

Is there a need for critical robotics research?

Sara Ljungblad ^{a,1}, Niamh Ni Bhroin ^b and Sofia Serholt ^c
Mafalda Samuelsson Gamboa ^a

^a*Department of Computer Science and Engineering, University of Gothenburg,
Chalmers University of Technology*

^b*Department of Media and Communication, University of Oslo*

^c*Department of Applied IT, University of Gothenburg*

Abstract. What research questions and approaches are currently missing in the field of human-robot interaction research? We discuss humanistic and critical dimensions of critical robotics - an initiative created to investigate a possible paradigm shift in the field of robotics research. Robotic solutions are increasingly presented as a potentially desirable and emerging technologies in different social domains and institutions. At the same time, both this vision and the reality of robots in society pose a great deal of challenges. We provide some background to the critical robotics movement and present examples of research that takes a critical robotics perspective.

Keywords. critical robotics, human-robot interaction, drones, social robots, re-configuring HRI, epistemology

Introduction

Technically working robots in research, art and society are in many aspects fundamentally different from the high level reasoning, planning, and learning entities presented in science fiction [1]. Only very specific parts of our world can be computed and materialised, yet many issues that could critically question the use of robotics as a design material envisioned for institutions in society are not addressed in HRI (Human Robot Interaction). Initiatives such as Responsible Robotics, Responsible AI and Trustworthy Robotics are discussing privacy, bias, deception, and ethics related to AI and robotics [2], but not necessarily from an human robot interaction design perspective. The conference series Robophilosophy has become a transdisciplinary and interdisciplinary venue where different aspects of social robots are discussed by philosophers, researchers from the humanities and social sciences and engineers [3]. In this context, Bischof et al [4] arranged a panel, a continuation of a previous workshop theme entitled "Re-Configuring HRI: The

¹Corresponding Author: Sara Ljungblad, University of Gothenburg, Chalmers University of Technology Gothenburg, Sweden; E-mail: sara.ljungblad@chalmers.se

mutual shaping of social robots and social institutions and how interdisciplinary work could support new perspectives in HRI”.

Over several years, we have worked on developing the concept of critical robotics, essentially aiming to reconfigure HRI. We aim to contribute to research and society by reflecting on how humanistic and critical perspectives can complement existing HRI research, provide alternative epistemological perspectives. Critical robotics research is formulated 1) through the support from the community and 2) by exploring alternative and complementing design goals and research directions [5,6]. One aspect of critical robotics is a call for the application of critical theories, methods, and perspectives to the exploration of how robotic technologies are situated in and relate to broader social processes and what the consequences of this might be. In social science, there is already a broader understanding of related social structures, agency, power and social dynamics, which essentially concerns critical theory [7]. The critical and re-configuring of robotics has been done in art for more than 100 years [8], and in the related field HCI (Human Computer Interaction), designers and researchers do critical design to challenge underlying assumptions, norms and illusions in technology [9].

We believe that the shaping and defining of critical robotics research is an ongoing process that will work best as a collective effort from the research community, where a few initiatives are emerging to strengthen humanistic and critical perspectives in the field, e.g. re-configuring through an STS lens [4], or through design research methods [10]. The field of Human-Computer Interaction (HCI), has already been through several paradigm shifts of the research practice. Harrison et al. [11] consider that these paradigm shifts can be traced in the underlying metaphor of interaction that is being discussed in the research field. Such metaphors guide research questions and appropriate methods, yet different metaphors can co-exist. We will now present the activities we have conducted so far and briefly outline how these have informed critical robotics research.

1. What could Critical Robotics Research concern?

In 2018, we arranged a workshop entitled “Critical Robotics – Exploring a new paradigm” [5]. The workshop was framed as a critique against solutionism to open up for critical, sustainable and humanistic perspectives, questioning how robotic solutions are shaped. While MacKenzi and Wajcman’s [12] work addresses the social context, there is also a long-standing art and literary tradition in exploring and articulating potential robot futures [8], and critical design [13] exemplifying critical practices and perspectives relevant for HRI. Thus, our initial goal was to critically address and problematize underlying values in robotic research and how these are communicated. In this way, we pointed to a potentially deceptive vision of robotic capabilities (see e.g. [1]).

