HRI Wasn't Built In a Day: A Call To Action For Responsible HRI Research

Micol Spitale[1,2], Rebecca Stower[3], Maria Teresa Parreira[4] Elmira Yadollahi[5], Iolanda Leite[3], and Hatice Gunes[1]

***Abstract*— In recent years, the awareness of the academy around *responsible* research has notably increased. For instance, with advances in machine learning and artificial intelligence, recent efforts have been made to promote ethical, fair, and inclusive AI and robotics. To better understand if and to what extent HRI is incentivizing researchers to engage in responsible research, we conducted an exploratory review of the publishing guidelines for the most popular HRI conference venues. We identified 18 conferences which published at least 7 HRI papers in 2022. From these, we discuss four themes relevant to conducting responsible HRI research in line with the Responsible Research and Innovation framework: ethical and human participant considerations, transparency and reproducibility, accessibility and inclusion, and plagiarism and LLM use. We identify several gaps and room for improvement within HRI regarding responsible research. Finally, we establish a call to action to provoke conversations among HRI researchers about the importance of conducting responsible research within emerging fields like HRI.**

I. Introduction

Human-robot interaction (HRI) focuses on the study of interactions between robots and humans, by encouraging researchers from robotics, cognitive science, computer science, human-computer interaction, and psychology to come together. Initially, HRI collaborations took place in workshops held during other conferences (e.g., CHI), which then evolved into an independent research field with the advent of established conferences in the early '90s, such as the IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) that was born in 1991. HRI as a field is, thus, still in its infancy. Nonetheless, similar to many other fields, progress in HRI is inherently fast-paced, with ever-changing technological resources and needs. HRI is often touted as an interdisciplinary field, with this arguably being one of its major strengths [1]. Yet, establishing, maintaining, and publishing fruitful interdisciplinary collaborations is a non-trivial undertaking, especially in such a nascent field. As a consequence, we observe the emergence of (mostly unwritten) heterogeneous expectations and requirements for how to conduct and publish HRI research, which results in more laborious reviewing, lower impact research, and higher barriers to entry for novice HRI researchers.

[1] Department of Computer Science and Technology, University of Cambridge, UK, [2] Department of Electronic, Information, and Bio-engineering, Politecnico di Milano, Italy, [3] Division of Robotics, Perception, and Learning, KTH Royal Institute of Technology, [4] Information Science, Cornell University, [5] School of Computing and Communications, Lancaster University, UK. micol.spitale@polimi.it

Responsible Research and Innovation (RRI) refers to a framework introduced in 2012[1] by the European Commission, defined as a set of methodological guidelines and recommendations aimed at considering the interaction between society and science, focusing on challenges such as sustainability, public engagement, ethics, science education, gender equality, open access, and governance [2]. Several research communities have followed the RRI framework, for example, the artificial intelligence (AI) community has attempted to address ethical challenges by proposing a set of guidelines that include principles and recommendations aimed at harnessing the potentials of new AI technologies [3]–[5]. Similarly, recommendations to improve research transparency and reproducibility have also been proposed in AI and psychology fields [6], [7]. The advent of the third wave in HCI [8] – which focuses on the pervasive and ubiquitous presence of technology in everyday life – incentivizes HCI researchers to put inclusion and accessibility at the centre of their research to provide people with disabilities more opportunities to use innovative technologies [9]. Within HRI, some recent reviews propose guidelines to conduct transparent and reproducible user research (e.g., preregistration, data sharing) [10], [11]. However, the field of HRI is still lagging behind other research communities in distilling guidelines and recommendations for **responsible HRI** – defined as thoughtfully and transparently designed and conducted research.

Considering all the different stakeholders within the research and publication cycle – authors, reviewers, institutions, publishing venues (journals and conferences) – consensus on what makes *a responsible HRI paper* is oftentimes divided. In this context, we believe that conferences play an important role in promoting responsible HRI by homogenising publishing expectations within the field, thereby serving as the ideal platform to define and establish responsible guidelines as a point of convergence among different perspectives and research backgrounds. Are venues that publish HRI research taking advantage of this opportunity? Are relevant conferences sufficiently and effectively promoting responsible research practices across the field?

We begin addressing these questions by conducting an exploratory review of conference venues publishing HRI works to better understand how to promote responsible HRI

[1] https://op.europa.eu/en/publication-detail/-/publication/bb29bbce-34b9-4da3-b67d-c9f717ce7c58/language-en

research within the community and encourage discussion among researchers around this topic. This work is not supposed to present a comprehensive list of guidelines or recommendations for HRI researchers; instead, it aims at sparking conversations around the importance of conducting responsible research. As we address the challenge of "Embracing Human-Centered HRI", with works on Ethical Issues in Human-robot Interaction Research and Trust, Transparency, and Explainability in HRI, the RO-MAN conference – with a great majority of topics related to HRI in their CfP – is the perfect venue to host such discussions. This paper may serve as a basis to motivate the HRI research community to take further steps towards implementing responsible research practices in a more systematic way that engage ethical, transparent, and inclusive interactions with humans as part of our research efforts.

