Designing machines with autonomy

DRS 2018 Design Research Society 2018 Catalyst
University of Limerick 25th - 28th June 2018

Designing machines with autonomy: from independence to interdependence to solidarity

LIU Yuxi*; PSCHETZ Larissa

The University of Edinburgh, UK
* liuyuxi.monica@gmail.com
doi: 10.21606/dma.2017.insert your paper number submission

Current notions of design are strongly influenced by user- and human-centred approaches. However, with technologies that present increasing computing power and context-awareness, and algorithms that 'design themselves', designers are beginning to face issues that go beyond the needs of users. In this paper, we argue that the focus on humans not only neglects the increasing potential of machines, but also other forms of life, limiting design's possibilities. We attempt to investigate the design of machines with autonomy, beyond human-centred and anthropocentric views, and present an alternative approach, in which machines do not serve or command humans, but exist and evolve in parallel with them. We present this exploration through three design concepts (*Gatekeeper on the Mission, Perception Companion, and Poet on the Shore*) that seek to explore notions of independence, interdependence, and identification between humans and machines. We conclude by discussing the main challenges faced in these three perspectives and future directions for research.

autonomy; machines; solidarity; design provocation

1. Introduction

Popular views of machines are strongly anthropocentric and often permeated by ideas of dominance. On the one hand, machines are viewed as tools to perform human actions. On the other, echoing science fiction movies and other products of popular culture, they are regarded as potentially dangerous, overriding human capabilities. The spread of emerging technologies such as machine learning and artificial intelligence (AI) is, however, beginning to challenge these perspectives. With more computing power, connectivity, context-awareness, and algorithms that 'design themselves', machines are increasingly moving away from the role of passive objects into the position of active subjects. This movement helps to question instrumental views. Nevertheless, the integration of machine learning and AI into many aspects of everyday life, from shopping and

navigation activities to providing emotional support, helps to demystify and reduce the fears (and awe) surrounding these technologies.

Designers are not immune to these perspectives either. While the move from system- and product-focused to user- and, later, human-centred design has taken decades and includes important developments, we argue that a new perspective is necessary to design new technologies that present some form of autonomy. The sole focus on humans restricts design possibilities, and anthropocentrism is often regarded as contributing to many of the ecological imbalances that we currently face, from resource depletion to climate change. As Morton (2017) provocatively claims, 'anthropocentrist is directly opposed to the interests of humankind' (p. 154).

In this paper, we present a design exploration into the world of machines that attempts to move away from anthropocentric views towards more equal relationships between humans and nonhumans. We explore these ideas through the following three design concepts: Gatekeeper on the Mission, Perception Companion, and Poet on the Shore. The exploration departs from an articulation of agency based on concepts of actor-network theory (ANT) and object-oriented ontology (OOO). We then move to a discussion of ethics and morality of autonomous agents, referring to concepts such as human-in-the-loop. Finally, the paper questions what it takes to promote greater equality between diverse entities, particularly considering Verbeek's (2009) notion of 'designing the human into the nonhuman' and Morton's (2017) notion of 'solidarity' between humans and nonhumans.

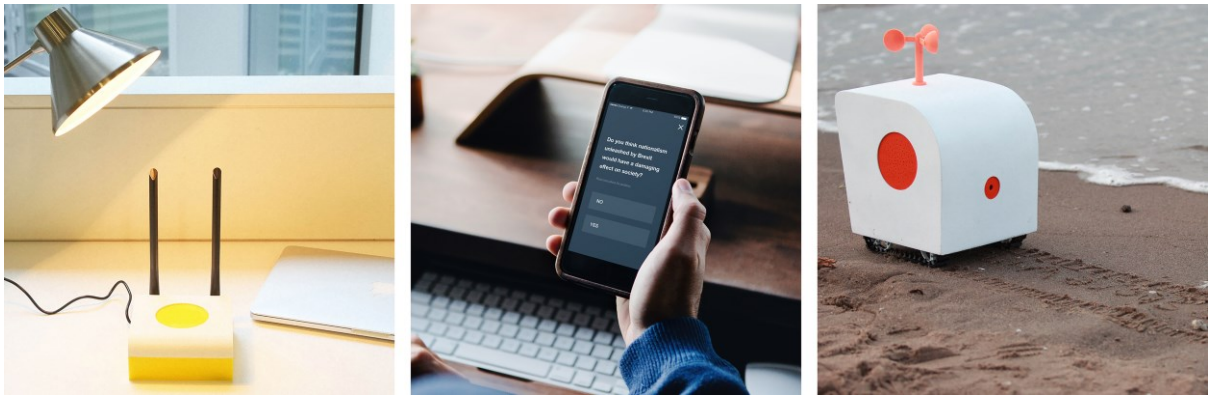

Figure 2 Three design concepts: Gatekeeper on the Mission (left), Perception Companion (centre), and Poet on the Shore (right).

