A Special session proposal for the
IEEE International Conference on Robot and Human Interactive Communication
(RO-MAN 2020)

Title

SOCIETAL CONSEQUENCES OF HUMAN-ROBOT INTERACTION

Abstract

As social robots will become increasingly autonomous and present in our daily life in the upcoming years, the field of HRI and Social Robotics should investigate the challenges that their pervasive presence could provoke, and set legal and ethical boundaries to their use. We welcome papers that explore how the interaction with norm-violating robots (e.g., cheating robots) could affect human behavior (e.g., conformity), and investigate human misconduct with robots in terms of discrimination (e.g., racism) and mistreatment (e.g., bullying). We also accept papers that present legal and ethical frameworks on social robotics' use and data collection.

Aims and Scope

This special session welcomes papers that address the societal consequences of Human-Robot Interaction (HRI). As social robots will become increasingly autonomous and present in our daily life in the years to come, the field of HRI and Social Robotics should investigate the potential challenges that their pervasive presence could bring about, and set legal and ethical boundaries to their use. With this special session, we aim at understanding: how will the pervasive interaction with social robots change human behavior? Do humans comply with norm-violating social robots? Do bias and discrimination observable in human-human interaction extend to HRI? How can we legally and ethically limit the misuse of social robots?

We accept papers that explore how the exposure to and interaction with norm-violating social robots (e.g., cheating robots, bribing robots, erratic robots) could affect human behavior (e.g., conformity), and how concepts such as *trust*, *familiarity*, and *rapport* affect people's reactions to such robots. Moreover, we welcome papers that assess human misconduct with robots in terms of discrimination (e.g., racism/racialization, sexism/sexualization) and mistreatment (e.g., bullying, verbal disinhibition, violent behavior), but also papers that investigate how we, as humans, react to robotic harm.

As the debate on the societal consequences of HRI should be addressed from a multidisciplinary point of view, we also accept papers that present legal and ethical frameworks on social robotics' use and data collection, especially concerning vulnerable populations, such as children and persons with disabilities (e.g., people with dementia). In this sense, topics of interest are liability and accountability, transparency and explainability in HRI, and deception.

Keywords: Social HRI, Norm-violation, Discrimination, Legal Frameworks, Ethical Frameworks

Exemplar Papers:

1. Aroyo, A. M., Rea, F., Sandini, G., & Sciutti, A. (2018). Trust and Social Engineering in Human Robot Interaction: Will a Robot Make You Disclose Sensitive Information, Conform to Its Recommendations or Gamble?. *IEEE Robotics and Automation Letters*, *3*(4), 3701-3708.
2. Barnes, M. L., & Sternberg, R. J. (1989). Social intelligence and decoding of nonverbal cues. *Intelligence*, *13*(3), 263-287.

3. Carlson, Z., Lemmon, L., Higgins, M., Frank, D., & Feil-Seifer, D. (2017). This robot stinks! differences between perceived mistreatment of robot and computer partners. *arXiv preprint arXiv:1711.00561*.

4. Villaronga, E. F., Kieseberg, P., & Li, T. (2018). Humans forget, machines remember: Artificial intelligence and the right to be forgotten. *Computer Law & Security Review*, *34*(2), 304-313.

5. Keijsers, M., & Bartneck, C. (2018, February). Mindless Robots get Bullied. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 205-214). ACM.

6. Rosenthal-von der Pütten, A. M., Krämer, N. C., Hoffmann, L., Sobieraj, S., & Eimler, S. C. (2013). An experimental study on emotional reactions towards a robot. *International Journal of Social Robotics*, *5*(1), 17-34.

7. Salomons, N., van der Linden, M., Strohkorb Sebo, S., & Scassellati, B. (2018, February). Humans conform to robots: Disambiguating trust, truth, and conformity. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 187-195). ACM.

8. Sánchez Ramos, A. C., Contreras, V., Santos, A., Aguillon, C., Garcia, N., Rodriguez, J. D., ... & Strait, M. K. (2018, March). A preliminary study of the effects of racialization and humanness on the verbal abuse of female-gendered robots. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 227-228). ACM.

9. Strait, M. K., Aguillon, C., Contreras, V., & Garcia, N. (2017). The public's perception of humanlike robots: Online social commentary reflects an appearance-based uncanny valley, a general fear of a "Technology Takeover", and the unabashed sexualization of female-gendered robots. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 1418-1423). IEEE.

10. Strait, M., Ramos, A. S., Contreras, V., & Garcia, N. (2018, August). Robots Racialized in the Likeness of Marginalized Social Identities are Subject to Greater Dehumanization than those racialized as White. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 452-457). IEEE.

11. Szczuka, J. M., & Krämer, N. C. (2019). There's More to Humanity Than Meets the Eye: Differences in Gaze Behavior Toward Women and Gynoid Robots. *Frontiers in psychology*, *10*, 693.

