems, corresponding with the *official vision*. For example, *Roboticist 2* talked about how there are different notions of autonomy, and how these are mostly implicit rather than explicit. *Roboticist 5* stated that there is no one vision of autonomy for robotics systems. He even sees an extraordinary level of confusion, and especially so in the contexts of care. He expressed grave worries about the situation in Europe, arguing that European policy makers have come to regard robotics as a solution to ageing, which he described as a disaster based in erroneous ideas of what robotics can and cannot deliver.

The general view was that the problematic aspects of robotics development lack a proper forum for discussion and problem resolution. *Roboticist 3* drew our attention to what he referred to as a huge discussion in the EU on autonomous robots for healthcare and home care. He underscored that there are problems of communication, however, he did not explain in great detail what those problems consist of. To paraphrase his main point, it is difficult to translate – carefully and ethically – how to explain the pressing ethical and social problems to those in power, mainly business-people and politicians.

These and similar accounts we heard, are indicative of the vagueness of what robot autonomy can stand for as a practical achievement—particularly of the purposes to which to put future robots. Even if the *official vision* describes 'smarter', more dynamic and more autonomous systems, there is nothing to indicate that roboticists have a 'master' idea of which technical and scientific problems need solving to achieve a particular version of robot autonomy for the purposes intended by policy-developers and politicians. Rather, it appears that various kinds of autonomy are implicitly understood amongst roboticists with no obvious link to the *official* vision.

We learn from the roboticists that visions of robot autonomy are primarily political, often driven by influences roboticists have little if any control over, in their own view. Working to resolve the practical and concrete problem domains is common practice, of which researchers can take some ownership. Other problems (ethical and socio-cultural) beyond the influence roboticists can have, are perhaps better addressed in ethical and (possibly) public discussions. Some of our participants also described

important developments being beyond their reach, such as the push for greater autonomy to replace human labourers which then results in considerable unemployment. When an Epinet researcher raised the question of unemployment, we were reminded of a whole history of labour redundancies in a number of application domains.

Roboticist 3 replied that the problem of human replacement and unemployment is beyond the control and influence of roboticists. He illustrated this with the case of Swedish forestry work which is highly specialised, yet the country does not have the human resources to fill the jobs. This is where robotics can help, regardless of unemployment. He further argued that other options may address an underlying problem of unemployment, however, such options are typically a matter of political decision-making, 'out of our hands and conditions we cannot change'.

Setting aside issues of labour redundancies and unemployment, participants confirmed the need to go along with certain aspects of what is expected of robot autonomy in recent times. This has to do with changes in research funding priorities, and innovation policy:

Roboticist 5 told of a need to be pragmatic, since he has to get grants. He then mentioned a tension in the new PPP, where there will be 50% private funding, and so changing the conditions for doing research compared to the FP7 funding regime which was 100% publicly funded. He also noted how the Commission insists on RRI being hard-wired into project designs with the necessary inclusion of ethical, legal and societal expertise. The implication is that we will see all sorts of tensions arise between the different actors who are now involved in taking the research agendas forward.

Summing up.

We learn that our participants shy away from directly addressing the question of autonomy as it appears in European policy visions and strategic planning of future robots. Rather, they relate strategically to robot autonomy on their own terms with reference to their respective professional backgrounds, experiences and experimental orientations. This way, roboticists render robot autonomy manageable. In some instances the problem-domains are very remote from dealings with legal and ethical complications. It is not as if the roboticists are insensitive to the character of legal, ethical and societal issues, problem-finding and resolution, but they feel that they lack the specific means and institutional support, perhaps also the public trust, to effectively address these kinds of complications.

As can be seen from the exchanges we report on here (although, not exhaustive of all the issues that came up on the day), there was a general willingness, even a sense of urgency, to discuss how to approach ethical considerations. Many of the roboticists already have experience from engaging with ethicists (and ethics committees) and so are not strangers to many of the ethical and legal complications that can emerge in the making of robot autonomy. Yet, as for direct relations with ethics and law we observed during the workshop how only one of the roboticists embarked on the difficult task of cutting across disciplines, i.e., *Roboticist 5*, with his elaboration of relative autonomy, controlled autonomy and dynamic autonomy in the creation of assistive robots for sensitive care settings. His efforts were well received, for example by *Lawyer 2* who stated how she appreciated the notion of dynamic autonomy, because it can be constructively dealt with by lawyers.

There appeared to be some feeling of insufficiency however, over communications with publics and 'society'. For example, many roboticists have to give serious thought to social acceptance and, among other things, they fear the stigmatization of machines. Many also described a wish to communicate with wider society, going beyond the mere inclusion of immediate users. There seem to be an issue here of establishing a legitimate venue for concrete dialogue and interaction over ethical, social and public concerns.

