TU Delft

Delft University of Technology

Document Version
Final published version

Citation (APA)
De Graaf, M. M. A., Perugia, G., Fosch-Villaronga, E., Lim, A., Broz, F., Short, E. S., & Neerincx, M. (2022). Inclusive HRI: Equity and Diversity in Design, Application, Methods, and Community. In *HRI 2022 - Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 1247-1249). (ACM/IEEE International Conference on Human-Robot Interaction; Vol. 2022-March). IEEE. https://doi.org/10.1109/HRI53351.2022.9889455

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Inclusive HRI: Equity and Diversity in Design, Application, Methods, and Community

Maartje M. A. de Graaf
Dept. of Information and Computing Sciences
Utrecht University
Utrecht, Netherlands
m.m.a.degraaf@uu.nl

Giulia Perugia
Dept. of Industrial Engineering and Innovation Sciences
Eindhoven University of Technology
Eindhoven, Netherlands
g.perugia@tue.nl

Eduard Fosch-Villaronga
eLaw Center for Law and Digital Technologies
Leiden University
Leiden, Netherlands
e.fosch.villaronga@law.leidenuniv.nl

Angelica Lim
School of Computing Science
Simon Fraser University
Burnaby, Canada
angelica@sfu.ca

Frank Broz
Dept. of Intelligent Systems
Delft University of Technology
Delft, Netherlands
f.broz@tudelft.nl

Elaine Schaertl Short
Dept. of Computer Science
Tufts University
Medford, USA
elaine.short@tufts.edu

Mark Neerincx
Dept. of Intelligent Systems
Delft University of Technology
Delft, Netherlands
m.a.neerincx@tudelft.nl

2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI) | 978-1-6654-0731-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/HRI53351.2022.9889455

***Abstract*—Discrimination and bias are pressing issues of many AI and robotics applications. These outcomes may derive from limited datasets that do not fully represent society as a whole or from the AI scientific community's western-male configuration bias. Although being a pressing issue, understanding how robotic systems can replicate and amplify inequalities and injustice among underrepresented communities is still in its infancy among social science and technical communities. This workshop contributes to filling this gap by exploring the research question: What do diversity and inclusion mean in the context of Human-Robot Interaction (HRI)? Here, attention is directed to three different levels of HRI: the technical, the community, and the target user level. Overall, this workshop will focus on the idea that AI systems can be created to be more attuned to inclusive societal needs, respect fundamental rights, and represent contemporary values in modern societies by integrating diversity and inclusion considerations.**

***Index Terms*—Diversity, Inclusion, Equity, Human-Robot Interaction, Accessibility, Global South, Gender, LGBTQAI+**

I. INTRODUCTION

Robot technologies help automate industrial, retail, and farming sectors and, lately, healthcare, education, and public service. While robots and Artificial Intelligence can increase resource efficiency and productivity, automating parts of society reserved once only to humans is nonetheless not straightforward and raises particular ethical, legal, and societal concerns [1]–[3]. A growing global concern is that AI systems may exacerbate and reinforce stereotypes and biases that different societies have with respect to gender, age, race, and sexual orientation [4]–[6]. For instance, face recognition systems having difficulty recognizing dark-skinned women and content moderator tools automatically flagging drag queens' explicit use of language as toxic, thus preventing them from freely communicating online [5], [7]. These outcomes may derive from limited datasets that do not fully represent the society as a whole or from the AI scientific community's structural and systematic configuration biases [8], [9], yet they are extremely influential [4], [10], [11].

Nowadays, there is an exponential growth of personal voice assistants that can socially interact with users [12]. A common feature of these artifacts is that they are given female names, have female voices, and usually display a servile personality engineered to please users all the time [12]–[14]. When artificial agents become embodied, the situation changes. A recent study has unveiled that only 15% of the robots in the Anthropomorphic roBOT (ABOT) dataset [15] are perceived as feminine, and none of them as androgynous [16]. Alesich and Rigby [12] also note how feminine robots are usually employed in gender-stereotypical roles: house-cleaners, caretakers, and child-minders, among the others. The use of female and masculine voices / robots in serviceable and medical contexts reinforces gender stereotypes about the role women and men should (or should not) play in society [17]. These are usually biases rooted in oppressive gender inequalities that have existed throughout history and are typically exacerbated by the lack of diversity of the technical teams developing algorithms and robots, which usually work in companies with significant gender disparities in their board of directors [18], [19]. Similar concerns are found in other AI applications, namely in algorithms for medical applications [20], gender

Authorized licensed use limited to: TU Delft Library. Downloaded on November 07,2022 at 06:54:09 UTC from IEEE Xplore. Restrictions apply.

classifiers for marketing, social media platforms, or recruiting practices, resulting in disparities in hiring [21]. Likewise, sex robotics mostly targets cisgender heterosexual men and objectifies women's bodies [22].