List of topics

Topics include but are not limited to:
- Conformity to social robots
- Social robots violating injunctive norms
- Cheating/Bribing robots
- Faulty robots/Erratic social robots
- Deceptive social robots
- Social Engineering in HRI
- Racism/Racialization of social robots
- Sexism/Sexualization of robots
- Unconscious bias towards social robots
- Bullyism towards social robots
- Misuse/Abuse of social robots
- Verbal disinhibition towards social robots
- Violent behavior towards social robots
- Human reactions to robotic harm
- Surveillance capitalism in Social Robotics
- Accountability/Liability in HRI
- Transparency/Explainability in HRI
- Deception in HRI

- Ethical implications of misbehaving robots
- Mapping responsibilities in HRI
- Unintended consequences of HRI

List of possible contributing authors

Giulia Perugia, Uppsala University (UU), Sweden
Maike Paetzel, Uppsala University (UU), Sweden
Sebastian Wallkötter, Uppsala University (UU), Sweden
Ginevra Castellano, Uppsala University (UU), Sweden
Eduard Fosch-Villaronga, Leiden University (eLaw), The Netherlands
Aurelia Tamo-Larrieux, University of Zurich, Switzerland
Alexander Mois Aroyo, Italian Institute of Technology (IIT), Italy
Alessandra Sciutti, Italian Institute of Technology (IIT), Italy
Hiroshi Ishiguro, Osaka University, Japan
Christoph Bartneck, University of Canterbury, New Zealand
Astrid Rosenthal-von der Pütten, RWTH Aachen University, Germany
Denise Geiskkovitch, University of Manitoba, Canada
Brian Scassellati, Yale University, USA
Tony Belpaeme, Ghent University, Belgium

List of organizers (including short bio)

Dr. Giulia Perugia, Postdoctoral Researcher at Uppsala University (Sweden)
E-mail: giulia.perugia@it.uu.se

Giulia Perugia is a postdoctoral researcher at Uppsala Social Robotics Lab (Department of Information Technology, Uppsala University, Sweden). Dr. Perugia received a degree in Design from the Academy of the Arts and New Technologies (Italy) in 2007; a BA in Literature and Linguistics from the University of Roma Tre (Italy) in 2011, a MSc in Cognitive Science from the University of Siena (Italy) in 2013, and an Erasmus Mundus Joint Doctorate (EMJD) in Interactive and Cognitive Environments (ICE) – with a specialization in Assistive Technologies – from the Polytechnic University of Catalonia (UPC; Spain) and Eindhoven University of Technology (TU/e; Netherlands) in 2018. Her research interest lies at the intersection of Social HRI, Social Psychology, and Multimodal behavior. As a scientist, she is interested in understanding and modeling the emotional and social linkage that the interaction with social robots triggers, how this can be used for assistive and educational purposes, and how it might impact and shape our society. Her doctoral project – ENGAGE-DEM: A Model of Engagement of People with Dementia – was aimed at building a model of engagement of people with dementia during interactions with social robots by means of nonverbal behavior and psychophysiology. Perugia was awarded the Distinguished Interdisciplinary Research Award by the Robotic Society of Japan (RSJ) and the Korean Robotic Society (KROS) at RO-MAN 2017 for her pioneering work on the measurement of psychophysiology in people with dementia. She was Program Committee member of HAI 2019 and ICSR 2019, and currently is Review Editor of Frontiers in Robotics and AI (Human-Robot Interaction section), Poster Chair of HAI 2020, and co-organizer of the workshop *Trust, Acceptance and Social Cues in Robot Interaction* for RO-MAN 2019, and of the homonymous Special Session for the International Journal of Social Robotics (IJSR).

Dr. Eduard Fosch-Villaronga, Postdoctoral Researcher at Leiden University (Netherlands)