Finally, politics was treated with some suspicion by the participants, i.e., any evidence of imposition and demand that cannot be scientifically attained. The general idea is that publics should be included in decisions on purposes and aims and on user scenarios and application domains, although the available means of communication could be better known, better adjusted and more effectively used.

Second Day – ethics, law and governance

On the second day, two questions were addressed by ethicists, legal scholars and others with significant investment in ELS issues:

Question 1: How do ethics and law engage with efforts to increase the autonomy of robotic systems for use in care and companionship?

Question 2: what are the main professional challenges faced by yourself in engaging with the creation of autonomous systems?

The session which took up most of the day, also included contributions from a representative of the recent PPP initiative, euRobotics, whose job is to accommodate ELS issues within the new RD&I platform. In the following we elaborate and reflect on the main findings and then move on to the 'Final round of discussion' among all participants in this workshop.

1. Autonomy is an ethical and philosophical term with a long history in philosophical treatises, for example, of free will, the good life and personal responsibility. The term is also implicit in legal frameworks, where it is qualified in relation to agency and legal responsibilities.

The ethicists and legal scholars were clearly used to communicating with roboticists. All of them were well aware that the others use and rely on different types of autonomy. As could be expected, the scholarly strategies to ascertain or depict autonomy also differ within ethics and law, and the contributions on the day allowed for clarifications of those differences among the participants.

Ethics.

Roboethics is now a burgeoning field of research which incorporates a variety of ethics traditions.[9] The variety of approach however, does not merely reflect upon different ethical or philosophical traditions, but also the concerns and preoccupations emerging in their working relations with roboticists and others (lawyers, users, policy

[9] A preliminary mapping revealed (at least) the following disciplines and perspectives to be relevant: Rights and principles based ethics, everyday ethics, pragmatism, (social) psychology, HCI, HRI, hermeneutics, virtue ethics, Kantianism, philosophy of mind, artificial intelligence, capabilities approach, vision assessment, phenomenology, feminist studies, disabilities studies, science and technology studies, animal rights, human rights law, privacy scholarship, engineering ethics, bioethics, RRI, machine ethics, robo-ethics, participatory TA, foresight.

makers, etc.). The topic of autonomy (understood as self-determination) is much closer to ethicists' most pressing concerns with the use of the term, than to those of roboticists. But, as the ethicists went on to depict and explain autonomy, they did not spend much time defining it, except in relation to specific threats or challenges to it. In short, the participants orient to various relational aspects of autonomy, although, quite differently from the roboticists.

Ethicist 1 was concerned about the different conceptions of autonomy emanating from robotics / robot designs, and ethical / philosophical / legal scholarships. Based in the euRobotics Green Paper (and Latombe 1991), he defined autonomy simply as the capability for the 'setting of one's own goals', however, that is not a realistic design option. *Ethicist 1* argued that roboticists design what are just tasks, not the setting of goals. Robots capable of setting their own goals would be a contradiction in terms. Autonomy in robotics is always derivative and always handed down from human planning and design. As he explained, the contradiction of goal-setting robots appears in different guises which renders difficult communication between ethicists, lawyers and roboticists. For example, in the work leading up to the Green Paper, the main problem was for the lawyers to understand exactly what the roboticists were conveying about autonomy and related issues. This contradiction will also resonate in relationships between (autonomous) robots and humans. There are fundamental differences between the actions and relations of robots and humans, which *Ethicist 1* illustrated with movie clips, showing degrees of what appears autonomous. Speaking about deception, he explained that autonomy is not merely a matter of authenticity. There is a machine triggering a kind of response/action. It is more subtle than that – perhaps, for any given application we should compare what we used to do and what we now can do with the machine, keeping in mind that the complexity of human relations is not a real relational task for a robot. Humans and robots are different in kind, but this difference tends to be ignored due to the similarities in many individual tasks humans and robots undertake. Consequently, the gradual expansion of robots throughout society, coupled with the propensity to anthropomorphize them,[10] may pose a danger to our perceptions of human autonomy and relations, and of robots in society. On a different note, robots capable of action in unstructured environments will scare people so, for such purposes, great care needs to be taken. To summarize, a key message from *Ethicist 1* was that autonomy generally means something like *free will* among publics, philosophers and lawyers. We therefore need to communicate very carefully in talking and discussing autonomous robots who actually have no free will.

[10] *Roboticist 2*: from where comes the humanness?
Ethicist 1: from the delegation of human tasks to machines.