II. HRI Research Venues

This section describes the methodology adopted to conduct our exploratory review, including the search query, eligibility criteria, selection process, and data extraction and analysis. The primary goal of this work was to spark conversations around the topic of responsible research. As such, our review of conference guidelines is not intended to be systematic, but rather a starting point for understanding how responsible research is approached in HRI conferences.

A. Search Query

We searched for venues which published human-robot interaction research in 2022 via the Scopus database (as in previous works [11], [12]) from January 2022 to December 2022. We decided to search for venues in the year 2022, because at the time of our search (6th October 2023) some of the venues/conferences had not yet been held, and/or their proceedings were not yet available. In addition, we only searched for publications within 1 year to guarantee that each venue had published their proceedings only once.

The search query, extracted papers, and corresponding venues are publicly available in a GitHub repository[2].

B. Eligibility Criteria

We defined the following inclusion criteria:

- only conference venues (i.e., excluding journals, book chapters etc.). We believe that conferences are the perfect venues in which this provocative conversation around responsible HRI can take place, and so they play an important point of convergence to establish responsible guidelines;
- venues contained at least 7 papers about HRI. This threshold was decided by computing the median number of published papers for conferences with more than three HRI papers;
- venues have their own proceedings (not relying on journal publications, e.g., *Lecture Notes in Networks and Systems*);
- venues were held in 2022. Therefore, any bi-annual conference that was not organised in 2022 was excluded.

Note that we have also included short papers (e.g., Late-Breaking Work in the HRI conference). Although we acknowledge that the acceptance rate for these works is often higher than for full paper tracks, we have included them as they still undergo peer review and are published as part of conference proceedings.

C. Selection Process

We first extracted all the manuscripts from the Scopus database using the search query defined in Section II-A. Then, we identified all the corresponding venues of the manuscripts, and we computed the number of published papers for each venue. After that, we screened the venues according to the eligibility criteria reported in Section II-B.

D. Data Extraction and Analysis

We identified the following set of variables to extract and analyse from the selected venues: conference description (i.e., verbatim description of the conference from IEEE, ACM or Scopus websites); conference inception (i.e., the first year the conference was held); requirement or recommendation in CfP (i.e., description of the guidelines reported about CfP in the venue website); introduction year (i.e., the year in which the corresponding requirement or recommendation was introduced); level of mandate (i.e., whether the corresponding requirement or recommendation was mandatory or not); and details of description (i.e., how extensive the description and clarification of the corresponding requirement or recommendation is).

We extracted a total of 1396 manuscripts and 386 corresponding venues from the Scopus database using our search query. Then we screened the venues based on the eligibility criteria. We first selected the venues that published at least 7 papers in the field of HRI in 2022, resulting in a total of 30 venues. Then, we excluded the venues that were not conferences and did not have their own proceedings. This resulted in a final number of 18 venues (reported in Table I). Following the identification of relevant venues, we checked for the most up-to-date conference websites. For conferences which had already released their 2024 call for papers, we used this website, otherwise, we used the 2023 conference website. These websites served as the basis for the data extraction described above. This process ensured all conferences were being evaluated based on their most current practices.

III. Findings and Discussion

We identified categories of responsible requirements and recommendations from the venues selected following the framework method for qualitative analysis [13] (i.e., using the Responsible Research and Innovation as an analysis framework while allowing for other emergent observations in the data). Specifically, we first familiarised ourselves with the extracted variables, such as the requirements/recommendations in the CfP, we then analysed the

[2]https://github.com/ElmiraYadollahi/ResponsibleHRI.git

TABLE I: List of conferences selected from the review of the 2022 venues, the corresponding HRI paper counts, and the responsible categories they fulfill according to their most current website (2023/2024). *Legend*: E&HP - Ethical and Human Participants Considerations; A&I - Accessibility and Inclusion Guidelines; T&R: Transparency and Reproducibility Requirements; PL: Plagiarism, and LLMs.