E-mail: e.fosch.villaronga@law.leidenuniv.nl

Dr. Eduard Fosch-Villaronga is a Marie Skłodowska-Curie Postdoctoral Researcher at eLaw - Center for Law and Digital Technologies at Leiden University in the Netherlands (2019-2021). Eduard investigates the legal and ethical aspects of healthcare robots and Artificial Intelligence (A.I.) technologies. Eduard has published the book "Robots, Healthcare, and the Law. Regulating Automation in Personal Care" (Routledge, 2019) and has been recently appointed as an expert at the "Sub-Group on Artificial Intelligence (A.I.), connected products and other new challenges in product safety" to the Consumer Safety Network, a consultative expert group chaired by the European Commission's Directorate-General for Justice and Consumers, which is responsible for E.U. policy on justice, consumer rights, and gender equality. He was also awarded the EURA Young Scholar Prize of the Jean Monnet European Centre of Excellence on the Regulation of Robotics and A.I. Eduard currently co-leads the Working Group on the Ethical, Legal, and Societal (ELS) Aspects for Wearable Robots at the H2020 European COST Action CA16116 and is interested in human-robot interaction, responsible innovation, and the future of law. Previously, he worked as a researcher at the Microsoft Cloud Computing Research Center at Queen Mary University of London (the U.K., 2018) investigating the legal implications of cloud robotics; and at the University of Twente (NL, 2017) as a postdoc, exploring iterative regulatory modes for robot governance. Eduard Fosch-Villaronga holds an Erasmus Mundus Joint Doctorate (EMJD) in Law, Science, and Technology coordinated by the University of Bologna (I.T., 2017), an LL.M. from University of Toulouse (F.R., 2012), an M.A. from the Autonomous University of Madrid (E.S.), and an LL.B. from the Autonomous University of Barcelona (CAT, 2011). Eduard is also a qualified lawyer in Spain, and his publications are available online.

Dr. Alexander Mois Aroyo, Postdoctoral researcher at the Italian Institute of Technology (Italy)

E-mail: alexander.aroyo@iit.it

Alexander Mois Aroyo earned his Ph.D. in Bioengineering and Robotics from the University of Genoa and the Italian Institute of Technology (IIT) in Genoa, Italy. His research was carried out in the Robotics, Brain and Cognitive Science group at the IIT, and at the Intelligent Robotics Laboratory at Osaka University, Japan. His dissertation topic "Bringing Human Robot Interaction towards Trust and Social Engineering" gives an insight on the key factors to develop trust between humans and humanoid robots, and tries to understand how over-trust could lead to social engineering techniques used to manipulate people, in both positive and negative ways. Currently he is working as a Postdoc fellow under the supervision of Alessandra Sciutti on the ERC Starting Grant wHiSPER, focused on the investigation of joint perception between humans and robots.

Dr. Alessandra Sciutti, Tenure-track researcher at the Italian Institute of Technology (Italy)

E-mail: Alessandra.Sciutti@iit.it

Alessandra received her Ph.D. in Humanoid Technologies from the University of Genova (Italy) in 2010. After a Post Doc at the Italian Institute of Technology (IIT) and two research periods in USA and Japan, she became the scientific responsible of the Cognitive Robotics and Interaction Laboratory of the RBCS Dept. at IIT. After being Assistant Professor in Bioengineering at DIBRIS University of Genoa, she is now Tenure-Track Researcher at the Italian Institute of Technology, head of the COgNiTive Architecture for Collaborative Technologies (CONTACT) unit. In 2018 she has been awarded the ERC Starting Grant wHiSPER, focused on the investigation of joint perception between humans and robots. She published more than 60 papers and abstracts and participated in the coordination of the CODEFROR European IRSES project. She is an Associate Editor of Robots and Autonomous Systems, Cognitive Systems Research and the International Journal of Humanoid Robotics and she has served as a member of the Program Committee for the International Conference on Human-Agent Interaction and IEEE International conference on Development and Learning and Epigenetic Robotics. The scientific aim of her research

is to investigate the sensory and motor mechanisms underlying mutual understanding in human-human and human-robot interaction.

Dr. Ginevra Castellano, Associate Professor at Uppsala University (Sweden)
E-mail: ginevra.castellano@it.uu.se

Ginevra Castellano is an Associate Professor in Intelligent Interactive Systems at the Department of Information Technology, Uppsala University, where she leads the Uppsala Social Robotics Lab. Her research interests are in the areas of social robotics and affective computing, and include social learning, personalized adaptive robots, multimodal behaviors and uncanny valley effect in robots and virtual agents. She has published 100 research papers on these topics, receiving over 2900 citations. She was the coordinator of the EU FP7 EMOTE (EMbOdied-perceptive Tutors for Empathy-based learning) project (2012–2016). She is the recipient of a Swedish Research Council starting grant (2016–2019) and PI for Uppsala University of the EU Horizon 2020 ANIMATAS (Advancing intuitive human-machine interaction with human-like social capabilities for education in schools; 2018-2021) project and of the COIN (Co-adaptive human-robot interactive systems) project, funded by the Swedish Foundation for Strategic Research (2016–2021). She was a member of the management board of the Association for the Advancement of Affective Computing; Program Committee member of ACM/IEEE HRI, IEEE/RSJ IROS, ACM ICMI, AAMAS, ACII and many others. Castellano was a general co-chair of IVA 2017. Castellano is an Associate Editor of Frontiers in Robotics and AI, Human-Robot Interaction section and IEEE Transactions in Affective Computing. Castellano received the 10-Year Technical Impact Award at the ACM International Conference on Multimodal Interaction 2019.