Agents of Value

Batya Friedman The Information School University of Washington +1.206.616.7210 batya@uw.edu

ABSTRACT

Agents, be they human or technological, act on behalf of others in the world. Autonomous software agents may also act as representations (as in simulations) of human concerns. To do so both effectively and ethically, these agents must in some meaningful way understand and account for what those for whom they act consider important in life. At stake is no less than the risk of designing technologies that undermine important human values. Yet, with our limited view, it is not at all obvious how to design agent technologies so that they are more likely to support the actions, relationships, institutions, and experiences that human beings care deeply about.

In this invited speech I will explore the question of how to design agent technology to be sensitive to human values and concerns. I will draw from over two decades of design work and theory development in Value Sensitive Design. Along the way, I will touch on how agent technologies instantiate human values, the locus of value tensions in multi-agent systems, and key value tensions. From there, I will turn to methods from value sensitive design - direct and indirect stakeholder analyses, value scenarios, Envisioning Cards – which can be used to improve sensitivity to human values in agent technology research and design work. I will end with a few research challenges for which cutting edge agent technology design has provocative implications for human futures. Throughout, I will take an interactional stance: that agent technologies shape human experience and our very being; and through our experiences and being, we continually re-imagine those very technologies.

Categories and Subject Descriptors

D.2.10 [Design]: Methodologies and Representations. H.1.2 [User/Machine Systems], Human Factors. I.2.11 [Distributed Artificial Intelligence]: Intelligent Systems.

General Terms

Design, Human Factors.

Keywords

Value Sensitive Design.

Short Biography

Batya Friedman is a Professor in The Information School, Adjunct Professor in the Department of Computer Science, and Adjunct Professor in the Department of Human-Centered Design and Engineering at the University of



Washington where she directs the Value Sensitive Design Research Lab. Dr. Friedman pioneered value sensitive design (VSD), an approach to account for human values in the design of information systems. First developed in human-computer interaction, VSD has since been used in information management, human-robotic interaction, computer security, civil engineering, applied philosophy, and land use and transportation. Her work has focused on a wide range of values, some include privacy in public, trust, freedom from bias, moral agency, sustainability, safety, calmness, freedom of expression, and human dignity; along with a range of technologies such as web browsers, urban simulation, robotics, open source tools, mobile computing, implantable medical devices, social media, ubiquitous computing and computing infrastructure. Dr. Friedman is currently working on multi-lifespan information system design and on methods for envisioning - imagining new ideas for leveraging information systems to shape our futures. Voices from the Rwanda Tribunal is an early project in this multi-lifespan information system design program. In 2012 Batya Friedman received the ACM-SIGCHI Social Impact Award and the University Faculty Lecturer award at the University of Washington. She received both her B.A. and Ph.D. from the University of California at Berkeley.

Websites

- Batya Friedman: ischool.uw.edu/people/faculty/batya
- Value Sensitive Design Research Lab: vsdesign.org
- Voices from the Rwanda Tribunal: tribunalvoices.org
- Envisioning Cards: envisioningcards.com

Appears in: Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), Ito, Jonker, Gini, and Shehory (eds.), May 6–10, 2013, Saint Paul, Minnesota, USA. Copyright © 2013, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.