The scientific community widely supports the idea that integrating gender and sex factors in research makes better science [23], [24]. However, many disciplines struggle to account for diversity. Authors continuously report that "inequality and a lack of gender diversity still exist in medicine, especially in academia and leadership" [25]; that "when we look to the diversity in immunology research labs, overwhelmingly, women, people of color, and LGBTQIA+ scientists are underrepresented among the laboratory head and top leadership roles" [26]. The AI community is by no means different in this respect, as highlighted by recent studies that explored gender biases in the community, i.e., "our results indicate a huge gender unbalance among authors, a lack of geographical diversity" [27]. However, missing sex and gender considerations in the development of robotics can lead to adverse consequences for society that range from exacerbating existing biases and stereotypes (which are prohibited by law) to the safety concerns related to misgendering a person in health-related applications [6], [20].

II. Inclusive HRI workshop

Although being a pressing issue, understanding how robotic systems and their AI can replicate and amplify inequalities and injustice among underrepresented communities is still in its infancy. This workshop contributes to filling this gap by exploring the research question: What do diversity and inclusion mean in the context of HRI?

This full-day workshop aims to provide a forum to share experiences and research insights on identifying, addressing, and integrating Diversity, Equity, and Inclusion (DEI) aspects in HRI. We cover DEI in terms of design and application of robotic technologies, as well as research methods and community. We build on seminal work developed in this area by the GenR workshop[1] at RO-MAN 2021, the ongoing Special Issue *Gendering Robots: Ongoing (Re)configurations of Gender in Robotics* on the International Journal of Social Robotics[2], the Gendering Algorithms initiative at Leiden University[3], and the Diversity, Equity, and Inclusion for Embodied AI (DEI4EAI) initiative[4] created by the 4TU Federation in collaboration with Leiden University in the Netherlands.

Our goal is to raise awareness of the importance of DEI in HRI to avoid exacerbating existing biases and stereotypes or creating new ones. Extended time for discussions will highlight and document promising research directions and approaches to encourage further work in this area. A large part of this effort is to bring together a community of researchers by strengthening existing connections, and building new ones under the topic of diversity and inclusion.

[1]https://sites.google.com/view/ro-man21-genr-workshop/home
[2]https://www.springer.com/journal/12369/updates/19292566
[3]https://www.genderingalgorithms.org/
[4]https://www.dei4eai.com/

III. Format

The workshop format will combine informative sessions, panel discussions, and audience engagement. We plan themed discussion sessions around the key topics raised by accepted paper submissions. A large part of the workshop will be devoted to discussing the next steps in making HRI more diverse, equal, and inclusive. Panel discussions and groups will be composed of representatives from different backgrounds to integrate the necessary perspectives in this endeavor. We will ask the authors of the accepted papers and the HRI community to provide questions or raise pressing issues that provide starting points to boost discussion. To do so, we will create a QA dedicated page on our website.

IV. Audience and Dissemination

In this workshop, we want to bring together researchers and practitioners from a wide range of backgrounds, including computer science, engineering, ethics, law, gender studies, and HCI, interested in making HRI more inclusive and diverse. We encourage researchers to attend the workshop even without a paper submission. Our goal is to maximize community engagement to further increase awareness of DEI issues. A careful dissemination plan will be laid out to ensure that the call for papers and later also a call for participation is distributed via various mailing lists, social media, and networks. A workshop website (https://sites.google.com/view/dei-hri-2022/) has been created to provide information about the workshop, disseminate the accepted papers, and promote community building. People wanting to be part of the DEI-HRI community can fill out the form on our website: https://sites.google.com/view/dei-hri-2022/community.

V. Submissions and Expected Outcomes

We invite authors to submit extended abstracts (up to 2 pages, excluding references) and short papers (up to 4 pages, excluding references) on a range of topics relevant to DEI in HRI. Since we hope to learn from other fields of knowledge and form new connections with related research communities, we also welcome submissions from researchers outside of the HRI community. We particularly welcome HRI and Social Robotics research focusing on accessibility, disability and ableism, LGBTQIA+, intersectional feminism, neurodiversity, the global south, gender, sexual orientation, race, ethnicity, disability, or religion.

All papers should be submitted in PDF format using the IEEE two-column format on EasyChair, and will be peer-reviewed based on their originality, relevance, technical soundness, and clarity. Paper acceptance requires that at least one author registers for and (virtually) attends the workshop. After the conference, we will provide online access to the workshop proceedings on the workshop website with the authors' permission. In addition, the organizers will coordinate a White Paper on the topic of Inclusive HRI with the contribution of authors and participants, thereby further disseminating ideas and discussions developed during the workshop and establishing a clear road map to pursue DEI in HRI.

Authorized licensed use limited to: TU Delft Library. Downloaded on November 07,2022 at 06:54:09 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] D. Schönberger, "Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications," *International Journal of Law and Information Technology*, vol. 27, no. 2, pp. 171–203, 2019.

[2] G. Wisskirchen, B. T. Biacabe, U. Bormann, A. Muntz, G. Niehaus, G. J. Soler, and B. von Brauchitsch, "Artificial intelligence and robotics and their impact on the workplace," *IBA Global Employment Institute*, vol. 11, no. 5, pp. 49–67, 2017.