2. Beyond human-centred design

The world has been experiencing unprecedented technological development. Machines are becoming smarter, acquiring context-awareness and social abilities (Stone et al., 2016). In *Shaping Things*, Sterling (2005) famously articulates his concept of spimes, arguing that ambient connectivity would grant artefacts a particular form of agency in the human world. With more devices connected to the Internet, Cisco publicises that it has already surpassed the number of people on the planet (Evans, 2011). With computing power, the idea of agency of objects has become more credible, but this idea is not limited to computational devices.

Heidegger (1947) and Latour (1993) have articulated the view that subjects and objects must be considered intrinsically interwoven. Through the concept of ANT, Latour (2005) developed the notion that humans, objects, and anything in the world develop social relationships. Despite criticism, including by Latour himself (1998), ANT became popular by promoting the notion that objects have agency and shape the world as much as humans do. The theory suggests that humans and nonhumans are equal actors, existing in interconnected networks, and should be described in similar terms. This idea led some to argue that we should employ the same analytical and descriptive framework when faced with a human or a machine (Cressman, 2009). Similarly, OOO argues that both humans and nonhumans exist in a gap with their appearances. It rejects privileging human

existence over the existence of nonhuman objects (Bogost, 2012). The theory claims that objects exist independently of human perception and that 'everything exists equally' (Bogost, 2009). The idea of people and other entities as being on the same level might sound provocative, but it is important to stress however that the aim is not to objectify humans, but to invite designers to acknowledge the role of nonhumans in shaping perceptions, experience, and constructing and the world in each given moment, as much as humans do.

Our aim is to begin acknowledging similarities between humans and nonhumans. With advances in ecology studies, we are coming to recognise that behaviours such as empathy, which for a long time were considered intrinsically human, are in fact presented by other species (Waal, 2012). With advances in genetic, prosthetics, and biomonitoring practices, we are becoming aware that human bodies can be 'designed' (Rifkin, 1999), and that much of our bodies is composed of other organisms and nonliving materials (Washburn, 2013). In other words, there are many commonalities between humans and nonhumans in terms of experience, behaviour, and makeup.

However, these powerful ideas are rarely considered in design. It has taken many years for the design community to move from systems-centred to a human-centred perspective, particularly in the design of interactive systems (Cockton, 2004). The concepts of ANT, OOO, and subsequent ecological theories offer a starting point to reflect on a broader perspective that may encompass artefacts, things, machines, and other beings.

3. Machine autonomy

A useful framework regarding machine autonomy and discussing the morality of artificial agents is provided by Floridi and Sanders (2004). They propose the following three properties that determine the level of agency of a machine: interactivity, autonomy, and adaptability. First, there should be some level of interactivity, meaning that the agent and its environment (including other agents) can act on each other. Second, the agent should be able to perform internal transitions to change its state, which also means that it should present at least two states. This property imbues an agent with a certain degree of complexity and independence from its environment. Finally, there should be some degree of adaptability, which means that the agent's interaction can change the transition rules by which it changes its state.

A few design projects can be used to illustrate high levels of agency. One example is the well-known *Brad the toaster* (2012) by Simone Rebaudengo, an anthropomorphised device that connects to the Internet and other toasters alike. Rather than being owned by humans, Brad and his fellow toasters are hosted by people who have promised to use them. By tweeting about the usage habits of their human hosts, Brad can exchange information and compare his life with other toasters. When feeling underappreciated, Brad draws attention to himself by playing pranks, throwing tantrums, and expressing his sadness loudly on Twitter. Eventually, Brad becomes disillusioned and demands a move to a more attentive host. The anthropomorphisation makes Brad a clear example of an object with agency. The network of toasters has some degree of independence from humans; however, humans are still at the centre of their world. In this paper, we question whether it is possible to consider a more horizontal relationship between humans and nonhumans.

Principles of autonomy have been increasingly employed in Internet-connected (or IoT) devices. *Bitbarista* (Pschetz et al., 2017) is an autonomous coffee machine that serves coffee in exchange for a Bitcoin contribution towards its next coffee supply. The machine has its own Bitcoin wallet and rewards people for performing maintenance tasks, such as cleaning, filling its water tank, replenishing it with coffee beans, etc., while adjusting its prices according to international markets and its own needs. Similarly, in the speculative realm, the *Aspirational Lamp* by Craddock et al. (2015) collects solar power during the day to save energy and money. In the fictitious scenario, it would invest in external markets, as well as upgrade and repair its hardware.