Ethicist 2 did not proceed from a classical ethics-based understanding of autonomy. She proceeded, partially at least, on the basis of the Roboethics Roadmap in which autonomy is defined in a way that is quite straightforward for roboticists to understand, i.e., autonomy as the capacity to learn. Learning is a prerequisite for social / assistive robots to work *socially*, but we also see a certain readiness—that people can increase their capabilities in and through the use of assistive robots, i.e., if implemented wisely. According to *ethicist 2*, machines mediate, and this is how their autonomy (learning capacities) are perceived. Also, inspired by works in science and technology studies (STS) she rejected classical notions of individual agency as the basis for assessing human autonomy. Rather, autonomy needs to be seen as a capability of peoples in communities.[11] The FP6 project, Cogniron, was mentioned here as an example of such an approach. However, autonomy in robots raises other challenges, mainly due to their unpredictable behaviours. *Ethicist 2* argued that unpredictability leads to gaps in claims to liability. Adequate laws are important here but not sufficient. There is also a moral responsibility to establish throughout a whole chain of responsibilities. Referring to the works of Sheila Jasanoff (e.g. on the Bhopal disaster), she described how efforts to rectify and compensate for damages induced by technoscience, raises meta-questions about the overall design of our societies. An adequate learning process here for society on the whole, would consist in efforts to restore the situation to its previous state, while learning how an innovation produced the disaster in the first place. In short, seeking to increase the autonomy in robotic systems with learning robots, will have to approach the community's collective capacity to deliberate and accommodate in the best possible way, both the dangers and the potentials.

Ethicist 3 briefly argued that autonomy implies being a master of your own destiny. The notion was not further expounded upon but used to pose the question of how robots affect perceptions of human autonomy. This is explicitly a human-centric approach insofar as the target of the question is the implications for humans (users and others as well as creators and manufacturers). *Ethicist 3* did not treat the definition or exact meaning of autonomy, but left it implicit in his presentation, while using the term to point back (reflexively) at the people making and using robots rather than at the robots themselves. In this account, with cues taken from hermeneutics and phenomenology, autonomy is relational insofar as the mediation of abilities at the

[11] In our interpretation, *ethicist's 2* position was based in a mix of communitarian ethics, robotics and science and technology studies. Neither of these exclude a strong position on autonomy, but require that it is qualified in terms of community/social relations and the ways in which these are mediated through machines.

very interface with the technology, remains of central importance. The same goes for the relevancy-constraints in which the question of autonomy arises and the ethical issues associated with it. *Ethicist 3* claimed that we need to pose the question if the purpose is to 'unveil the secrets of embodied intelligence'; to help people; to help helpers; or if the purpose is to solve societal and economic problems. Against this background, he outlined a number of ethical problems that relate to robot and human autonomy, but that are not limited to it. For example, western healthcare already has a number of problems without robots. Loneliness and a lack of self-mastery is a problem among the elderly and the quality and ethics of care are faulty. The quality and ethics of human-robot relations are uncertain, as well as the relationship between device-use and identity. Automation continues to come at the loss of skill and human dignity is under pressure, with or without machines.

Law.

Topics raised by the two legal scholars on the day, included questions such as: do we need to change existing laws and principles to accommodate for autonomous robotics, or can we adapt existing principles to the new cases? What about relations with other disciplines and with governance? Will law have to depend on them to a greater degree? What about the autonomy of *law*? Both presenters remained ambivalent. On the one hand, they were trying to accommodate for new challenges to law. On the other hand, they were trying to stick with the inherent positivism and conservative aspects of law. This could be seen clearly insofar as both of them put forward proposals, such as 'electronic personhood' and 'autonomous systems as authors of creative works', but later expressing doubt or downright criticism about those proposals. Both presenters were practitioners of *civil law,*[12] not human rights lawyers or privacy lawyers. Hence, there was something of a distance to the ethicists, since civil law may possess fewer tools for dealing with concepts such as dignity or self-determination. Whereas law remains much closer to the human subject than robots as subjects, it does not define the key concepts of autonomy and agency, dignity, self-determination etc., except in ways that can be legally qualified.

Lawyer 1 specialises in civil law with a strong interest in questions relating to non-human agency and liability. She was involved in the writing of the Green Paper on robotics. She started out by pointing to the fact that law is very human-centric, and that the human is the first subject of the law. A central focus was the shortcomings of the existing European Liability Directive and tort and liability regimes. *Lawyer 1*

[12] Dealing chiefly with matters such as corporate and business, family law, tort and liability.

argued that we may have to change the law or look for new legal principles since a robot cannot be treated simply as a thing. Possible analogies considered in addition to 'things' included animals and children. Yet, due to unpredictable events and environments these entities are not within the control of the person in charge. She thus encouraged the others to consider the chain of involved parties, such as: modellers of robots, producers, programmers, owners, users, teachers. The chain is increasing and product liability is not sufficient. We may therefore have to create a new legal status for robots, that of electronic personhood. Resources for trying to address this problem were taken mainly from two resources: first, the notion of legal personhood already enshrined in corporate law; secondly, previous thinking (Solum) about Artificial Intelligence and software agents.[13] A second (complementary) possibility included more interdisciplinary and organisational measures, for example, ethicists or technology assessors working for each company, and commitment to privacy by design and by default. Referring to the emerging General Regulation on Data Protection, she highlighted the example of how regulation is mandatory and emerging from a legal regulation. Yet, even as *Lawyer 1* proposed these possible solutions, she remained sceptical about their feasibility and/or necessity.