During the critical robotics workshop we used a “backcasting” method. The method was initially used to advance society towards greater sustainability, developed by The Natural Step [14]. We adapted the method to articulate a desirable future for HRI research (i.e. desirable future research questions, methods, etc), and then to check this vision against the existing reality of conducted research. Finally, we “backcasted” from our desirable future to visualize the steps needed to reach this future. The steps (i.e. approaches and research perspectives) were suggestions for specific activities that would be needed in order to change the field in the long term. Several participants actually had

a dystopian science fiction inspired view of a future with robots, but a more humanistic vision of the kind of research that would be important for society. The workshop resulted in several suggestions for HRI research including: enhancing the transparency of the research process by reporting on failures and problems, the need to study existing practices, and to avoid too broad robot scenarios. Other suggestions included research-based ethical perspectives to reflect dilemmas encountered in the research context and avoid high-level ethical discussions which are not informed by empirical research. Other suggestions were to nuance more dystopian views of robots with empirical research, to do more "activist" and qualitative research, to recruit diverse research participants, and to address tensions between academia and industry.

In 2022, we published a special issue on critical robotics in AI and Society [6] and reached out to the community to contribute in defining and exploring critical robotics as a specific research direction. The key contributions of the included articles concerned:

- 1) The robotification of the society. E.g. discussing how robotics can disrupt socio-cultural life worlds for better and worse ([15].
- 2) Oversimplification of care and service work. E.g. presenting the risks of oversimplifying work with aging people and in health care [16] and how care practices are being deconstructed [17]
- 3) The (not so) straight forward path to robot design. E.g. how participatory design still may lead to undesirable solutions [18,19] and the value of STS [16]
- 4) Problematizing socio-emotional relationships with robots. E.g. how it is more important to foster reciprocity among humans than for robots [20]
- 5) Exploring and challenging sociality in Human robot interaction. E.g. questioning relationships with robots through art [21]

2. What would HRI be without critical research practices?

Do we really need to question robotics in HRI research? If the field would reject humanistic and critical perspectives, such practices would still exist but be found elsewhere such as in STS communities [4], in robotic art [8] and in critical design communities in HCI [13]. HRI would miss out on highly engaging, sometimes humorous and provocative work, with strong humanistic rather than technological foundations, decomposing robotic dystopia. One example of what critical practices brings to HRI is Lupetti and Van Mechelen [10], who conducted an explorative design study using a robot in a school to promote critical thinking in children. In their work, they capitalize on the explicitly deceptive nature of robots and suggest to "embrace and magnify robot aspects with deceptive potential, such as personality and intentionality, so that they become explicitly controversial and may be used as a springboard for a critical debate" [10]. A critical stance can also be part of the methodology. Ostrowski and Breazeal [22] recruit a participatory design approach to discuss social justice through design fiction.

What if HRI researchers would not be encouraged to reflect on their own bias and visions? A reflexive stance is not necessarily seeking to remove bias, but rather to be mindful of the researcher's positionality. Bardzell and Bardzell [9] recommend self-disclosure: "Practice should involve a disclosure of the researcher's position in the world, her or his goals, as well as the researcher's position in her or his intellectual and, to an appropriate extent, political beliefs". A related challenge is the need to reflect on the con-

ceptualization of practices that inform the design and development of robotic services, as stereotyped views of people and practices may lead to over simplification, even when participatory design practices are used (see e.g. [17,18,19]). Even if some critical approaches to re-configuring robotics has been introduced through art for more than 100 years [8], and activist and reflective research is found in the related field of HCI [13] we believe that a focus on critical robotics research deserves a more prominent role in the field of HRI. Such research could address norms, values and power structures embedded in the field and incorporate more thoughtful social and technological critique which can in turn contribute to both social and technical progress.