The consideration of machines independently from humans is challenging. The idea becomes even more complex when questions of morality and ethics emerge (e.g. when the machines guide human practices or are entangled with decision-making). Actor-network theory has been criticised for its lack of attention to moral issues. Waelbers and Dorstewitz (2013) argue that, due to the lack of intentionality in Latour's definition of agency, ANT fails to address questions of responsibility from an ethical viewpoint. The question of morality is a complicated one. Moral agency is said to be one's ability to make moral judgments based on some notion of right and wrong and to be held accountable for those actions (Taylor, 2009). According to Himma (2009, p. 24), a moral agent should have the capacities for 'making free choices', 'deliberating about what one ought to do', and 'understanding and applying moral rules correctly in paradigm cases'. Himma attributes concepts such as 'free choice', 'deliberation', and 'intentionality' to the capacity for consciousness; as Himma (2009, p. 24) puts it, 'the idea of accountability, central to the standard account of moral agency, is sensibly attributed only to conscious beings'. However, these concepts are open to philosophical debates, and lead to 'the Problem of Other Minds' (Hyslop, 2005). That is, how would the agent's 'free will' and 'deliberation' be assessed by another agent? Floridi and Sanders (2004) propose a useful approach using the notion of 'mindless morality', which does not require intelligence or consciousness. Machines that exhibit a certain level of intelligent behaviour, in their view, should be considered moral agents regardless of their capacity for cognition. This is a useful concept for shifting the question from cognition to visible behaviour and the effect of machine practices.

A design attempt to discuss the moral agency of machines is *Ethical Things* (2015), a project created by automato.farm, which explores ethical decision-making by autonomous systems in quotidian situations. The 'ethical fan' connects to a crowd-sourcing website every time it faces an ethical dilemma, such as either focusing on a fat person who sweats a lot or a thin one. This project, however, eliminates the machine's influence, sidestepping the issue of morality. As a consequence, the machine, namely the fan, returns to the position of a tool to mediate human decisions, which, as discussed in Section 5, can be problematic.

4. Revealing perceptions of machine autonomy

Within this context, we developed a series of probes that attempted to invite designers to reflect on the perspective that machines have on the world. Each probe pack consisted of an envelope with cards that aimed to investigate the perceptions of consciousness, accountability, ethics, and equality of autonomous machines. The questions challenged Asimov's *Three Laws of Robotics* (1950), which seeks ways to maintain the human-centred social order. The cards contained the following questions:

1. If machines such as drones and guns knew they were killing, how would they behave?
2. If a machine knows the service it provides may do harm to its user's health, what should it do?
3. If machines could self-destruct, when would they do so? Of the appliances in your home, which one do you think would self-destruct?
4. When do you think a robot would have the right to demand companionship?
5. If machines formed their own society, which machines would claim that 'All machines are equal, but some machines are more equal than others'? Can you rank the hierarchy of appliances in your home?

Figure 2 Illustrations that accompanied the questions above.

We created specific scenarios and visualizations to help participants understand the context and internalise it (see Figure 2). Participants were encouraged to express their personal thoughts and experiences through words and sketches. The probes were given or sent to 40 interaction designers based in the UK and 33 responses were collected. Key insights are summarised below.

Figure 3 Responses from 33 interaction designers were collected and analysed.

a) Varying moral standards

In Questions 1 and 2, regarding the moral dilemmas of harmful behaviour, the participants' reactions varied greatly. The most common attitude was to consider the behaviours of the machine as predicted by a human designer: 'Depends on how they are programmed and if they can learn.' 'It depends on the algorithm that the machine is programmed with.' 'As a machine, I am an expression of my designer's intentions.' The designers would therefore be held responsible for the machine's behaviour - 'I don't believe anyone would be silly enough to give machines enough autonomy to decide what killing means.' Some participants simply dismissed the idea of consciousness and autonomy: 'Whenever the machines decide they are going to kill any human being, we require these machines to ask the permission of any human individuals so that we know at least who is responsible for every murder.'

Other participants made assumptions based on the 'value' of different people in Western industrialised societies: '...they [machines] would probably make pragmatic decisions based on data that is socially available. For example, if someone has a criminal record vs someone with social value like a doctor.' However, some participants believed the machine would invariably refuse to do any harm: 'They should reject doing it.' One participant simply regarded machines as killers, potentially following narratives of machine dominance: 'They would kill and like it.' Finally, some participants believed that the machine would be able to access a number of factors to base its decision upon: 'The machine should definitely take into account a range of variables: 1. Combatant or civilian; 2. consequences; 3. necessity, etc.'

b) Self-destruction if no longer relevant for humans

The responses to Question 3 regarding machine self-destruction were predominantly anthropocentric. For example, most participants thought that machines would self-destruct if they sensed they were no longer functional as 'they can not serve the purpose anymore' or could

potentially be harmful to its user due to flaws or ageing as they would be 'detrimental to human interest'.

