

Operationalising Design Fiction for Ethical Computing

Joseph Lindley
HighWire Centre for Doctoral Training
Lancaster University
United Kingdom
joseph.lindley@gmail.com

Dhruv Sharma
HighWire Centre for Doctoral Training
Lancaster University
United Kingdom
d.sharma2@lancaster.ac.uk

ABSTRACT

Design fiction is a type of speculative design, where story worlds are crafted to then be used as a canvas upon which so-called diegetic prototypes can be sketched [9]. Because these prototypes exist only within story worlds they are not constrained by currently available technology; because of this design fictions are excellent means to open up space for critical conversations about the future [2,6]. This project experiments with using design fiction as a novel way to explore the complexities of technology and ethics. We focus on one specific case to demonstrate the method we adopted, however the contribution is general in nature and may be applicable to other cases too. The work consists of two parts, this paper and a ‘design fiction documentary’ film, ‘Care for a Robot’ [12]. The paper and film are intended to be viewed together.

Categories and Subject Descriptors

K.4.1 [Computing Milieux]: Public Policy Issues – *ethics*.

General Terms

Design, Experimentation, Human Factors

Keywords

Design fiction, domestic robots, care for the elderly, radical digital interventions, accessible ethics.

1. INTRODUCTION

We are Dhruv Sharma and Joseph Lindley, we are both doctoral students at the HighWire Centre (Lancaster University). Because of the novel format of this work we have included this section to provide some context and make clear what our personal interests are, why we are doing this research, and how we think it relates to ETHICOMP. Joseph is researching the relatively immature concept of design fiction, he’s interested in understanding what design fiction’s kernel is and the range of ways it can be used. Dhruv is researching loneliness among the elderly. In particular his research is interested in how ‘radical and digital interventions’ [18] may be used to reduce the negative impacts of loneliness among the elderly.

The example case that this work revolves around is domestic care robots. Although not commercially available at present, current

discourse around medical robots designed to care for people leads us to believe that having an accessible and meaningful debate about the ethical implications of these, potentially pervasive, technologies is essential given the breadth of their impact when (or if) they do become viable [14]. The ideas presented here signify our early response to this challenge. How can we prepare for potentially pervasive technologies in the offing? The work isn’t intended to be a manifesto or statement of truth about what ethical stance ‘should’ be adopted vis-à-vis domestic care robots, nor is it intending to posit the ‘best’ method to address the challenge of preparing for an ethical debate around caring robots. Rather it describes the concepts, theory and practicalities behind one possible way of accessing the debate and making it more meaningful. As such we think this approach may be replicable for other cases, and we also see this as a general contribution to studies of computing and ethics.

1.1 Radical Digital Interventions

Sharma et al.’s review of existing age-related loneliness interventions, highlights that the majority demonstrate an incremental approach to addressing the problem [18]. They argue that 1) there are relatively fewer interventions that are ‘radically’ different and that 2) use of digital technology is underrepresented in this area. In order to explore possible strengths - or limitations - of this type of intervention we should pay extra attention to radical-digital interventions and strive toward experimentation and innovation in this area.

The distinction between incremental and radical interventions is akin to *reformist versus radical* departures in environmental discourses [5]. Reformist departures seek solutions within familiar modes of rational management, whereas radical departures argue for a comparatively significant movement away from industrial modes of living and being. Manzini suggests that incremental innovations represent our existing ways of ‘thinking and doing’ whereas innovations falling outside our current ways of ‘thinking and doing’ represent radical innovation [15]. Norman and Verganti define *incremental* innovation as “improvements within a given frame of solutions” and “doing better, what we already do” but describe *radical* innovation as “a change of frame” or “doing what we did not do before” [16].

Improvements upon ‘what we already do’ are usually backed up by reflective practices and learning from past experiences. Radical ‘changes of frame’ however are either the product *of*, or ultimately lead *to*, uncharted territories. It is therefore impossible to predict the ramifications, implications and impact of radical innovation unless we speculate about what forms those innovations may take. Practices such as design fiction offer us with an academically grounded approach to crafting, interpreting, and making sense of these speculations [cf. 1,8,11].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference ’10, Month 1–2, 2010, City, State, Country.

Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

1.2 Design Fiction

There are scant arguments for clearly bounding precisely what design fiction is and how it should be used. It has demonstrably been used as a prototyping tool, research method, ideation aid, and as a communication tool [4,10,13,23]. Design Fictions harness the power of speculative design thinking to holistically imagine how ideas from the present would manifest in the future. Designers and practitioners create design fiction artefacts in a huge range of shapes, sizes and media: film, text, objects, and combinations of all of them [13]. The most popular definition of design fiction refers to the purposeful application of diegetic prototypes to encourage a suspension of disbelief about change [21] (refer to [11,13] for a more in-depth deconstruction of this definition).

Lindley's pragmatics framework for design fiction proposes some categories of design fiction intended to make communications about applications of design fiction clearer. As part of that work Lindley introduces a nomenclature to differentiate between design fictions that are created from scratch (intentional design fictions) and other entities that coincidentally share the properties of a design fiction (incidental design fictions) [13].

'Anticipatory ethnography' proposes using observations of design fictions as part of design ethnography research projects. Where design ethnographers tend to do 'quick and dirty' ethnographic studies of people and places in order to design things better, anticipatory ethnographers might do similarly quick and dirty ethnographic studies, but of the people and places in a design fiction world as opposed to the real world. A straightforward example of how one might use anticipatory ethnography is to take a piece of science fiction cinema that meets the criteria of being an incidental design fiction. Watch the film to take detailed ethnographic notes of the action and situations, and then to apply methods of design ethnography in order to generate actionable insights pertaining to the world and diegetic prototypes depicted in the film. If the film's ability to suspend disbelief with diegetic prototypes is strong, then anticipatory ethnography should generate powerful insights [11].

1.3 Robot and Frank

We cast the 2012 film *Robot and Frank* [17] as a piece of incidental design fiction. Set in an unspecified near future where today's modern hybrid cars are aging and rusty, and the local library is finally withdrawing paper books. The film depicts an elderly man called Frank, his children, and the introduction of a caring robot into Frank's life. Some aspects of how the robot interacts with humans in the film might appear unrealistic, however we argue that on the whole the diegetic prototypes in the film *are* able to suspend disbelief about change, and therefore it passes the test of being an incidental design fiction. A full discussion of what can or cannot be considered incidental design fiction is unfortunately beyond the scope of this paper.

Used as a stimulus, *Robot and Frank* was essential to producing *Care for a Robot*, however it is not necessary to actually watch the film in order to make sense of the work and take some value from it. However, we personally recommend it as being a simple, yet thought-provoking film, and also to further contextualize this

work. Some sequences from *Robot and Frank* appear in our film *Care for a Robot*.¹

1.4 Care for a Robot

This work is slightly unusual in that it has a two-dimensional relationship with design fiction. It *extends* the incidental design fiction that is *Robot and Frank*, in order to then *create* an intentional design fiction, *Care for a Robot*. Furthermore the format of *Care for a Robot* is, as far as we are aware, the first of its kind: a design fiction documentary.

The film was made by first showing selected clips from *Robot and Frank* that depict interactions between humans and the robot to the contributors who would eventually be the interviewees in *Care for a Robot*. The clips were selected to be deliberately thought provoking and encourage debate around whether the interactions shown were possible, plausible, or desirable.

The clips were shown to the interviewees, then we briefly introduced the relevant concepts in an informal discussion (including radical digital interventions, design fiction, anticipatory ethnography and our vision for *Care for a Robot*). Before filming any interviews we then asked interviewees to imagine they were living in a world where caring robots, just like the one they had seen 'diegetically situated' in *Robot and Frank*, were a reality and that either they or somebody close to them had experience of working with or owning these robots. Through dialogue between ourselves, and the interviewees, we developed a range of scenarios and personas that you see in the finished film. These are varied and include: a prospective customer buying for her father in law; a hacker who wants revenge after her robot's data was commandeered by the manufacturer; an employer who has appropriated care robots in order to access cheap labour; an academic who bought, and then returned, a care robot for his elderly parent.

