Robotics and Personhood: Towards an Ethical Experience-Centred Design

Expert Statement: Royal Irish Academy Engineering and Computer Science Committee

July 2017

Expert Statement: Royal Irish Academy Engineering and Computer Science Committee

Robotics and Personhood: Towards an Ethical Experience-Centred Design

Recent advances in robotics have in many ways enhanced the human experience, improving medical care, providing a safer work environment and bringing novel educational interventions, for example. But the line between what we allow robots to do and where we give primacy to human activity and experience is an ethical matter that may be best addressed through policy; as yet no such policy framework exists. In this statement, we argue that, as we become more accustomed to intelligent technologies, it is worth considering the risks, as well as the benefits, that the design and everyday use of such technology may pose for our personhood – our interpretation of ourselves and others. The ubiquity of computers, robots and intelligent devices, in our working and social lives, which are increasingly designed to function as a person, indicates a need for an agreed framework supported by international legislation that would clearly delineate ethical boundaries (see Sharkey, 2008). Here, we argue for the need for such a framework, using a focus on 'social' robots, such as service and care robots in the workplace and home, to demonstrate the potential for robot-human conflict. It is here that design practice has the potential to greatly enhance personhood or detract from it.

The rise of robots in the workplace

Robots in the workplace have taken over dangerous and difficult jobs, but have also displaced workers more widely in manufacturing and many other areas of work. Frey and Osborne (2013) estimate that 47% of total US employment is at risk and warn that 'as technology races ahead, low-skill workers will reallocate to tasks that are non-susceptible to computerisation … For workers to win the race, however, they will have to acquire creative and social skills' (p. 45).

Even if this percentage overestimates the risk, the case for a high volume of job displacement is convincing. The argument follows that in order to avoid future generations of high unemployment, a significant educational response is needed, emphasising skills that are distinctly human. But, on its own, this is a fatalistic response, accepting with little questioning the inevitability of technological unemployment (see Ford, 2015 and Kaplan, 2015). While such a response may spark innovative economic approaches to ensuring income for those who are not employed or those who are under-employed, and focus the economy on serving public interest rather than the other way around (see for example Kaplan, 2015), it underplays the personal and societal value of satisfying and suitable work.

In addition to industrial robots, we are seeing greatly increasing sales of service robots for professional use, with use of defence systems (e.g. unmanned air systems), medical robots, rescue and exploratory robots, as well as privately owned service robots (vacuum cleaners, lawn mowers, educational robots, etc.) developing rapidly. Leaving aside military applications and the threat they pose, by far the most important segment of this market in terms of the relationship between information technology and personhood concerns what have been called 'care robots'. The fact that we are developing robots to interact closely with, and take care of, children and vulnerable people, in the absence of an agreed policy framework, should be a matter of great concern to us.

‘Care robots’ in the home

Robots designed to entertain children are already available for those that can afford them. Some are capable of recognising voices and faces, can call children by their name and appear to hold a conversation with them. As an educational tool robots are potentially very useful. However, is there a risk that, as such robots become more sophisticated, they might provide part-time (or full-time) replacement of primary carers (Sharkey & Sharkey, 2010)? A child’s cognitive and linguistic development, emotional adjustment and social attunement, his or her burgeoning ‘theory of mind’ (the ability to take another person’s perspective and to appreciate that person has a mind with thoughts and emotions of his or her own), all depend on interactions with people – other children, adults and importantly ‘more knowledgeable others’ (Vygotsky, 1930/78).

For a child, the experience of being cared for by a robot responding contingently to the child’s actions with touch, speech or emotional gesture is qualitatively different from the experience of being cared for by a person responding sensitively and appropriately. When a person responds to a child who has fallen or is frustrated, they do so with empathy because they too have experienced pain and frustration. Their response comes from the recognition and imagination of shared experience, a deeply contextualised sensitivity that people utilise with skill from a young age. No matter how refined the robot’s ability to classify and respond with a matching expression, it is not a response to the cause of the cues. A response to the cause of the cues is integral to enabling a child to take the other’s perspective.

The cognitive skill that enables even a young child to appreciate the difference between a contextually sensitive response and a solely contingent one may also support the child in understanding the difference between make-believe and reality. Some of the qualities that are being built into childcare robots’ interactions to make them more engaging for children may blur the boundary between real and make-believe. As Sharkey and Sharkey (2010, p. 171) note, ‘an infant entertaining a relationship with a robot may not be in a position to distinguish this from a relationship with a socially and emotionally competent being.’ The effect on development is unknown, and yet there is an increasing emphasis on ‘social robots’, including as companions (see Richardson, 2015).

Robot companions

Recent years have seen rapidly increasing availability of hi-tech robots to assist older people as well as a very large, rapidly expanding market in telecare, memory devices and other assistive technologies, often designed for people diagnosed with dementia. These areas receive significant investment from national and international research councils, yet no policy framework is in place that would require key social and ethical questions to be addressed in the design of such systems.