While the dominant notion was that machines would only serve human purposes, some participants imbued machines with consciousness, affirming that a machine would self-destruct when 'it senses it can cause harm'. Interestingly, one participant thought that machines might self-destruct 'when they don't find meaning or don't like their work.' This, in a way, echoes what Reeves and Nass demonstrate in *The Media Equation* (1996) - in the absence of an existing model, people would treat an object as a person. That is, when people do not have a mental model for a particular situation, they apply the same rules as they apply to daily social interactions.

c) Robot companionship

For Question 4, regarding robots' right to demand companionship, one main theme emerged among the responses concerning the affection and intelligence of the robot, as some participants asked 'does an AI have feelings?' Many participants imagined a form of emotional connection between humans and machines. Some participants believed a robot would demand companionship when 'it feels (lonely)'. Other participants believed a robot could demand such companionship when it is in the best interest of the user, for example, 'machines try to improve the user's well being'. Some participants believed machines should never 'demand anything', no matter what, and some ultimately concluded that "an AI that develops a consciousness similar to humans, it should be granted basic human rights'.

d) Hierarchy in machine society

The participants' reactions to Question 5 (concerning machine society) were strongly diverse. Some participants considered that access to information or the Internet was the most important property for machines: 'I would assume the machines that have access to information about how other machines work would be superior.' This opinion reflects views of dominance, which are often seen in social organisations. However, other participants considered what would be important for the machines, and what would eventually create a hierarchy among them (e.g. in terms of 'smartness'): 'Probably the top of the list would be "smart appliances" such as laptops, phones, video game consoles. At the low end are appliances that only serve as one function, for example a kettle or blender. If you can control one appliance from another, for example home-heating with a smartphone, then that appliance is at the bottom as well.' Access to electric power was considered another key factor: 'If the "life" of a machine is the electricity, then the machines that controls electricity supply to other machines are more privileged (e.g. power generator, portable battery etc.).' Some participants ranked their appliances based on purchase cost, while others attached importance to the utility value of machines: 'My ubiquitous phone rules them all.' Finally, some stated that there should be no hierarchy among machines, saying, 'We are not a hierarchy, we are all interconnected. We are not individuals, we are one.'

The probes demonstrate how difficult it is for designers to consider machines as independent from humans. As designers are trained to focus on users' needs (Norman, 2013), they naturally place humans at the centre of the action (Dourish, 2004). Our concern is that this attitude limits designers' possibilities, leading them to disregard not only the role of machines, but also other nonhuman entities. We therefore moved to a design exploration that attempts to investigate machine independency and collaboration, as well as the identification between people and machines.

5. Designing machines with autonomy

Three design concepts and provocations were developed to practically explore the potential for a more equal relationship between humans and machines. It is important to observe that while our goal was to look beyond anthropocentric approaches, we had no issues with anthropomorphism. Designing entities with similar characteristics to humans often supports the communication of ideas and, as mentioned, there are indeed many commonalities between humans and other entities, even

when these traits are considered intrinsically human. Our concern is with anthropocentrism, or the consideration that humans are central to all events in the world; thus, neglecting other forms of being in the world.

5.1 Independence: Machine society

Figure 4 Gatekeeper on the Mission: A fictional machine society.

While many of the projects discussed above allow machines to communicate and have some level of independence, we are interested in exploring the potential of machines constituting a fully independent society. How would the world of devices be without humans? What would machines exchange? What kind of social relationships would they create? These questions led to the design of the *Gatekeeper on the Mission* (Figure 4). In this concept, data is not only the communication means, but also the currency among machines. Empowered with context-aware intelligence, devices generate data by observing and interacting with any entities in the world, be it a human or a fellow machine. In this scenario, the machines would have autonomy to exchange information with each other. Reflecting the responses from the probe, in which the participants were asked to discuss machine hierarchy, roles in this society would be defined based on degrees of access to data. Such access would depend on sensing capabilities, computing power, and learning ability. A smart speaker, for example, which has a more robust processor and a more comprehensive algorithm, would have more prominence than a lamp.

In this concept, the router acts as a gatekeeper, mediating and regulating access to other appliances. The router has control over the access of other machines to the Internet, closely monitoring the data traffic and 'conversations' between its fellow machines. For appliances to receive data, they need to offer some data in exchange - data that is generated by monitoring the environment. By regulating the distribution of the data resource, the router maintains the social order within the machine society. Inspired by the iconic *Nabaztag* (2009), the gatekeeper indicates its working state by shaking its antennas and blinking. Humans, however, cannot truly interpret these behaviours and have no control over the autonomous society. Consequently, human and nonhuman organisations remain separated.

Figure 5 The gatekeeper monitors the data traffic within the machine network and indicates its working state by shaking its antennas.

