

This is Krístrún's draft copy, freely available for fair use, (see http://en.wikipedia.org/wiki/Fair_use) !

Published version is in *Studies in Ethics, Law, and Technology* 4(3): Article 1

Editorial: The convergence of the physical, mental and virtual

Guest editor: Krístrún Gunnarsdóttir
Cesagen, Lancaster University

This editorial introduces a special issue of *Studies in Ethics, Law, and Technology*, centring on the convergence of the physical, mental and virtual. The idea of publishing a special issue on this matter came about at a conference, *ICT that makes the difference*, organised by the consortium of a FP7-funded project, ICTethics.¹ In particular, we wanted to foreground some of the material presented and debated in sessions on the role of assistive robotics, the use of RFIDs and other implants for brain/body-device interactions, and issues surrounding ‘medical access to the brain’.

The special issue takes as its point of departure the gap that exists between the visionary work and experimentation undertaken by scientists, and the results of theoretical and practical reflection on issues of ethical, legal and social relevance. One of the objectives of the ICTethics project is to investigate how ELSA studies can be operationally embedded in the early stages of ICT design and development, as well as in agenda setting for S&T research. But to what extent do scientists, policy-makers, ELSA scholars and other stakeholders network and communicate to bring about improved conditions for good governance and professional accountability? The special issue brings together cutting-edge experimenters, philosophers and ELSA scholars, as both authors and commentators, to explore some of the latest developments that manifest convergence of the physical, mental and virtual, and relate them specifically to issues of selfhood, identity and responsibility, empathy, medical ethics, social robustness and accountability. In doing this, we hope to set an example of how radically different disciplines can communicate and complement each other’s work.

Robots as pets and friends?

A policy perspective published by the European Commission (European Commission, 2008) states that “[t]omorrow’s robots will not be confined to industry, but work in the ‘real world’, providing solutions for many societal issues”. Indeed, we learn from EU policy frameworks such as the i2020 initiative (European Commission, 2010) that current societal challenges will be met with a wave of innovations to which new and emerging ICTs are essential, including robotics.

¹ The international conference, *ICT that makes the difference*, took place at Hotel Métropole in Brussels, 22-25 November 2009. ICTethics is funded under FP7 Science-in-Society-2008-1.1.2.1, Ethics and new and emerging fields of science and technology (Capacities Programme), contract nr. FP7-230368.

Robots will be assistants, co-operating with people in everyday environments, and in ways that are sensitive to ordinary goings-on—even helping to optimise social decision-making (see also EUROP, 2009).

We might look forward to sharing our lives with perceptive and cognitive robots, with integrated sensors and free-range mobility. Engineers have come under growing pressure to develop machine intelligence that is more *in-hand* and *at-home* in both production and common everyday affairs. Representative models of mind gave way to models of neural networks and reinforced learning, complex sensory technologies, embodied intelligence and, more recently, models of social intelligence and affective experience (e.g. Breazeal, 2003). The last on the list serves the attempt to embed artificial agents in emotionally sensitive and socially relevant activities.

Although the robots we see in films such as *iRobot* and *Terminator* are not on the horizon, a significant development in recent years turns on the companionship we may have of robots as pets or friends. For example, ‘Paro’ is a therapeutic robot seal, shown to reduce patient stress, improve motivation, and stimulate interaction between patients and their caregivers.² Paro is promoted as a companion to the elderly and those suffering from dementia or autism. But what does ‘companionship’ stand for in human-robot relations? This question is explored by Mark Coeckelbergh in the special issue where he challenges the objection that such companionships are deceptive—that a robot is always merely a machine and cannot meet social and emotional needs. Accordingly, to rely on robots in the care of vulnerable people would be morally apprehensible (see e.g. Sparrow and Sparrow, 2006 on the future of aged care). What we learn from Mark’s argument is that human-robot relations can be seen as something more akin to the relationships children form with dolls and teddy bears, or adults with their pets and favourite gadgets—relations that mirror our own vulnerabilities, evoke empathy and a sense of ‘fellow feeling’, however, we should be careful not to believe that such relations can replace the company of other humans.