The motivation for developing these robots is to help older people maintain independence, often in their own homes. An important consideration, however, is an appreciation of the experience of being old, having dementia, or caring for someone diagnosed with dementia, an appreciation that relies on human contact and would result in a quite distinctive person-centred approach to designing support for older people.

Kitwood (1997, p. 8), the founder of person-centred care for people diagnosed with dementia, defined personhood as the 'standing or status bestowed upon one human being by others in the context of relationship and social being.' For Kitwood, the personal and social identity of the person with dementia arises out of what is said and done, and responsive communication between people is the experience out of which personhood emerges. This definition recognises the importance of autonomy as well as the value of social relationships, their interconnectedness and interdependence in preserving personhood. This must be a key consideration before the assistive technology is built; experience and personhood have to be a key focus of the negotiation of technological imaginaries and concepts at the very earliest stages of design.

Public policy and experience-centred design

What we allow robots to do and where we give primacy to human activity and experience is clearly a matter for policy. This has recently been indicated by the European Parliament Committee on Legal Affairs' (2016) call for a guiding ethical framework for the design, production and use of robots and the IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems (2017) the purpose of which is to 'ensure every technologist is educated, trained, and empowered to prioritize ethical considerations in the design and development of autonomous and intelligent systems'. An immediate focus for policy should be a clear articulation of the values of robotics design practice, an ethical approach to which would be informed and motivated by a concern for human experience and the ways in which people make sense of it.

We finish with two examples of (the many) areas in which design of robotics could give primacy to human activity and experience.

The first is privacy. Privacy, like many of the human concerns that ought to be considered when designing robots, is culturally and situationally sensitive. What is intrusive in one culture or context may not be in another. Privacy is a key value for people, and is influenced by quite sophisticated calculations and trade-offs. For example, most people would consider surveillance of a young child via a baby monitor to be appropriate; the need for safety overrides the child's right to privacy. But what of an older child interacting with a robot, who is unaware that he or she is being covertly monitored via a camera? There is no simple design approach to giving primacy to a need for privacy, but a variety of design approaches that have emerged with respect to Smart Cities may be differentially relevant. They include transparency through, for example, notification and consent, accountability and respect for the boundedness of agreements (implicit or otherwise) to compromise on privacy.

The second example is care, which is a process of protecting and looking after the needs of another, and often requires intimate communication of needs, concerns and recognition. Care requires an empathic relationship in which people mutually accept each other and are responsive to each other's feelings. That does not mean that robots cannot be used to entertain, stimulate, educate, protect or foster connections between people. Rather, their use in these contexts should be subject to an overriding concern for the primacy of human contact in care, the minimum requirement to sustain personhood.

We are beginning to see a growing and welcome focus on the ethics of robotics and on the user-perceived quality of interactions with robots. Over time, and with political will, public policy and law may provide a protective framework for human–robot interaction. Over time, people will also learn about the potential and the limitations of social robots. In the meantime, if only to ensure that the potential for beneficial development of robotics is not hampered, robotics design should give due weight to personhood, human activity and experience.

References

European Parliament Committee on Legal Affairs (2016) Draft report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL). Available at: http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML%2BCOMPARL%2BPE-582.443%2B01%2BDOC%2BPDF%2BV0//EN

Ford, M. (2015). *Rise of the robots: technology and the threat of a jobless future. Basic Books*.

Frey, C.B and Osborne, M.A. (2013). *The future of employment: How susceptible are jobs to computerisation?* Oxford Martin Programme of the Impacts of Future Technology. http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems (2017). https://standards.ieee.org/develop/indconn/ec/autonomous_systems.html

Kaplan, J. (2015). *Humans need not apply: a guide to wealth and work in the age of artificial intelligence.* Yale University Press.

Kitwood, T. (1997). *Dementia reconsidered: the person comes first.* Open University Press.

Richardson, K. (2015). *An anthropology of robots and AI: annihilation anxiety and machines.* Routledge.

Sharkey, N. (2008). 'The ethical frontiers of robotics'. *Science* 19 December 2008 (322: 5909), pp. 1800–01. DOI: 10.1126/science.1164582

Sharkey, N.E. and Sharkey, A.J.C. (2010). 'The crying shame of robot nannies: an ethical appraisal'. *Interaction Studies*, 11:2, pp. 161–90.

Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press. Original work published in 1930.

The Royal Irish Academy's Engineering and Computer Science Committee acknowledges and thanks the authors of this expert statement; John McCarthy (University College Cork) and Fiona Lyddy (Maynooth University) who is a member of the Committee. The views expressed herein are not necessarily shared by individual Members of the Academy.

The Engineering and Computer Science Committee also acknowledges and thanks the contributions of the anonymous reviewers of this expert statement.

For further information please contact the Royal Irish Academy at info@ria.ie

Royal Irish Academy
19 Dawson Street
Dublin 2
D02 HH58

info@ria.ie
00 353 (0) 1 676 2570

www.ria.ie