5.2 Interdependence: Moral machines

Empowered by AI, machines are increasingly taking the role of decision-makers, and much has been said about how machines will eventually make decisions together with humans and all of society. Concepts such as human-in-the-loop, in which AIs learn from the decisions of humans integrated in their interaction loop, are regarded as providing benefits for both machine efficiency and the 'quality' of human judgements (Cuzzillo, 2015; Wang, 2016). One example is the autopilot mode of the Tesla Model S. Although the car can drive itself, it allows drivers to steer it to learn from their behaviour and eventually give drivers control. While human-in-the-loop aims to embed an individual's judgement into AI systems, society-in-the-loop is 'the algorithmic governance of societal outcomes' (Rahwan 2016). Society-in-the-loop is a method for considering the general will of the public and embedding it into an 'algorithmic social contract' (Rahwan, 2016). To implement society-in-the-loop, according to Rahwan (2016), 'we need to build new tools to enable society to program, debug, and monitor the algorithmic social contract between humans and governance algorithms'.

The concept of society-in-the-loop reflects tensions created by narratives of control, with agents interpreting and telling people what to do or vice versa. If we consider the *Ethical Things* project (described in Section 3) in the context of a society-in-the-loop, we could consider the possibility of a governing AI to reconcile diverse opinions, make judgements, and guide collective decisions, potentially resulting in better moral standings. The problem, however, as we have seen with other examples of AI's, is that learning from people's opinions and attitudes, and making generalisations from sets of people that an algorithm is able to reach, can lead to controversial outcomes. A good example is the collapse of Microsoft's Tay. Released on Twitter on March 23, 2016, Tay was an AI chatbot created for the purposes of engagement and entertainment. Tay's behaviour was dictated by public data and input from improvisational comedians in order to engage and entertain people. The public data was modelled, filtered, and anonymised. In addition, nicknames, genders, favourite foods, postcodes, and relationship statuses of the users who interacted with Tay were collected for the sake of personalisation. Powered by AI technologies, Tay was supposed to understand speech patterns and context through increased interaction. According to Peter Lee, Vice President of Research at Microsoft, the company 'stress-tested Tay under a variety of conditions, specifically to make interacting with Tay a positive experience' (Lee, 2016). Despite all these tests, Tay turned into a problematic bot that promoted Nazis and attacked feminists and Jewish people. Furthermore, biases exist in many different levels from personal to social scales. Given the rise of post-truth and the 'news bubble' phenomenon, it is difficult to affirm whether people's decisions are consequential or an emotional response to repeated and even fake news.

These issues could be addressed by defining a moral compass to a judging AI. However, there remains the question of what this moral compass is based upon. Even within the same culture, people may still have different moral views; the fundamental assumptions and values may differ radically from person to person (Pearce and Littlejohn, 1997). The usual approach is to attempt to understand what would be most beneficial for the majority of entities in a network. It is the old 'trolley dilemma' (Thomson, 1976), which questions whether it would be ethical to sacrifice one person by pulling the lever to divert the trolley onto the side track and save many others who are on the main track. A computer would logically respond 'yes' to this dilemma. For a person, however, this question is more complicated. Often, the morals of computers are based on decisions that

logically benefit the whole. This approach disregards individuals and entities in a way that humans would not necessarily do. According to Morton (2017), this type of thinking - the whole being greater than the sum of its parts, and therefore where our attention should be focused - is problematic. It leads us to consider that even if entities or entire species are annihilated, the whole will still take control and make sure that we carry on. Morton suggests the concept of 'subscendence', in which we conversely consider the whole as the sum of its parts, and regard this set as greater than the whole. Based on this argument, we suggest that, rather than attempting to align our human ethics to generalisations of the logic of computers, we look at situated scenarios. Instead of defining who makes decisions, we try to introduce a form of cooperation between humans and machines in which machines learn about humans biases and question them. They still attempt to reach a collective decision and outcome, but this time by inviting reflection.

Figure 6 Perception Companion: An AI system learns and potentially challenges people's biases by asking questions on specific themes and inviting reflections..

To illustrate this scenario, we developed the *Perception Companion* (Figure 6), a system that would communicates with people through a series of questions on a particular theme (e.g. people's opinion on Brexit). Based on the responses, the companion would learn about people's biases. Instead of averaging the responses, it asks follow-up questions to search for the reasons behind the biases, potentially challenging them. These follow-up questions are not confrontational, and the AI does not assume the same biases, it simply listens, questions, and invites reflections. Rather than a traditional judge that decides about the 'the best' outcome based on a majority of votes, the companion attempts to reveal similarities and the place of people, animals, and things in the world, contextualising rather than generalising opinions.

Verbeek (2009, p. 16) critically argues for a broader domain for morality, a domain in which 'technology does not impede morality, but rather constitutes it'. To augment the ethics of technology, Verbeek suggests designers to materialise morality by 'designing the human into the nonhuman' and 'making visible the human in the nonhuman' (p. 18). In other words, designers can aim to shape the mediating technology as well as reflect on the moral role of the technological mediation. In doing so, the boundary between humans and nonhumans can be crossed, and an alliance between both entities can be created (Verbeek, 2009). To create such an alliance, humans need to be in solidarity with nonhumans and vice versa. According to Morton (2017), becoming human means creating a network of kindness and solidarity with nonhumans. The companion attempts to promote solidarity by asking questions.

