

The Emerging Policy and Ethics of Human Robot Interaction

Laurel D. Riek
Computer Science & Eng.
University of Notre Dame
lriek@nd.edu

Woodrow Hartzog
Cumberland School of Law
Samford University
whartzog@samford.edu

Don Howard
Philosophy Dept.
University of Notre Dame
dhoward1@nd.edu

Ajung Moon
Mechanical Engineering
University of British Columbia
ajmoon@mech.ubc.ca

Ryan Calo
School of Law
University of Washington
rcalo@uw.edu

ABSTRACT

As robotics technology forays into our daily lives, research, industry, and government professionals in the field of human-robot interaction (HRI) in must grapple with significant ethical, legal, and normative questions. Many leaders in the field have suggested that “the time is now” to start drafting ethical and policy guidelines for our community to guide us forward into this new era of robots in human social spaces. However, thus far, discussions have been skewed toward the technology side or policy side, with few opportunities for cross-disciplinary conversation, creating problems for the community. Policy researchers can be concerned about robot capabilities that are scientifically unlikely to ever come to fruition (like the singularity), and technologists can be vehemently opposed to ethics and policy encroaching on their professional space, concerned it will impede their work. This workshop aims to build a cross-disciplinary bridge that will ensure mutual education and grounding, and has three main goals: 1) Cultivate a multidisciplinary network of scholars who might not otherwise have the opportunity to meet and collaborate, 2) Serve as a forum for guided discussion of relevant topics that have emerged as pressing ethical and policy issues in HRI, 3) Create a working consensus document for the community that will be shared broadly.

Categories and Subject Descriptors

I.2.9 [Artificial Intelligence]: Robotics; K.4.1 [Computers and Society]: Public Policy Issues

Keywords

Human-robot interaction; robotics; ethics; policy; law; morphology; design; assistive robots; health technology; privacy

1. WORKSHOP OVERVIEW

The workshop is a single-track, day long meeting that includes the topics of deployment of robots in healthcare settings,

Permission to make digital or hard copies of part or all 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. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

HRI'15 Extended Abstracts, March 2–5, 2015, Portland, OR, USA.

ACM 978-1-4503-3318-4/15/03.

<http://dx.doi.org/10.1145/2701973.2714393>.

robot morphology, and autonomy, all within the context of robots and humans working together collaboratively. Each topic will be discussed by an invited panel, comprised of individuals from HRI-related disciplines (e.g., robotics, computer science, psychology, design), and policy-related disciplines (e.g., ethics, philosophy, law). Each panel will be moderated by one of the organizers. The end of the day will be dedicated to creating a working consensus document for the community. Goals include: (1) establishing common language for multi-disciplinary discussion, (2) identifying key areas of concern, (3) establishing baseline approaches to solving policy and ethical problems in HRI.

1.1 List of workshop topics

Healthcare: HRI researchers and practitioners often deploy robots in therapeutic settings with vulnerable populations; for example, to help treat children with autism spectrum disorders [14] to reduce stress and encourage pro-social behavior among older adults [6] and to help children with developmental disabilities [7]. They also have been used to assist people with daily living tasks, such as bathing, manipulation [16], mobility [1], and other activities to support independent living and aging-in-place [4].

However, the use of robots to aid with these therapeutic and intimate tasks with vulnerable populations raises substantial concerns: How will clients' privacy be protected? Should the use of robots for these tasks be regulated by an administrative agency such as the FTC, Health and Human Services (HHS), or a new agency specifically created to address the issues raised by robots? Could robot caregivers displace human caregivers, negatively affecting both clients' welfare and healthcare providers' jobs?

Morphology: Robots can range in appearance from looking mechanical to anthropomorphic in appearance [13]. Morphology is a richly debated topic in the community, with many studies showing people will anthropomorphize and form attachments to nearly anything conveying animacy [3, 2, 10]. Some worry that increasingly humanoid representations not only convey inaccurate expectations to people about a robot's capabilities, but may also be unethical when treating vulnerable populations [12].

Robots have significant ability to harm people not just physically, but through manipulation. People can develop feelings for robots, when can be leveraged to coerce these people to act in the interest of the robot's designer or controller. Imagine a robot designed to intentionally exploit emotional attachment to solicit payments for mandatory “upgrades” to keep the person's beloved friend “alive.”

Autonomy: With the anticipated, rapid introduction of semi-autonomous robotics technology in human social environments, an ever more important form of HRI will be handoff of control from robot to human controllers at various points of operation. In the HRI research community, these handoffs of control are often discussed in the context of instrumenting shared autonomy and designing for acceptability [9, 15]. Others have explored how Wizard-of-oz is used in HRI research settings [11]

However, the handoff of control to robots puts the onus of ethical and policy considerations, which must be made ex ante, on designers. Designers that assume control over actions from humans through automation will be subjecting themselves to legal liability and ethical problems with respect to when the handoff is appropriate and the substance of decisions once the handoff has been made. Entire liability and risk management regimes like insurance must be re-calibrated in light of the handoff of control. Care should be taken to ensure proper designing for ease of handoff without significant interruption of control functionality, and designing for avoidance of unwarranted human operator habituation to automatic controls. Significant open questions remain as to when the handoff of control to robots should be legally or ethically required.