Lawyer 2 also works within civil law but focusing especially in IPR related issues. Her presentation oscillated between two different notions of autonomy. First, there is the question of robot autonomy, which was not investigated in full but rather focused on the question of intellectual property rights for machine-generated works. She provided several examples of such works.[14] There are different opinions within different legal regimes: The European Court of Justice has upheld the notion that work must be created by a real person to be protected, whereas the UK common law provides incentives also for machine-generated works. This presenter remained sceptical about the possibility, but found it interesting, theoretically speaking. Her preferred concept of autonomy was *Roboticist's 5* dynamic autonomy but how to adapt this to IPR did not become clear.[15] Secondly, there is the question of autonomy of and for the law. What is autonomy for law? According to *Laywer 2* it does not really exist, save through attributions of legally relevant acts. In general, this presenter

[13] A possible problem here is that these lack characteristics normally used to identify agency, such as personhood, soul, etc. According to *Lawyer 1* this is not so from a legal perspective. However, we may pose the question whether this conclusion would be as straightforward from the perspectives of human rights law or privacy law.

[14] Such as Harold Cohen's Aaron, David Cope's musical experiments, Patrick Tresset's drawing machine, the works of Anna Mura & Leire Guerela.

[15] Roboticist 5 told that in parts of his research on evolutionary robotics he sees evolution of new behaviours not even conceived by humans by the time they appear. There is thus a distance between the human creator and the creation. Here the matter seems to be not so much the property rights as such but rather the (im-)possibility of holding someone responsible.

was in favour of law proceeding stepwise and in a conservative manner: We should think carefully before we introduce new concepts. Law is a normative instrument that always has to balance different considerations and values to find the optimal mix between innovation, protection of weaker parties, welfare, sustainability, problems of ageing populations, etc. Still, law comes under pressure, she claimed, and law is always reactive. It cannot be created outside of societal processes. She argued that there is a real danger here because industry will push too hard for certain laws to be passed. To this claim, Epinet *researcher 3* replied that calling law instrumental is only one side of the coin. There are also the fundamental rights and freedoms, which can be used creatively.

PPP representative.[16] One representative from the newly founded PPP joined the workshop on the day. His background is in robotics, and he has worked for the robotics company CEA-LIST, doing research into AI and applied uses of robotics in hazardous environments, manufacturing and health care. He is now in charge of the ELS activities of the PPP. He explained autonomy as the capacity of a machine to execute tasks independently from human interventions. In order to do so, the machine needs to be capable of learning, generalization and action. Within the PPP there is no one solution to how to bring about robot autonomy. Rather, there is openness toward all kinds of technical solutions. He sees autonomy as especially relevant in the field of health care and assistive technologies for the elderly, but also as a requirement from the legal field, from which he mentioned the examples of non-contractual liability and labour law. Autonomy must therefore be seen as a broad umbrella-term that seeks to capture the efforts of several disciplines. The aim of the PPP is to become world leader, a goal that can only be reached if a number of hindering obstacles are overcome. Examples of such obstacles are non-harmonized legal frameworks and markets in Europe, protectionism, problems addressing non-contractual liability, and a lack of clear answers from societal actors on how industry should proceed to promote robotic technologies. Another problem is the poor levels of appreciation for ELS issues inside the robotics community. This is also a pressing concern since poor communication and misunderstandings with society can lead to the stigmatization of robotics. Europe should aim to become world leader also in ELS issues. The PPP aims to take a broad and inclusive approach to ELS issues, as exemplified in a new coordination action called Rock EU. This project offers the possibility, through its topic group on ELS, to create links with the other topic groups inside the PPP for collaborative purposes. He invited ELS practitioners to take part in this work.

[16] This presenter could have appeared in the roboticist session on the previous day as well. However, he only arrived in time for the ethics/law session.

2. Facing problems

The disciplines of ethics and law come under pressure from robotics development, and they are challenged by having to rely on roboticists as their main source of knowledge about current and future research trends. Another form of pressure comes from working within 'unhealthy' power structures, where many practical real-world problems are not informing decision-making. While the ethicists in this workshop related their views to philosophically-based understandings of autonomy, the key issues they emphasised were based in very similar kinds of social and technological interdependencies as the ones faced by makers and users of robots. Some of these dependencies pose a threat to human autonomy and to social cohesion, they argue. The lawyers, on the other hand, proceeded quite differently, given the need to qualify their analyses in terms of legally relevant acts and principles. Their analyses were brought up against the limits of existing legal practices and principles, sometimes triggering a search for new concepts and institutional arrangements.