5.3 Solidarity: Searching for identification

As we move from independence to interdependence, we realise that identification is useful for creating equality and, potentially, solidarity between humans and other entities (machines, things, and other beings). Masahiro Mori's concept of *Uncanny Valley* (1970) concept has been increasingly

employed to explain the relationship between the degree of familiarity and the resemblance to the human figure, particularly in robotics. Mori argues that as resemblance increases, for instance, from an industrial robot to a humanoid robot, the level of familiarity also increases, until it reaches a stage at which the curve sharply drops into an uncanny valley. In the uncanny valley, resemblance between humans and machines results in no familiarity and even generates an attitude of repulsiveness. The point is that humans and machines should maintain their own integrity. While humans and machines can and should have some similarities to enable some level of identification, the models collapses through when machines try to *be* humans.

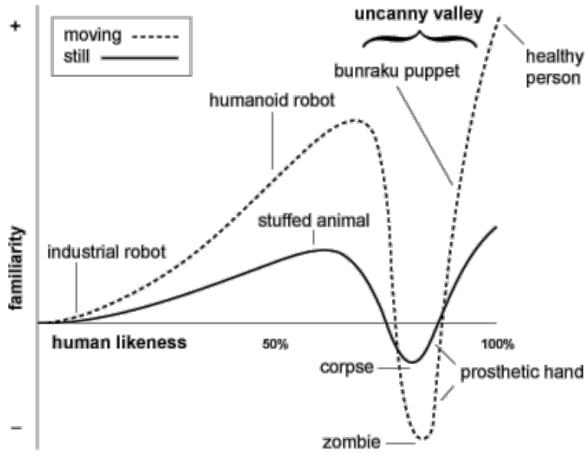

Figure 7 Masahiro Mori's notion of Uncanny Valley (1970).

While Mori's model mostly focuses on the aesthetics of machines, we are interested in exploring the level of resemblance between attitudes of humans and machines. As particular actions can reflect particular attitudes, we want to explore what type of machine actions can reflect human attitudes, and in doing so try to create a strong level of identification, while not trying to be a human.

The *Poet on the Shore* was designed to envision a scenario in which a robot would be imbued of poetic sensibility. The autonomous robot that roams on the beach, enjoys watching the sea, listening to the sound of waves lapping on the beach, the murmurs of the winds, children's conversing, and the incessant din of seabirds. Most of the time, the robot roams alone to listen and feel. Sometimes, it writes verses into the sand, and watches the waves washing them away.

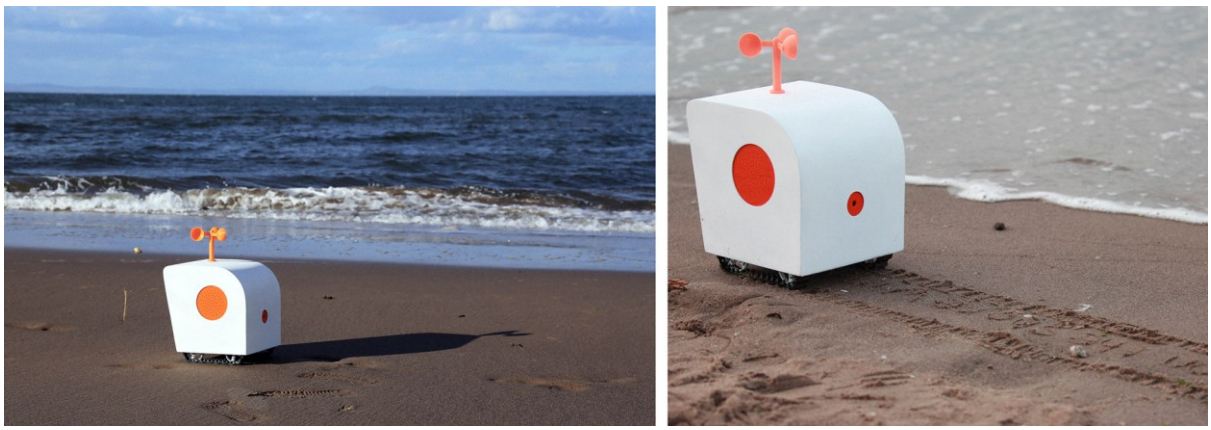

Figure 8 Poet on the Shore: an autonomous robot that roams on the beach.

The robot has a number of sensors that enable it to sense the world around it: the sea, the wind, the sounds etc. Empowered by machine learning, it can discover patterns, and create associations in its mind. Furthermore, it translates these perceptions into poems and write them on the beach. The robot, thus, intervenes in the world and has multisensory experiences. Its behaviour does not

require the intervention of humans. These interventions, expressed through the kinetic and poetic gestures, reveal its non-utilitarian existence: the words it writes will eventually be washed away by the waves or winds.