2. PANELISTS

- **Meg Leta Ambrose**, *Communication, Culture, & Technology, Georgetown University*
- **Kenneth Anderson**, *College of Law, American University*
- **Peter Asaro**, *Media Studies, The New School*
- **Jodi Forlizzi**, *Human Computer Interaction Institute, Carnegie Mellon*
- **Michael Goodrich**, *Computer Science, Brigham Young University*
- **David Luxton**, *Naval Health Research Center*
- **Jason Millar**, *Philosophy, Queen's University at Kingston*
- **Ayse Saygin**, *Cognitive Science, University of California San Diego*
- **Jean Scholtz**, *Visual Analytics, Pacific Northwest National Laboratory*
- **Bill Smart**, *Mechanical Engineering, Oregon State University*
- **John Sullins**, *Philosophy, Sonoma State University*
- **Aimee van Wynsberghe**, *Philosophy, University of Twente*
- **Eric Valor**, *Team Gleason Initiative and SciOpen Research Group.*

3. WORKSHOP ORGANIZERS

- **Laurel Riek**, *Computer Science and Engineering, University of Notre Dame,*
- **Woodrow Hartzog**, *Cumberland School of Law, Sanford University,*
- **Don Howard**, *Philosophy, University of Notre Dame,*
- **AJung Moon**, *Mechanical Engineering, University of British Columbia,*
- **Ryan Calo**, *School of Law, University of Washington,*

4. DOCUMENTATION PLAN

A cross-disciplinary glossary and topic primer will be made available to participants on the workshop website. This will help establish common ground between participants from different

disciplines, and will serve as a useful educational tool for students wishing to explore the HRI ethics and policy space but who may not know where in the literature to begin.

Following the workshop, the organizers will produce an experiences report that will be widely distributed via the workshop website [8], the Robohub website [5], and various social media roboethics-related groups. The report will be licensed under a Creative Commons license, and deposited in an archive repository which will be accessible through a persistent ID.

5. REFERENCES

- [1] T. Carlson and Y. Demiris. Human-wheelchair collaboration through prediction of intention and adaptive assistance. In *IEEE International Conference on Robotics and Automation*, 2008.
- [2] B. R. Duffy. Anthropomorphism and the social robot. *Robotics and autonomous systems*, 42(3):177–190, 2003.
- [3] N. Epley, A. Waytz, and J. T. Cacioppo. On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, 114(4):864, 2007.
- [4] J. Forlizzi, C. DiSalvo, and F. Gemperle. Assistive robotics and an ecology of elders living independently in their homes. *Human-Computer Interaction*, 19, 2004.
- [5] S. Hauert and et. al. Robohub. <http://robhub.org>, 2015.
- [6] C. D. Kidd, W. Taggart, and S. Turkle. A sociable robot to encourage social interaction among the elderly. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2006.
- [7] H. Kozima, M. P. Michalowski, and C. Nakagawa. Keepon. *International Journal of Social Robotics*, 1(1):3–18, 2009.
- [8] L.D. Riek, W. Hartzog, D. Howard, A. Moon, and R. Calo. The Emerging Policy and Ethics of Human Robot Interaction. <http://openroboethics.org/hri15>, 2015.
- [9] C. A. Miller and R. Parasuraman. Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *HUM FACTORS*, 2007.
- [10] C. Nass and Y. Moon. Machines and mindlessness: Social responses to computers. *Journal of social issues*, 56(1):81–103, 2000.
- [11] L. D. Riek. Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines. *Journal of Human Robot Interaction*, 1:119–136, 2012.
- [12] L. D. Riek and D. Howard. A code of ethics for the human-robot interaction profession. *We Robot 2014*, pages 1–10, 2014.
- [13] L. D. Riek, T.-C. Rabinowitch, B. Chakrabarti, and P. Robinson. Empathizing with robots: Fellow feeling along the anthropomorphic spectrum. *Affective Computing and Intelligent Interaction (ACII)*, 2009.
- [14] B. Scassellati, H. Admoni, and M. Mataric. Robots for use in autism research. *Annual Review of Biomedical Engineering*, 14:275–294, 2012.
- [15] L. Takayama, E. Marder-Eppstein, H. Harris, and J. M. Beer. Assisted driving of a mobile remote presence system: System design and controlled user evaluation. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2011.
- [16] K. M. Tsui and H. A. Yanco. Simplifying wheelchair mounted robotic arm control with a visual interface. In *AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics*, pages 97–102, 2007.