The human-machine merger

Two of the articles in the special issue are written by scientists, Kevin Warwick and Rolando Meloni, et al., who make it clear that they are first and foremost experimenters in their respective fields. They are not ethicists, legal scholars or sociologists. They agreed to describe the experiments they are involved in, and explain the latest trends in their fields to the readers of the journal. They express the understanding that reflection on the ethical, legal and social implications of the work they do is both timely and necessary but, also, that they themselves may not be best suited to identify potential complications. Therefore, their papers are complemented with comments by relevant experts.

Most of the special issue is dedicated to these exchanges, dealing with what Warwick calls

² See <http://www.parorobots.com/> .

‘human-machine merger’. This is a broadly defined category, involving the development of human-machine interfaces to directly insert and/or extract signals. Warwick explains to the reader how brain cells are grown on a computer chip, fitted to control a simple robot body. We can ask what the limitations are to the size and capacity of such a brain, say, if the brain cells were human and the robot body equipped with cutting-edge sensory technologies. He also discusses the use of RFID tags for the tracking and positioning of persons, and experiments with active implants to control objects and devices, and modify brain functions. For example, neuro-degenerative diseases are dealt with by targeting brain cells with electromagnetic signals such as in the treatment of Parkinson’s disease. Can these deep brain stimulations be controlled in a ‘smarter’ way, he asks. Meloni, et al., also explain in the special issue how deep brain stimulation could be complemented with the insertion of genetic material to restore or prevent further degeneration in the brain. They aptly note however, that direct access to the brain for medical purposes is still very complicated and far from risk free, even if scientists have proof of principle and some success. Similarly, Elisabeth Hildt draws attention to the experimental character of current methods to extract brain signals for medical purposes. We can make use of brain activity by extracting signals from scalp readings, she explains. These are non-invasive techniques offering unique opportunities to persons with motor impairments to control mechanical devices and better communicate with those around them. But, even if scalp readings are making their way into computer gaming experience, they are still applied in controlled environments for medical purposes and riddled with complications. The invasive technologies, brain implants, are far less developed but the attraction is that they promise to deliver much higher resolution of brain activity, taking human-computer interaction to a whole new level, prosthetic limb control for example.

Hildt has had the opportunity to work in clinical and experimental settings and gives an account in her article of common ethical concerns. Apart from the multitude of potential physical complications when devices are implanted, there are questions of how to obtain informed consent, say, from ‘locked-in’ persons. There are significant risks of psychosocial affects relating to disappointment and frustration when a brain-computer interface fails to deliver in spite of extensive training, when devices and systems are withdrawn after a successful trial, and so on. There are questions of hope and promise, and the extent to which quality of life can actually be improved with more independence, privacy and social participation. James Giordano also takes medical ethics as a point of departure in his comment on the work of Meloni et al. In addition to the concerns raised by Hildt, he raises the question of how to establish the legal ramifications and how to address issues of justice: “how shall decisions of which patients receive these state-of-the-art technologies be addressed, deliberated and articulated?”

Together, these questions draw attention to the manner in which benefits, risks and burden are estimated and articulated in experimental and clinical practice—in communication among experts, with policy-makers, patients, and other relevant stakeholders. They point to a pervasive problem of how to regulate new and emerging technologies in ways that can clarify legal ramifications and

improve professional and regulatory accountability. Furthermore, there remain unresolved questions of how to distribute access to treatments that are most likely to be beneficial and how to avoid the most vulnerable persons from receiving risky treatments in the name of ‘empowerment’. But the opportunities that lie in convergent technologies, such as human-machine mergers and the use of implants, go well beyond clinical purposes (Nordmann, 2004; Kjølberg et al, 2008). There are further ethical, occupational and organisational challenges (e.g. Rodotà and Capurro, 2005 on the ethical aspects of implants). For example, Warwick suggests that children or those with dementia could be micro-chipped for safety and Juliet Lodge, in her comment, reminds us that there are always some persons relying on someone else, some RFID reader, some other machine, to respond to warning signs. Despite all good intentions, there is no guarantee that technical solutions like implants will deliver better safety or better organised service, nor that they will not be used for malign purposes. We need to think through the challenges in a socially robust manner, sensitive to the ways in which we know humans to be resourceful and competent, but also lazy and inconsistent—sensitive to questions of responsible and sustainable innovations, taking into account the occupational and organisational challenges they engender, as well as new powers and economies.