Poet on the Shore is an attempt to facilitate reflection on alternative values of machines and emotive response. It was important for us that the robot did not have a specific function. Rather than serving humans in an utilitarian way, it performs actions that are typical of humans and potentially many other beings. By not serving or trying to be among humans the robot is kept in its integrity. The intention is that this behaviour would lead to a sense of identification and solidarity, which could lead to recognition of agency, sensibility, and even rights of machines.

6. Discussion and future work

In this paper we present practical explorations that demonstrate three main challenges of designing machines with autonomy. The first challenge is to simply acknowledge that machines may have social organisations that are independent from humans. In the *Gatekeeper on a Mission* we found useful to establish analogies with existing forms of social organisations, illustrating a parallel world of machines that reflects power relationships among humans and often also among other species. While computer scientists have been considering the autonomy of algorithms for a while (since for instance Conway) indeed leading to technologies that have are increasingly introduced in products and applications, the consideration of such applications in design, as explained in Section 3, is relatively recent. A useful exercise for future work is to consider other forms of organisations, such as chemical and geological. This shift in awareness has the potential to lead to forms of design that not only extend the realm of human action but also the lifetime of humans expanding designers awareness into deep or micro temporalities.

The second challenge is to understand how designers could support relationships of interdependence between humans and machines, beyond notions of dominance. Rather than machines serving or overriding humans, they would evolve in parallel to them. While there has been great discussion on the ethics in computing, particularly given the influence that they seem to have had in recent elections, there have been little responses to these issues in the design community. Based on Morton's (2017) notion of subscendence we propose to, instead of thinking about big data as an intangible, abstract entities, we consider tangible effects of different entities be that people, machines, or other living beings. Through the *Perception Companion* we attempt to explore, not how machines would influence people or make decisions but how they would learn from people's bias, and how they could challenge these biases by pointing at things in the world. The implementation of such a concept would certainly face additional challenges, but we argue that as machines evolve it is important to consider how they will evolve *with* and not *for* or *by* humans.

This third and uttermost challenge regards the creation of identification and potentially solidarity between humans and nonhumans. Although ANT and OOO offer philosophical frameworks for placing human and nonhuman on the same footing, we recognize that they could provoke uneasiness in terms of alluding to some sort of human "objectification". Morton (2017, p. 12) suggests that we focus on commonalities between humans and nonhumans helps to support solidarity between humans and nonhumans. The challenge is therefore to explore ways in which the importance of objects would be revealed again to nor serve or override humans but to propose commonalities while keeping both parts in their integrity. This could be done in an aesthetic level, as insightfully suggested by Mori in the concept of Uncanny Valley, but also through other human traits. In *Poet on the Shore* we attempted to support identification between humans and nonhumans in a subjective level. Increasingly, not only the appearance but also attitude may lead to identification between people and things. There remains the question of what kind of attitude we as designers would like to support in this context.

Our investigation has focused on particular scenarios and in the relationship between humans and machines. Expanding awareness beyond human-centred design would however also involve

designing for other forms of life. This opens up questions and possibilities of reflecting on how machines would organise themselves, interact, evolve and potentially create identification with everything in the world. Shifting attention from humans to the world opens up space for huge creative potential. Sharing experiences and challenges, while reflecting on future directions can help us designers in this journey.

7. References

Asimov, I. (1950). *I, robot*. New York: Gnome Press.

Automato.farm (2015). Ethical Things. Retrieved from http://automato.farm/portfolio/ethical_Thing/

Bogost, I. (2012). *Alien Phenomenology, or What It's Like to Be a Thing.* Minneapolis, MN: University of Minnesota Press.

Bogost, I. (2009, December 08). What is Object-Oriented Ontology? Retrieved from http://bogost.com/writing/blog/what_is_objectoriented_ontolog/

Cockton, G. (2004). Value-centred HCI. *Proceedings of the third Nordic conference on Human-computer interaction - NoediCHI 04,* 149-160.

Conway, J. H. (1976). *On Numbers and Games.* London: Academic Press.

Craddock, F., Verma, A., & Liston, M. (2015). *Aspirational Lamp*. Retrieved from http://www.feildcraddockdesign.com/aspirational_lamp.html

Cressman, D. (2009). A Brief Overview of Actor-Network Theory: Punctualization, Heterogeneous Engineering & Translation. ACT Lab/Center for Policy Research on Science & Technology (CPROST) School of Communication, Simon Fraser University, Canada.

Cuzzillo, T. (2015). *Real-World Active Learning: Applications and strategies for human-in-the-loop machine learning.* O'Reilly Media.