There was a general sense that, given time and opportunity, many of these problems could be worked out. Learning takes place between ethicists, lawyers and roboticists. Also, both ethicists and lawyers repeated some of the same problems addressed by roboticists as 'politics interfering with their work'. Participants spoke of the unhealthy reliance on top-down decision making in the EU, as well as too strong an influence of business interests and drive toward competitiveness:

Ethicist 1 argued that the problem is not robotics or robots, rather, it is powerful people who think this is a solution to the problem of ageing.

Lawyer 2 argued that there is a real danger to law insofar as industry will push too hard, and try to shape the law to its own ends.

According to *Ethicist 2*, Europe is controlled too much from the top. Relating to ethics, this results in a doubly moralistic attitude, a 'Jesuit attitude'. On the one hand, Europe elevates ethical principles to the highest status. On the other hand, it will simply order society around, and seemingly without any respect for those very same principles.

The push from above which many of the participants mentioned, is problematic as well in terms of the pressure it places on roboticists to seek public acceptance of their robots amidst cultural fears or even a stigmatization of them. *Ethicist 1* called for extreme care being required in communicating real potentials and limits in the application of robotics, and in carefully working out the ethical and social issues. *Ethicist 2* argued that ethics and law are the prime sources to support expressions of public concern, however, that is not enough. The question was raised, how the EC will allow public concerns to propagate from the ground up to the top levels.

3. Solutions were proposed to improve on the relations between robotics and society: technical, disciplinary and institutional relations.

Both the ethicists and the lawyers proceeded to propose solutions, e.g., to better protect human autonomy and social cohesion, to improve interdisciplinary collaborations and institutional arrangements, and possibly include larger parts of the public for improved public relations. Among the main proposals in this vain, we summarize the following:

Standardization and standardization bodies.

Along with the recommendation of more bottom-up decision-making, *ethicist 2* proposed specifically to focus efforts on the development of standards. She claimed that roboethics principles must be shaped by research and development from the bottom up, while it is still too early to state definitely what the principles should be. Regarding the ISO, roboethics principles should be matched with specific standardizations… In her opinion, standards will be very important in implementing better tools to address the social and ethical issues relating to robotics.

Integrating technology assessment, ethics and law into robotics projects.

Several participants from different professional backgrounds, expressed their support for a proposal put forward by *Lawyer 1*, suggesting better integration of ethics and technology assessment into robotics projects. The approach could be toward self-regulation, however, made mandatory through law. The ongoing work on the new General Data Protection Regulation (GDPR) was mentioned in this respect as an

example of best practice—how to establish an ethical code of conduct including engineers, philosophers and lawyers. She posed the question of why we do not think it obvious that there should be at least one ethicist working for each company. Again it was argued that this seems to be the approach taken in the GDPR proposal which introduces, for example, privacy by design, privacy by default, etc.

Allowing for public concerns to be expressed bottom-up.

Better public engagement and more consultation were held up as important means to improve robotics-society relations (similar is proposed in the report *Taking the European Knowledge Society Seriously, 2007*). *Lawyer 2* mentioned the example of Dutch media authorities implementing a panel including citizens, and how that was given more importance than voices of bureaucrats and politicians. *Ethicist 2* argued in favour of such activities in relation to standardization work; she also argued in favour of much stronger emphasis on Europe's young people and their ideas being reserved a place at the table.

Final round of discussion – Implications for innovation and governance

In this section we provide a summary of the discussions, based on written notes whereby the contributions are paraphrased, not quoted. It is difficult to sum the discussions up in terms of some particular direction of debate. Our participants addressed different issues in no particular order, some of which engendered response, some which serve as standalone comments. This illustrates quite well the character of the deliberations as they took place across the different disciplinary domains. In the next (and final) section we sum up the overall conclusions from the workshop, including what we see as key implications for governance.

Ethicist 1: I want to highlight two things: 1) There is a strong incentive to keep humans in the loop. Regulations could be used for restricting (robot) autonomy in certain domains, such as military robots. 2) Technology Assessment (TA) emerges as a tool for law, and also as a way of increasing participation from different disciplines.

Epinet researcher 2: regarding discussions in public fora - there is a pressure to speed up public discussion. For example: I attended another workshop on robots for the police – 35 civil servants and researchers. There was a willingness to skip basic rights entirely under pressure to speed up the discussion. We are still in a kind of a utopian phase – people are blinded by possibilities, we don't see the negative sides.

Epinet researcher 4: I agree with what *Epinet researcher 2* says. We are supposed to advice Brussels – they want a tool/tech fix to deal with discrepancies and conflicts of disciplines. In this kind of format (i.e. this workshop) discussions have been very good. In situations of lopsided power relations, the quality of the exchange is lesser. How to transport this kind of wisdom to inform and change the culture of communication? It is a struggle.