None of these 'workshop' sessions were longer than 30 minutes. We did not script any of the responses, and the footage you see in the finished film is constructed from entirely improvised or 'off the cuff' responses to interview questions. *Care for a Robot* is not chronological and instead focuses on highlighting themes that emerged in the interviews.

2. RELEVANCE TO ETHICOMP

The primary purpose for this paper is to present a method for exploring the ethical considerations of radical digital interventions. In our example case the radical digital intervention is a domestic care-giving robot, however we suggest that the same method may be applied to other cases too. Although the method itself is the significant contribution here, we have included some examples quotes from *Care for a Robot* in the paper too (see 2.3). It is important to stress that we intend this work to initiate a discussion about how to use design fiction as a means to explore ethics as opposed to adopting a didactic position. This work is a first step.

2.1 We Are Not Ethicists

Although it should be clear by this point already, we want to reiterate that we are not ethicists. Neither were the interviewees that feature in *Care for a Robot*. However we believe that this fact

¹ The copyrighted materials from *Robot and Frank* are included under 'fair use' as part of a research project.

- that could be seen as a shortcoming for a paper about ethics – is not detrimental to the kernel of this work.

Design fictions tend to present the future as mundane. The future is an accretive space that may well include the buzzing of a cathode ray tube screen right alongside the sheen of a super-thin curved 3D-capable display. In *Robot and Frank*, rusty and ageing first generation hybrid cars are depicted sharing the roads with super-modern all-electric models. The future will not be a white-walled utopia but will be inhabited by a menagerie of semi-broken technologies and protagonists that, as we do today, are mainly motivated by everyday considerations [3,7]. By leveraging the future mundane (as it's shown in *Robot and Frank*), filtering those situations through the everyday perspectives of our interviewees, then finally packaging the outcome into a digestible format, is how this work creates meaning and generates value.

Being able to produce and contain insights pertaining to radically different ideas, while encoding the essence of everyday mundanity is how this work proposes to bring something new to the ethicist's toolbox. Because we're trying to tease out the 'warts and all' character of the future scenario being explored, it doesn't appear to be the case that our position as 'non-ethicists' has been too much of a hindrance.

2.2 Ethics and The Future

The challenges of understanding the ethics of technology appear to be necessarily bound to the future. We agree that as regards the ethics of technology "At bottom, these issues reduce to traditional ethical concerns having to do with dignity, respect, fairness, obligations to assist others in need, and so forth" [24]. The core ethical issues tend to remain quite static, meanwhile radical technological advances change the situations that these issues apply to considerably. It is the nature of these innovations, and the specifics of the situations they create, that are the largest challenge for ethicists. Design fictions naturally tend towards developing plausible concepts aligned with the trajectory of change, while also *communicating* these concepts with a high degree of 'situativity' [cf. 11,22].

Second, if we want to explore these possible scenarios - which of course are plural - then we need a means to ask meaningful 'what if' questions, as well as a means to understand the answers. There are various ways in which one might approach asking these questions [cf. 19,20]. We feel this design-fiction orientated approach has some distinguishing factors. First it has the ability to interrogate technologies radically different to those currently available; second that it does so within the brackets of a future mundane; third the ideas contained in the design fiction stimulus are filtered by the everyday and human responses of the interviewees. This results in insights that we refer to as 'diegetically situated'.

2.3 Example Quotes

We are clear that this work's primary aim is to describe and advocate for using design fiction as a tool to open a discursive space from which insights about ethics may emerge. As self-professed non-ethicists we're tentative about making any direct claims to do with ethical insights. More important than our own interpretations we hope that presenting this work at ETHICOMP 2015 will stimulate discussion and encourage interrogation of the idea such that it may be developed further, perhaps adapted, and hopefully adopted in other projects.