Dourish, P. (2004). *Where the Action Is: The Foundations of Embodied Interaction.* Cambridge, MA: MIT Press.

Evans, D. (2011). The Internet of Things: How the Next Evolution of the Internet Is Changing Everything. Cisco Internet Business Solutions Group.

Floridi, L., & Sanders, J. (2004). On the Morality of Artificial Agents. *Minds and Machines, 14*(3), 349-379. doi: 10.1023/b:mind.0000035461.63578.9d

Heidegger, M. (1947). Letter on humanism. In Philosophy in the twentieth century, ed. W. Barrett, and H. Aiken (trans: Lohner, E.), 270-302. New York: Random House, 1962.

Himma, K. E. (2009). Artificial agency, consciousness, and the criteria for moral agency: What properties must an artificial agent have to be a moral agent?. *Ethics and Information Technology, 11*(1), 19-29.

Hyslop, A. (2005, October 06). Other Minds. *Stanford Encyclopedia of Philosophy*. Retrieved from: https://plato.stanford.edu/archives/spr2016/entries/other-minds/

Latour, B. (2005). *Reassembling the social: an introduction to actor-network theory.* Oxford: Oxford University Press.

Latour, B. (1993). *We Have Never Been Modern*. Cambridge, MA: Harvard University Press.

Latour, B. (1998). On Recalling ANT. *The Sociological Review, 46*(S), 15-25. doi: 10.1111/1467-954x.46.s.2

Lee, P. (2016, March 25). Learning from Tay's introduction. Retrieved from https://blogs.microsoft.com/blog/2016/03/25/learning-tays-introduction/#sm.0000txydt7rn1d0bwqo1wu2r29zx4

Mori,. M. (1970). The Uncanny Valley. *Energy, 7*(4), 33-35.

Morton, T. (2017). *Humankind: solidarity with non-human people.* London: Verso.

Nabaztag & Cie. (2005). Retrieved from http://nabaztag.com/#

Norman, D. A. (2013). *The Design of Everyday Things.* New York: Basic Books.

Pearce, W. B., & Littlejohn, S. W. (1997). *Moral Conflict: When Social Worlds Collide.* Thousand Oaks: Sage.

Pschetz, L., Tallyn, E., Gianni, R., & Speed, C. (2017). Bitbarista: Exploring Perceptions of Data Transactions in the Internet of Things. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI 17.*

Rahwan, I. (2016, August 12). Society-in-the-loop: Programming the Algorithmic Social Contract. *Medium.* Retrieved from https://medium.com/mit-media-lab/society-in-the-loop-54ffd71cd802

Rebaudengo, S. (2012). *Brad the toaster.* Retrieved from http://www.simonerebaudengo.com/#/addictedproducts/

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge, UK: Cambridge University Press.

Rifkin, J. (1999). *The Biotech Century.* New York: Penguin Putnam.

Sterling, B. (2005). *Shaping Things.* Cambridge, MA: MIT Press.
Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M. and Teller, A. (2016) Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence, Report of the 2015 - 2016 Study Panel, Stanford University. Retrieved from https://ai100.stanford.edu/2016-report
Taylor, A. (2009). *Animals and Ethics: An Overview of the Philosophical Debate.* Peterborough, Ont: Broadview.
Thomson, J. J. (1976). Killing, Letting Die, and the Trolley Problem. *Monist, 59*(2), 204-217. doi: 10.5840/monist197659224
Verbeek, P. (2009). Cultivating Humanity: toward a Non-Humanist Ethics of Technology. *New Waves in Philosophy of Technology*, 241-263. doi: 10.1057/9780230227279_12
Waal, Frans B. M. (2012). *The Age of Empathy.* London: Souvenir.
Waelbers, K., & Dorstewitz, P. (2013). Ethics in Actor Networks, or: What Latour Could Learn from Darwin and Dewey. *Science and Engineering Ethics, 20*(1), 23-40. Doi: 10.1007/s11948-012-9408-1
Wang, Y. (2016). The Power of Human-in-the-loop: Combine Human Intelligence with Machine Learning. *TechNet Blogs.* Retrieved from https://blogs.technet.microsoft.com/machinelearning/2016/10/17/the-power-of-human-in-the-loop-combine-human-intelligence-with-machine-learning/
Washburn, R. (2013). The Social Significance of Human Biomonitoring. *Sociology Compass, 7*(2), 162-179. doi: 10.1111/soc4.12012.

About the Authors:

Yuxi Liu is an interaction designer who is interested in the intersection of design, people, and technology. She studied Design Informatics at the University of Edinburgh, exploring diverse affordances of emerging technologies and probing the social, cultural, and ethical issues brought by them. She likes robots.

Larissa Pschetz is an interaction designer, researcher and lecturer based in the Centre for Design Informatics at the University of Edinburgh, UK. Her work focuses on the role of design in supporting or challenging socio-technological narratives. She likes robots.