Ethicist 2: In my view, Europe is too much controlled from the top. In other places you have younger people inventing, and such innovation is not shaped by the usual top-down forces. We need to unleash the powers of the young in Europe. To think, apply and invent, it does not require a lot of money.

Epinet researcher 1: given the expectations created within the policy environment, the tendency is to simplify – what problems are relevant, etc. The Epinet challenge to decision and policy makers is an exercise in deliberative complexification – questions of which issues *could* be raised. We need more effort – there are no final solutions. Honesty about occupational capacities – this is the real ignorance everywhere. We need honesty and transparency.

Roboticist 1: In the future we need more people that are generalists in this field, -to combine more perspectives. We need people who can serve as guardians of robots. In care situations (& consumer issues): to oversee issues and problems, connect responsibilities and situations (example: the blind owner. For children, -if parents cannot take care then others need to see to it that they are taken care of).

Roboticist 3: I worry about EU support for assistive robotics, taking place without the possibility to adequately discuss 'what is a robot' - people have too high expectations of what robots can do. We as researchers need to make it clear that there are many things that robots cannot do and that there are many problems that cannot be solved. But due to the structure of funding we need to sell ideas. I like *lawyer 2*'s idea about philosophers/ethicists working in the companies / research labs. They can make clearer the limitations of what robots can do and what they cannot do.

The aim of the present PPP is to make new business work – and this is very risky.

Roboticist 2: I'm still excited about this event – I have learned a lot. What became very clear to me pertains to the use and misuse of words, how they work across fields and barriers, and this is even more problematic when approaching the public. There is a great need to involve ethicists and lawyers. I learned a lesson: we need to be careful about how we use words.

Lawyer 1: This kind of interdisciplinary forum is really important – we need more funding for this kind of research. Our centre for instance is really under-funded. We are in-between fields [law+STS], it is difficult to exist in-between conventional disciplines.

Roboticist 4: Robotics can be applied everywhere, so everybody should be included, especially young people. Many fields are really far away from the roboticist's point of view. For instance, medical doctors at a Tufts university workshop were difficult to engage: doctors are really difficult, nurses are better…

PPP representative: Robotics is one of the main tech revolutions for the coming age. We need to anticipate, to provide a framework that is clear and simple – to become leaders in the world. There is a meeting of the PPP on March 12 – you are all invited – we will work practically on issues of how to take ELS issues into account in the program. I agree with *lawyer 1*: a main problem is financing – it is difficult to bring in the ELSi people – we need to fund and reinforce this.

Roboticist 4: What do you think about adding ELSi competencies to technical fields?

PPP representative: The ELSi people should be included, and also lead some actions, as in Robolaw. We are not aiming to exclude anybody. We are trying to define the important things to be done in the PPP. As for the problem of jobs: robotics creates more jobs than it destroys. If we cut unskilled labour we need to see how such elements can be brought into account.

Epinet researcher 3: To raise the issues of the role of law. Do we have the proper tools for doing our jobs well? Regarding the normative fields [hard/soft law], it seems that law is under transformation, both law and legal scholarship. As a knowledge form – how does it play out against TA or other related practices such as privacy by design, impact assesments, etc?

Roboticist 5: I am very mindful of the dangers of robo-exceptionalism. Still, the following seem to me to be peculiar to the field:

– the unrealistic expectations

– robots have agency

– robots provoke emotional attachments and dependencies

1. As for future EU projects in H2020 and the PPP: they should not be approved unless they address societal actors, sustainability issues, etc.
2. The time has come for us to treat robots as we treat clinical trials and experiments on human subjects. Specific ethical approval should be recommended.
3. The question of impact on jobs. I would like to see companies and manufacturers undertake a social compact. For every job that is removed the company undertakes to re-train or redeploy that person. If not, they can pay a tax. Replacement costs will have to be found in welfare budgets.

Epinet researcher 1: That is very radical, since it goes against the grain of the policies of the day.

Roboticist 5: As it is now, the private sector gets richer and the public sector gets poorer.

Roboticist 1: There is a race to the bottom between different legislatures. Regulation will have to take place at EU level, at the very least. Cf. also the concept of *in-sourcing* in the US, where manufacturing is brought back.

Roboticist 2: SMEs would be scared even more. Socially speaking I agree, but economically…it could perhaps apply to big factories.

Epinet researcher 4: What is work, and what is work going to be in the future?

Ethicist 2: Robots could be cleaning up the Atlantic Ocean: just think about it.

Conclusions

Drawing together the conclusions reached over the two days, there is no single meaning, use or scientific content to be associated with the concept of *autonomy* for robots.