Despite intending for this work to, first and foremost, be a 'jumping off point' for further discussion, we have included a small selection of quotes from the interviews in the film in order to highlight some provocative examples.

2.3.1 Price vs. Value

Quite separately from the monetary value of the robot, or the cost to the user, the interviewees demonstrated a range of differing opinions about how to quantify the value of the robot carers.

"I would argue that this is a trade-off... it depends on what we would trade off for the services we have"

This interviewee accepts that the companies providing the robots may take something back in order to offset the cost of the robot, perhaps by monetising the data gathered by the robots. This seems consonant with 'free' services available on the web today, for example Google's suite of applications, or the services made available by numerous social networks.

"We have three wonderful kids but they give our sitters a hard time... I know they're not intended to take care of children"

The interviewee's children are apparently notoriously difficult for baby-sitters to handle, whereas using a robot carer to perform baby-sitting duties – which may be more expensive monetarily – appears to be preferable for her.

"We got it as a robot carer and what it was turning into was a research tool for the company"

During a year-long contract this interviewee became aware that, in accordance with the terms and conditions set out by the service provider, data gathered by the robot would be used in a number of unexpected ways, which are perhaps undesirable, and were not clear at the outset.

2.3.2 How Robot Carers Are Perceived and Used

As well as the intended application – to be domestic care robots for elderly users – some of our interviewees appropriated their robots to do jobs and tasks that way were not, perhaps, intended.

"I've found them to be extremely useful as flexible labour"

An entrepreneur, this interviewee has purchased many robots to work across his service-industry business as a cost-saving measure: human labour is unable to compete in terms of bottom-line hourly cost.

"..on the off chance.. if the robot happened to capture information from his medical records.."

This interviewee remotely reviews logs of the robot caring for his grandfather in order to discern what medication his grandfather is taking. It is unclear whether monitoring this level of detail is done with consent, and whether that was the intended use of this function.

"The robots outlook is that 'the best way to take care of elderly people is to have robot carers in their homes'"

This interviewee has become convinced that the robot caring for his wife's parents is trying to influence their behaviour, by, for instance, arranging their walk times so that they will encounter other people with caring robots.

2.3.3 Service Provision

All of our interviewees assumed that large corporations were providing the robot carers, either in a traditional ownership model or 'as a service'.

"We helped them buy a microwave, so they weren't about to go and buy a robot on their own"

Installing a care robot to care for an elderly relative may well necessitate dealing with highly technical issues, where the end-user might not be technologically savvy enough to have a full comprehension.

"They offer a personalised service... obviously you can't just unbox them and let it go... Somebody goes into his house and monitors his interactions with people so they can pre-program the robot"

This interviewee is very positive about the pre-sales support and level of personalisation that the company offered to support the installation of a care robot at her father in law's house.

"Any 3rd party service providers had to sign a disclaimer [if the robot was in the house]... it's like those messages saying 'this call may be monitored for training purposes'"

This interviewee was not initially aware that the contract with the robot provider insisted that *anyone* entering the house was required to sign a disclaimer allowing the company to use data gathered during their visit.

3. IMPLICATIONS

First and foremost we would like this work to stimulate a conversation with the ETHICOMP community. Does this design fiction centric approach to opening a discursive space about the ethical implications of radical interventions hold any merit? If so what frameworks could be applied to critically examine design fictions like Care for a Robot?

This work, that considers a Hollywood film as a piece of incidental design fiction, adapts the ideas within anticipatory ethnography, in order to then produce a new design fiction documentary, is a first. By focussing on domestic care robots, and in particular trying to discern insights about the ethical implications of this technology, our approach attempts to bound the design fiction, encouraging the discursive space to converge on around a single theme.

Although we have focussed this work on a single type of radical digital intervention we are keen to experiment with applying this approach to other types of radical innovation, perhaps those that have not been conceived yet.

Finally we would like to understand if applications of design fiction might be complimentary to more traditional research into the ethics of computing. Can the relationship between these disciplines be mutually beneficial?

