Technological Forecasting & Social Change 84 (2014) 74–85

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

The empathic care robot: A prototype of responsible research and innovation ☆

Bernd Carsten Stahl *, Neil McBride, Kutoma Wakunuma, Catherine Flick

De Montfort University, Centre for Computing and Social Responsibility, The Gateway, Leicester LE1 9BH, UK

ARTICLE INFO

Article history:
Received 1 November 2012
Received in revised form 24 May 2013
Accepted 4 August 2013
Available online 5 September 2013

Keywords:
Responsible research and innovation
Creative prototyping
Robotics
Ethics
Emerging technology
ICT

ABSTRACT

Science fiction prototypes are often used to visualise or represent novel technologies or other techno-scientific innovations. The present paper follows this tradition and describes a prototype of a care robot that is endowed with affective capabilities. The paper describes some of the potential ethical problems arising from such a technology. This aspect of the paper is based on prior research in a European-funded technology foresight project that explored the ethical issues of emerging ICTs. The paper goes beyond the description of technical innovation and its ethical consequences. The recognition of the ethical relevance of research and innovation has spawned a discourse around responsible research and innovation. The paper draws on this discourse, which aims at anticipatory technology governance to ensure the social acceptability and desirability of technologies. The prototype vignette of the paper explores how responsible research and innovation could be realised in practice and how it could be used to address ethical issues such as those of affective care robots. The paper reflects the likely controversies that responsible research and innovation is likely to create and it uses the ethical dilemma of the care robot to draw the reader's attention to possible theoretical and practical conclusions.

1. Introduction and theoretical context

The development of products and services normally has a specific and clearly defined aim. At the same time it tends to lead to unexpected consequences and can have predictable negative consequences. This realisation is not new and has led to much discussion on how and when such consequences can be identified and addressed. A key contribution to this discussion was made when Collingridge [1] pointed out that there is a dilemma because at early development stages consequences are difficult to predict whereas at later stages where consequences become clearer the trajectory of the development becomes more difficult to change.

This question of how we can identify possible problems arising from scientific and technological research, development or innovation has led to numerous answers. These include professional bodies with specific regulations, codes of ethics, technology assessment, and participative design approaches to name just a few. Presently this debate focuses on the concept of responsible research and innovation (RRI). While RRI is not yet a clearly defined concept, the discourse surrounding it renders it clear that there are some core drivers, issues, and themes that fall under this idea [2–5]. The importance of this debate around RRI for the present paper is that it contains many ideas that are likely to influence the way in which we will in the future evaluate scientific and technical research and innovation. Understanding them is therefore an important part of future casting, of creating a language and a shared understanding of possible futures in order to engage in a discourse about which future we want, which is, according to Johnson [6], the purpose of prototyping.

This paper starts from this idea and develops a radio play that focuses on the way in which concepts of responsibility

* Corresponding author. Tel.: +44 116 207 8252.
E-mail address: bstahl@dmu.ac.uk (B.C. Stahl).

0040-1625/$ – see front matter © 2013 The Authors. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.techfore.2013.08.001

may be employed in future research and innovation. It does this by telling a story of robot development in the medium term future, maybe around the year 2030. The setting of the story is the development of a novel type of care robot which has emotional capabilities. This technical innovation is the main theme of the story. The second focus of the narrative, however, is the processes of RRI that are interwoven in the storyline.

Both the technical description and the RRI aspects arise from a recent European foresight study that looked at "Ethical Issues of Emerging ICT Applications" (ETICA, www.etica-project.eu). This project identified a number of emerging information and communication technologies (ICTs) as well as ethical issues these are likely to raise. On the basis of this identification and a subsequent expert-based evaluation exercise, ETICA developed a set of recommendations for policy makers and ICT researchers and developers [7,8]. Here, we build on these in two ways. On the one hand this research informs the description of the novel type of care robot (MARTA) which is endowed with many of the technologies that are currently being researched. On the other hand the radio play assumes that the recommendations of the ETICA project have been implemented and that this structures the way in which researchers engage with innovation. The purpose of the short story is to envisage how RRI could play out in the future. This is a key question for the special issue because this science fiction (SF) prototype will "let us imagine the future, […] think through the ethical implications of technologies, play with possible benefits, explore possible tragedies and ultimately engage in a deeper conversation about science, technology and our future"[6].