Insofar as robot autonomy is conceived as major contribution toward solving the grand challenges of an ageing society and sustainable healthcare, the *official* vision is flawed.

Autonomy for robots – insofar as it can be put to meaningful use – and autonomy for humans are not the same thing. Autonomy for robots is always derived from human concerns and interests. Forgetting or ignoring this could lead to undesirable social and ethical consequences.

We learn how differently roboticists, ethicists and lawyers conceive of autonomy. One common trait is that all conceive of autonomy as relational, and never as a free-standing property of either humans or machines.

Roboticists and lawyers tended towards seeing the ethical and societal issues arise along the trajectories of gradually increasing robot autonomy. This was different to the ethicists who tended towards situating autonomy, human subjectivity and community at the centre of their analysis. We identify a lack of attention in the law to human rights and privacy, the branches of legal scholarship that are closer to ethics articulations. This has been pointed out in the Roboethics Roadmap, but has not been much emphasised in the more recent ELSi activities of EUROP and euRobotics.

In spite of the many differences in approach and orientation, there were clear indications that participants – most of whom are experienced in c-disciplinary exchanges about ethical and societal issues – were quite capable of relating to the practices, ideas and concepts of each other.

Provided ample time and opportunity, many ethical and social issues could be better worked out across the disciplines present in this particular workshop. There are, however, serious constraints in terms of time, opportunity and funding.

Perhaps the most taxing issue that kept cropping up is how many see their practices, the values and goals they are meant to serve (better care, human dignity, smarter robots, improved human-machine interactions, etc.) as pressured by top-down structuring and control of innovation agenda setting and funding priorities. It is quite possible that these tensions will be further accentuated by the new PPP.

Appendix I – What about autonomy?

(Kristrún Gunnarsdóttir, Lancaster University)

Although the RoboCom (RCC) proposal was rejected, the case study partners chose to stick with the question of machine autonomy as a *policy problem*. There are a number of reasons for this choice, one being the lack of a coherent vision of what machine autonomy stands for and what it can achieve, along with uncertainties about the extent to which one or another form of autonomy should be delegated to machines. This policy problem was used to guide the programme of the networking event:

> How does autonomy figure in the making of the next generation of robotics, and how should machine autonomy feature as a policy issue with implications for the safety of persons and operations?

It was therefore of some value to our approach at ULANC that the RoboCom vision has faded into the background—that the event afforded a more open discussion of a range of understandings of autonomy, according to who is articulating or demonstrating its workings: engineers, philosophers, lawyers and other. We welcomed very much the opportunity to discuss a variety of definitions and products already competing, thus, demonstrating a number of ways in which autonomy is defined, enacted and recognised. In this respect, the event was particularly instructive to our approach. Not only did it confirm some of the findings and views we elaborated in previous reporting and forthcoming publication (Rommetveit, Gunnarsdóttir, van Dijk, et al. forthcoming). It allowed us to investigate *autonomy* in reference to the aim our research started out with.

1. To clarify the role of these visions (here RoboCom) in reflecting on the relationship between humans, animals and machines, as well as their role in robotics development as a scientific tool to test our theories of intelligence, behaviour, sociality, emotionality and embodiment.
2. To better understand the scientific and technological problem domains in robotics development in relation to and apart from these visions—what we refer to as the *challenges of enablement*.

What did the event teach us about autonomy?

1. Autonomy is partial and negotiable in and through human-robot interaction.
    a. This is particularly the case in the development of assistive robotics where the key to achieving the desired results rests on adequate acquisition of data (on humans and the immediate environment) which then are translated into interaction cues that allow for some improvisation of who does what and how. There are enormous privacy issues to consider in these interactional scenarios, given the kind of information collected, processed and relayed.

2. Autonomy is an apparition, a make-believe, *as if* a machine had human-like capabilities.
    a. A bit of a Wizard of Oz is often put to work in robotics, e.g., the design of Paro and other therapeutic robots, where the perception of autonomy rests on theatrics: a show of fellow-feeling, sentience and intelligence.

3. Autonomy is emergent in and through adaptiveness.
    a. A vision of machine adaptiveness is pushing toward the increase in machine autonomy as the machine adapts to its users and the environment, however, there is also the adaptiveness of human-robot relations to consider, in and through which new (semi-autonomous) agencies emerge.

4. Autonomy is perceived/detected in relation to our own.
    a. While there is no sign of a key-enabling general problem-solver that can replace a human, we can consider the idea of and attempts at autonomous systems with reference to the effect upon the perception we have of our own autonomy—how it is lived, recognised, understood and indeed protected.