4. ACKNOWLEDGMENTS

Many thanks to everyone in the closing credits of the film, your help and contributions were invaluable to doing this work. We're grateful to the stars and makers of Robot and Frank, for their film inspired this work. We would like to thank Robert Potts for his help developing our original formulation of 'anticipatory ethnography'. Thank you to our supervisors at Lancaster University. This work was funded by the UK Digital Economy Programme (Grant Reference EP/G037582/1).

5. REFERENCES

1. Auger, J. Speculative design: crafting the speculation. *Digital Creativity* 24, 1 (2013), 11–35.
2. Bleecker, J. Design Fiction: A short essay on design, science, fact and fiction. *Near Future Laboratory*, (2009).
3. Bleecker, J. A Design Fiction Evening with the Near Future Laboratory. 2013. <http://vimeo.com/84826827>.
4. Blythe, M. and Buie, E. Chatbots of the Gods: Imaginary Abstracts for Techno-Spirituality Research. *Proc. NordiCHI 2014*, (2014), 227–236.
5. Dryzek, J.S. *The politics of the earth: environmental discourses*. Oxford University Press, 2005.
6. Dunne, A. and Raby, F. *Speculative Everything*. The MIT Press, London, 2013.
7. Foster, N. The Future Mundane. *hellofosta.com*, 2013. <http://hellofosta.com/2013/10/07/the-future-mundane/>.
8. Hales, D. Design fictions an introduction and provisional taxonomy. *Digital Creativity* 24, 1 (2013), 1–10.
9. Kirby, D. The Future is Now: Diegetic Prototypes and the Role of Popular Films in Generating Real-world Technological Development. *Social Studies of Science* 40, 1 (2010), 41–70.
10. Lindley, J. and Potts, R. A Machine. Learning: An example of HCI Prototyping With Design Fiction. *Proceedings of the 8th Nordic Conference on Human Computer Interaction*, (2014).
11. Lindley, J., Sharma, D., and Potts, R. Anticipatory Ethnography: Design fiction as an input to design ethnography. *Ethnographic Praxis in Industry Conference*, (2014).
12. Lindley, J. and Sharma, D. Care for a Robot. 2015. <https://www.youtube.com/watch?v=VKKInpNueaY>.
13. Lindley, J. A pragmatics framework for design fiction. *Proceedings of the European Academy of Design Conference*, (2015).
14. Le Maitre, J. Medical robotics: the solution for our demographic challenge of an aging population? *Innorobo*. <http://innorobo.com/medical-robotics-the-solution-for-our-demographic-challenge-of-an-aging-population/>.
15. Manzini, E. Making Things Happen: Social Innovation and Design. *Design Issues* 30, 1 (2014), 57–66.
16. Norman, D.A. and Verganti, R. Incremental and Radical Innovation: Design Research vs. Technology and Meaning Change. *30*, 1 (2014).

17. Schreier, J. *Robot and Frank*. USA, 2013.
18. Sharma, D., Clune, S., and Blair, L. Radicalising the designer: Combating age-related loneliness through radical-digital interventions. *Cumulus Conference: The Virtuous Circle*, (2015).
19. Stahl, B.C., McBride, N., Wakunuma, K., and Flick, C. The empathic care robot: A prototype of responsible research and innovation. *Technological Forecasting and Social Change* 84, (2013), 74–85.
20. Stahl, B.C. IT for a better future how to integrate ethics, politics and innovation. (2011).
21. Sterling, B. Bruce Sterling Explains the Intriguing New Concept of Design Fiction (Interview by Torie Bosch). *Slate.com*, 2012.
http://www.slate.com/blogs/future_tense/2012/03/02/bruce_sterling_on_design_fictions_.html.
22. Suchman, L. *Plans and situated actions: the problem of human-machine communication*. Cambridge University Press, 1987.
23. Tanenbaum, J. Design fictional interactions. *Interactions* 21, 5 (2014), 22–23.
24. Tavani, H.T. *Ethics and Technology*. John Wiley & Sons, 2011.