[3] L. Righetti, R. Madhavan, and R. Chatila, "Unintended consequences of biased robotic and artificial intelligence systems [ethical, legal, and societal issues]," *IEEE Robotics & Automation Magazine*, vol. 26, no. 3, pp. 11–13, 2019.

[4] S. U. Noble, *Algorithms of oppression*. New York University Press, 2018.

[5] I. D. Raji and J. Buolamwini, "Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products," in *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 2019, pp. 429–435.

[6] E. Fosch-Villaronga, A. Poulsen, R. A. Søraa, and B. Custers, "A little bird told me your gender: Gender inferences in social media," *Information Processing & Management*, vol. 58, no. 3, p. 102541, 2021.

[7] A. Gomes, D. Antonialli, and T. Dias-Oliva, "Drag queens and artificial intelligence. should computers decide what is toxic on the internet," 2019.

[8] J. Zhao, T. Wang, M. Yatskar, V. Ordonez, and K.-W. Chang, "Men also like shopping: Reducing gender bias amplification using corpus-level constraints," *arXiv preprint arXiv:1707.09457*, 2017.

[9] M. Roopaei, J. Horst, E. Klaas, G. Foster, T. J. Salmon-Stephens, and J. Grunow, "Women in ai: Barriers and solutions," in *2021 IEEE World AI IoT Congress (AIIoT)*. IEEE, 2021, pp. 0497–0503.

[10] M. Willson, "Algorithms (and the) everyday," *Information, Communication & Society*, vol. 20, no. 1, pp. 137–150, 2017.

[11] J. Ito, "Supposedly 'fair' algorithms can perpetuate discrimination," 2019.

[12] S. Alesich and M. Rigby, "Gendered robots: Implications for our humanoid future," *IEEE Technology and Society Magazine*, vol. 36, no. 2, pp. 50–59, 2017.

[13] J. Liu, "Social robots as the bride? understanding the construction of gender in a japanese social robot product," *Human-Machine Communication*, vol. 2, no. 1, p. 5, 2021.

[14] J.-C. Giger, N. Piçarra, P. Alves-Oliveira, R. Oliveira, and P. Arriaga, "Humanization of robots: Is it really such a good idea?" *Human Behavior and Emerging Technologies*, vol. 1, no. 2, pp. 111–123, 2019.

[15] E. Phillips, X. Zhao, D. Ullman, and B. F. Malle, "What is human-like? decomposing robots' human-like appearance using the anthropomorphic robot (abot) database," in *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, 2018, pp. 105–113.

[16] G. Perugia, S. Guidi, M. Bicchi, and p. y. Parlangeli Oronzo, booktitle=Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction, "The shape of our bias: Perceived age and gender in the humanoid robots of the abot database."

[17] A. Danielescu, "Eschewing gender stereotypes in voice assistants to promote inclusion," in *Proceedings of the 2nd Conference on Conversational User Interfaces*, 2020, pp. 1–3.

[18] M. West, R. Kraut, and H. Ei Chew, "I'd blush if i could: closing gender divides in digital skills through education," Tech. Rep., 2019.

[19] F. Rahman and E. Billionniere, "Re-entering computing through emerging technology: Current state and special issue introduction," *ACM Transactions on Computing Education (TOCE)*, vol. 21, no. 2, pp. 1–5, 2021.

[20] D. Cirillo, S. Catuara-Solarz, C. Morey, E. Guney, L. Subirats, S. Mellino, A. Gigante, A. Valencia, M. J. Rementeria, A. S. Chadha *et al.*, "Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare," *NPJ digital medicine*, vol. 3, no. 1, pp. 1–11, 2020.

[21] S. Park and J. Woo, "Gender classification using sentiment analysis and deep learning in a health web forum," *Applied Sciences*, vol. 9, no. 6, p. 1249, 2019.

[22] K. Richardson, "The asymmetrical'relationship' parallels between prostitution and the development of sex robots," *Acm Sigcas Computers and Society*, vol. 45, no. 3, pp. 290–293, 2016.

[23] L. Schiebinger, "Scientific research must take gender into account," *Nature News*, vol. 507, no. 7490, p. 9, 2014.

[24] C. Tannenbaum, R. P. Ellis, F. Eyssel, J. Zou, and L. Schiebinger, "Sex and gender analysis improves science and engineering," *Nature*, vol. 575, no. 7781, pp. 137–146, 2019.

[25] O. Ekmekcioglu, L. Evangelista, and J. Kunikowska, "Women in nuclear medicine," *European Journal of Nuclear Medicine and Molecular Imaging*, pp. 1–2, 2021.

[26] J. R. Groom, "Diversity in science requires mentoring for all, by all," *Nature immunology*, vol. 22, no. 9, pp. 1065–1065, 2021.

[27] A. Freire, L. Porcaro, and E. Gómez, "Measuring diversity of artificial intelligence conferences," in *Artificial Intelligence Diversity, Belonging, Equity, and Inclusion*. PMLR, 2021, pp. 39–50.

Authorized licensed use limited to: TU Delft Library. Downloaded on November 07,2022 at 06:54:09 UTC from IEEE Xplore. Restrictions apply.