5. Autonomy is put to effect as companionship.

    a. Science fiction narratives about robots often represent the human rather than the robotic. The *other* in us is encapsulated in a robot with which the human develops companionship, however, the adaptiveness of the robot and the companionship will play on fear—that robots are a threat to human freedom and autonomy (e.g. I Robot and Terminator).

    b. One can argue that a *companionship* with robots (fictional or real) is a novel way for humans to relate and relay their emotions, their empathy and vulnerabilities—that the autonomy of the device lies in its delivery of a mirroring of the humanly.

6. Autonomy is put to effect as a way to connect with other humans.
    a. Research on the elderly population suggests that if assistive devices promise improved social engagement, then people are likely to sacrifice quite freely some of their privacy and sense of dignity—a way of holding onto their own *thick* rather than *thin* autonomy.

7. Autonomy is implicit in solving *autonomously* particular problems: some mechanical function; a visual recognition and response function; data analysis and response function; an act of social etiquette.
    a. The kind of autonomy built into assistive systems for the disadvantaged, appear to be primarily about the robotic devices having access to some kind of data or data streams, and built-in capabilities to act upon such data in pre-defined ways.

    b. Some acts can be broken down into autonomous task-specificity, resulting in the ability to perform social etiquette, say, handing a knife over with the handle facing forward.

8. Autonomy is autonomously *knowing* what to do next.

    a. When autonomy becomes implicit in solving particular functional/mechanical problems, it raises the question of whether the robot needs to be a sentient intentional agent to exhibit that autonomy. There is nothing to indicate that this is achievable.

9. Autonomy is integral to the history of the automata—the *robotic*.
    a. Looking at the robotics industry and market trends to-date will provide a good overview of progressions whereby *autonomy* as a descriptor of robot abilities is not necessarily very helpful. Progressions to-date are strongly linked with the history of the automata—the automaton which is only autonomous in executing pre-defined tasks (e.g. acting like a *machine)*.

10. Autonomy is linked directly to the problematic timeless question of *intelligence*
    a. One can ask what is intelligent and how it is intelligent, for example, in care scenarios. Do translations between scripts and functions that *can* and *need* to be known in advance for design purposes, require *machine understanding* of the resulting actions, for example, that no harm is intended, that the purpose is to explain or to entertain, that the matter is emotional, sensitive, private, and so on?

    b. One can ask what is machine sentience even if such phenomenon is nowhere on the horizon.

c. One can consider human adaptability to human-robot interfaces. Experience shows that people need extensive training with robotic devices in order for the intended (by design) human-robot interactions to work properly. The stability of devices can be a major challenge here. Users get into interactional problems and the devices have to be re-figured over and over again.

d. One can consider potential learning capabilities and adaptability, but there are very specific technical problems associated, for example, with how to bring comprehension abilities of robots to an acceptable level so they can be assigned simple household tasks. There are major challenges in building machine learning into mobile supports. Some of the ethical challenges here include provisions to make robots and robotics development safe, to always keep robots in the control of some human, not to employ them to kill, and make sure to have a clarity on command hierarchies in complex scenarios. It should be noted that liability issues are not only about whether or not robots can do harm but also what the implications are of robots being harassed and vandalised by humans.

11. Autonomy is not envisioned singularly as a distinct phenomenon, but in multiple ways by considering the purposes for which different forms of autonomy are deemed necessary.

    a. Anticipating fully autonomous systems is completely misleading, possibly a disastrous idea as a solution to a societal challenge. We should also be very careful with economic modelling that promises to save money. Different forms of semi-autonomy however, of interactive and assistive capabilities, could be very helpful.

    b. Building helpful assistive devices and systems is shown to be achievable: some care capacities *can* be freed up; some hygienic assistance (non-human) *can* be applied; wearable sensors *can* contribute to the preservation of a person's mobility; and, simple assistive capabilities *can* contribute to a person's independence.

Relating our findings to assessment traditions

It is not new in the STS study tradition of social epistemology to learn that the expectations of scientists and technologists are ill-considered in terms of the societal and culturally-specific assumptions they make. It is not new either in the study tradition of exploring phenomenal fields of practice, to come across incompatibilities between the kinds of things that can be said about executing a task or solving a problem, say, for AI or CHI design purposes, and how tasks and problem-solving actually unfold. Our research takes advantage of the findings and methods available to these study traditions, to elaborate our own critical evaluation. But, more important here are the findings that are leading us to a much deeper understanding of how uncertain the study objects are within this domain of innovation, in particular, that of autonomy as a problem domain, and what the implications of that are for future developments, for policy and regulation, for ethical considerations and so on. In this respect, the contributions of our invitees at the networking/embedding event as a whole was particularly helpful to unravel definitions and product examples along with the many ways in which autonomy is defined, enacted and recognised. The constitution of our case study group is particularly useful as well, given this focal point of our research. It offers rich material to legal scholarship, to ethical enquiry and uncertainty analysis—to join up our investigations and assessment to-date, along with the learnings we take from this networking/embedding event.