2. Background to the radio play vignette

The background section covers two main sub-sections: first, it discusses the current research on ethical issues of emerging information and communication technologies as explored during the ETICA project. It then moves on to a discussion of the current discourse around RRI. Both of these are key to the following vignette that reflects aspects of the RRI review of the new care robot that displays many of the technical capabilities explored during the ETICA project.

2.1. Ethics of emerging ICTs: the ETICA project

The ETICA project (Ethical Issues of Emerging ICT Applications, www.etica-project.eu) was a research project funded under the EU's 7th Framework Programme. It ran from 2009 to 2011. Its aim was to identify emerging information and communication technologies (ICTs), the ethical issues these are likely to raise, evaluate those and develop governance recommendations to help address these. The project can best be understood as a foresight project [9,10] that aimed to explore possible futures to allow the development of appropriate policy. As a result of the identification phase, ETICA developed detailed descriptions of 11 emerging ICTs (see deliverable 1.2 "emerging technologies" on the project website for detail and see [8] for details on methodology). By an emerging ICT the project meant a socio-technical system that had the potential to substantially affect the way in which humans live their lives.

The one technology that is most prominently used in the technology prototype developed below is that of affective computing, sometimes also called emotional computing [11–13]. There are three main aspects of affective computing. One is the recognition of human emotions, the second is the expression of emotions in ways humans can understand and the third is the modelling of emotions. Current technology has made considerable progress in the first two areas. The vignette in this paper revolves around the third one, where the scientists develop an "empathy machine", which contains a model of empathy and which also causes the main problems of the new robots. In addition to affective computing, the vignette implicitly includes several other emerging technologies, notably robotics, artificial intelligence and ambient intelligence. It would be easy to develop the story so as to include further technologies but for the sake of simplicity we have refrained from doing so.

During the ETICA project, each of the technologies we identified was investigated to find out whether there are specific ethical questions that are likely to arise from the technology and that the current literature on ethics and ICT has discussed. For affective computing the issues that were identified were conceptual questions, for example what counts as an emotion, questions of measurement and error, problems arising from persuasion and coercion, the appearance of computers as humans (anthropomorphism), questions of privacy, cultural differences with regard to emotions, questions of informed consent and responsibilities arising from affective computing. Many of these play a role in the vignette. (For a detailed discussion of the ethics of affective computing as well as the ethical aspects of the other emerging ICTs as identified by the ETICA project, please see deliverable 2.2, the Normative Issues Report, which can be downloaded from the ETICA website.)

In addition to the specific ethical issues associated with the individual technologies, the ETICA project furthermore explored shared or overlapping ethical issues that are related to several of these technologies. One can distinguish between those ethical issues that are well discussed and those that are less obvious [7]. Obvious issues are those that are already regulated by legal or other means such, most notably privacy and questions of intellectual property. Less widely explored issues include questions of individual and collective identity and changing patterns of social interaction and culture due to novel ICTs.

The ethical issue directly related to the "empathy engine" described in the vignette is relatively straightforward. It relates to the question what counts as empathy and an apparently simplistic implementation of the engine. This ethical issue is embedded and surrounded by other social and ethical issues, ranging from competing responsibilities (the role responsibility of the scientist vs. his responsibility towards his family) to questions of life and death and assisted suicide. The ethical issue at the heart of the vignette is thus not a general and easily identifiable one but highly context dependent.

This leads us to the final part of the ETICA project, namely the question of how ethical issues can be addressed proactively. The project issued a set of recommendations to policy makers and to ICT researchers. It was recommended to policy makers to institute an ethics impact assessment across ICT research, to develop an observatory for ethics and ICT and to establish a forum for stakeholder involvement. Researchers and research institutions were asked to embrace ethics as a positive factor and to implement processes that would allow them reflexivity in projects. These recommendations can be seen as an attempt to contribute to responsible research and innovation in ICT, which is a broader discourse worth discussing in more detail.