Finally, different methods of directly accessing the brain suggest new capabilities which can radically change what it means to be human, to have a self and identity, to be liable for one’s actions, and so forth. The classical line of reasoning here rests on an ancient notion of humans as bounded entities with inherent attributes. Conversely, the problems raised by merging humans and devices centre on questions of how much human agency can be extended to devices, for what purposes we want to enhance human capabilities, if the individual is no longer responsible for her actions when a device overrides her will or the self no longer that self. Another way of looking at these developments is to say that new hybrid agencies emerge when devices and humans are intimately interdependent while performing certain tasks. Accordingly, any new capabilities for action or human enhancement more generally, will have to be assessed in their own right. For example, Ruth Chadwick asks in her comment on Warwick’s work whether active implants, enabling direct communication between two human nervous systems, is necessarily an improvement of ordinary communication and how we usually get to know each other. Assessing improvement “may involve a moral or aesthetic judgement” rather than a judgement strictly about effectiveness. Søren Holm and Teck Chuan Voo also underscore the differentiation between philosophical speculation about self and individual responsibility, and how we might justify compensations if hybrid agents offend. If necessary, recompense is established in legal proceeding. It is negotiated with reference to normative judgements, reasonable certainties and doubts about the ordinary behaviours of humans and things in particular circumstance. As Mireille Hildebrandt has pointed out, we inscribe laws into computational functions and further developments in that direction may spell the end of ‘law’ as we know it (Hildebrandt, 2009). Computationally inscribed laws make certain tasks possible and other tasks impossible. Laws as we ordinarily know them

however, can always be broken and the court of law has always been an elbowroom for negotiation and mitigation. Legal proceedings and judgement are not an exact science but a space within which reasonable certainties and reasonable doubts are established.

As we can see, enormous opportunities lie in the latest developments which are grouped here under convergence of the physical, mental and virtual. We see opportunities to restore or modify normal human functions, to support communication and companionships, and sort out matters of safety and security. The special issue only begins to explore some of the implications for individual and social well-being. We believe there is a need to do much more in order to ensure responsible and sustainable innovations in this area, for which we hope this publication makes a contribution.

Bibliography

- Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems* **42**(1). pp. 167-75.
- 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).
- European Commission (2010). *Europe 2020: a strategy for smart, sustainable and inclusive growth*. Communication of the European Commission.
- European Commission (2008). *ICT Research. The Policy Perspective: European robots, getting smarter, safer, and more sensitive*. European Commission. <http://cordis.europa.eu/ictresults>
- Hildebrandt, M. (2009). Technology and the end of law. In E. Claes and B. Keirsbilck (eds) *Facing the limits of the law*. Heidelberg: Springer. pp. 443-64.
- Kjølberg, K., Delgado-Ramos, G. C., Wickson, F., et al. (2008). Models of governance for converging technologies. *Technology Analysis & Strategic Management* **20**(1). pp. 83-97.
- Nordmann, A. (Rapporteur) (2004). *Converging Technologies – Shaping the Future of European Societies*. European Commission - HLEG Foresighting the New Technology Wave.
- Rodotà, S. and Capurro, R. (2005). *Ethical aspects of ICT implants in the human body: Opinion produced on the direct initiative of the EGE*. European Group on Ethics in Science and New Technologies to the European Commission.
- Sparrow, R. and Sparrow, L. (2006). In the hands of machines? The future of aged care. *Mind and Machines* **16**(2). pp. 141-61.