References

- [1] Y. Fernelius, M. Jacobsson, S. Ljungblad and L.E. Holmquist, Are we living in a robot cargo cult?, in: *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, 2009, pp. 279–280.
- [2] M. Brandão, M. Mansouri and M. Magnusson, Responsible Robotics, *Frontiers in Robotics and AI* **9** (2022).
- [3] J. Seibt, M.F. Damholdt and C. Vestergaard, Five Principles of Integrative Social Robotics., in: *Robophilosophy/TRANSOR*, 2018, pp. 28–42.
- [4] A. Bischof, E. Hornecker, A.L. Krummheuer and M. Rehm, Re-Configuring Human-Robot Interaction., in: *HRI*, 2022, pp. 1234–1236.
- [5] S. Ljungblad, S. Serholt, T. Milosevic, N.N. Bhroin, R.T. Nørgård, P. Lindgren, C. Ess, W. Barendregt and M. Obaid, Critical robotics: exploring a new paradigm, in: *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*, 2018, pp. 972–975.
- [6] S. Serholt, S. Ljungblad and N. Ní Bhroin, Introduction: special issue—critical robotics research, Springer, 2021, pp. 1–7.
- [7] C. Fuchs, *Social media: A critical introduction*, Sage, 2021.
- [8] C. Csíkszentmihályi, An Engineer’s Nightmare: 102 Years of Critical Robotics, in: *Presented at the “Re-Configuring Human-Robot Interaction”, at the ACM/IEEE international conference on Human robot interaction*, 2022.
- [9] S. Bardzell and J. Bardzell, Towards a feminist HCI methodology: Social science, feminism, and HCI, in: *Conference on Human Factors in Computing Systems - Proceedings*, 2011, pp. 675–684. ISBN ISBN 9781450302289. doi:10.1145/1978942.1979041.
- [10] M.L. Lupetti and M. Van Mechelen, Promoting Children’s Critical Thinking Towards Robotics through Robot Deception, in: *HRI ’22: Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, 2022, pp. 588–597. ISBN ISBN 9781665407311.
- [11] S. Harrison, D. Tatar and P. Sengers, The three paradigms of HCI, in: *Alt. Chi. Session at the SIGCHI Conference on human factors in computing systems San Jose, California, USA*, 2007, pp. 1–18.
- [12] D. MacKenzie and J. Wajcman, *The social shaping of technology*, Open university press, 1999.
- [13] J. Bardzell and S. Bardzell, What is” critical” about critical design?, in: *Proceedings of the SIGCHI conference on human factors in computing systems*, 2013, pp. 3297–3306.
- [14] The Natural Step www.thenaturalstep.org, Accessed: 2022-06-04.
- [15] M. Nørskov, Robotification & ethical cleansing, *AI & SOCIETY* **37**(2) (2022), 425–441.
- [16] D. Burema, A critical analysis of the representations of older adults in the field of human–robot interaction, *AI & SOCIETY* **37**(2) (2022), 455–465.
- [17] A. Maibaum, A. Bischof, J. Hergesell and B. Lipp, A critique of robotics in health care, *AI & society* **37**(2) (2022), 467–477.
- [18] A. Weiss and K. Spiel, Robots beyond Science Fiction: mutual learning in human–robot interaction on the way to participatory approaches, *AI & SOCIETY* **37**(2) (2022), 501–515.
- [19] A. Dobrovestnova, G. Hannibal and T. Reinboth, Service robots for affective labor: a sociology of labor perspective, *Ai & Society* **37**(2) (2022), 487–499.
- [20] A. van Wynsberghe, Social robots and the risks to reciprocity, *AI & SOCIETY* **37**(2) (2022), 479–485.
- [21] C. Yolgormez and J. Thibodeau, Socially robotic: making useless machines, *AI & SOCIETY* **37**(2) (2022), 565–578.

- [22] A.K. Ostrowski and C. Breazeal, Design Justice for Robot Design and Policy Making, in: *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '22, IEEE Press, 2022, pp. 1170–1172–.