Source title	**Count papers**	**Acronym**	**E&HP**	**A&I**	**T&R**	**PL**
ACM/IEEE International Conference on Human-Robot Interaction	174	HRI	✓	✓		
RO-MAN 2022 - 31st IEEE International Conference on Robot and Human Interactive Communication	113	ROMAN				
IEEE International Conference on Intelligent Robots and Systems	41	IROS				
Proceedings - IEEE International Conference on Robotics and Automation	38	ICRA				
Conference on Human Factors in Computing Systems - Proceedings	25	CHI		✓	✓	
HAI 2022 - Proceedings of the 10th Conference on Human-Agent Interaction	20	HAI			✓	
IEEE-RAS International Conference on Humanoid Robots	16	Humanoid				
2022 IEEE International Conference on Robotics and Biomimetics, ROBIO 2022	12	ROBIO				✓
2022 IEEE International Conference on Development and Learning, ICDL 2022	12	ICDL				✓
Proceedings of the IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics	12	BIOROB				
Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics	11	SMC				
2022 10th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos, ACIIW 2022	10	ACII	✓			
Proceedings of the 2022 IEEE International Conference on Human-Machine Systems, ICHMS 2022	9	ICHMS				
UMAP2022 - Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization	8	UMAP	✓	✓		
Proceedings of Interaction Design and Children, IDC 2022	8	IDC	✓			
ICARM 2022 - 2022 7th IEEE International Conference on Advanced Robotics and Mechatronics	8	ICARM				✓
Proceedings of the Annual Hawaii International Conference on System Sciences	7	HICSS				
IEEE/ASME International Conference on Advanced Intelligent Mechatronics, AIM	7	AIM				

selected theoretical framework, namely RRI, and finally identified the following categories: *Ethical & Human Participants Considerations* (E&HP), *Accessibility & Inclusion Guidelines* (A&I), *Transparency & Reproducibility Requirements* (T&R), and *Plagiarism, & LLM Use* (PL). We have merged together topics like transparency and reproducibility or ethics and human participants consideration that are interconnected among each other. For instance, even if research transparency refers to a range of open practices including registering studies, sharing study data and research reproducibility to instances in which the original researcher's data and codes are used to regenerate the results, their interplay contributes, for example, to replication of studies, i.e., refers to instances in which a researcher collects new data to arrive at the same scientific findings as a previous study. In the following sections, we discuss how each conference fits in the identified categories.

A. Ethical & Human Participant Considerations

Ethical considerations should include principles to protect the rights of participants, improve the validity of research, and guarantee scientific integrity [14] as for RRI recommendations [15]. Regarding HRI-specific ethical aspects, there are debates around the topics of deception by robots [16], dehumanisation of participants [17], and discrimination of participants and robots [18].

The UMAP, IDC, HRI, and ACII conferences included in their call for papers on their website specifications about ethics and human participants at different levels of priority. Since 2018, the IDC conference has required a mandatory section titled "Selection and Participation of Children" to include information related to the recruitment and ethical implications of involving children in user studies. They guide authors by providing a set of questions (e.g., "What benefits were there, if any, to the participants and their broader communities of being involved in this research?") to frame

their statement. The ACII conference introduced in 2022 a mandatory section entitled "Ethical Impact Statement", in which authors should discuss ethical aspects by following a detailed checklist provided by the conference[3]. The ACII community encourages authors to reflect and guard against the potential risks that could be raised by affect-related technology, including affective robots [19], [20].

Since 2022, the UMAP conference requires that the authors discuss ethical considerations and "comply with ethical standards and regulatory guidelines associated with human research, including research involving human participants and research using personally identifiable data". However, they do not provide specifics or guidelines on how these ethical discussions should be conducted. Since 2021, the HRI conference has included a link on their website to the ACM guidelines[4] about the involvement of human participants, without providing any customised recommendations, requirements, or checklists.

Currently, for HRI conferences, standards for ethical reporting are unclear. For example, many institutions also do not have access to Institutional Review Boards and thus cannot report ethical approval in such a way. As reported in Table I, the RO-MAN conference website does not have specific guidelines on this regard. Comprehensive guidelines on how to report ethical approval in the submission process would aid both the authors and the reviewers in assessing the ethical standards of submission.

B. Accessibility & Inclusion Guidelines

Accessibility, defined as "the qualities that make an experience open to all, regardless of ability", and inclusion, the process of designing technologies "for all" to promote diversity [21], represent fundamental principles within the RRI framework. Within the HRI community, no laws, rules, or standards specifically address HRI's accessibility needs. Only recently, Qbilat et al. have distilled a set of guidelines for accessible HRI [22], which included what users need to perceive, understand and operate about robot components during HRI. Conferences have referred to the concept of accessibility and inclusion not only in terms of technology (e.g., accessible robots) but also in terms of accessible and inclusive conferences. For example, many conferences have promoted the compliance of accessible papers (e.g., for people with visual impairments) and promoted inclusive participation (e.g., enabling hybrid participation in conferences).

Three conferences, namely UMAP, CHI and HRI, included accessibility and inclusion aspects in their submission guidelines mainly with a focus on accessible and inclusive conferences (rather than technologies). UMAP introduced on their website a step-by-step tutorial on how to make the manuscript accessible with a link to the SIGCHI accessibility guidelines in 2022. Since 2014, CHI added a dedicated page on their website[5] for accessibility guidelines. They provide a set of Q&As that can help participants understand how the conference is made accessible (e.g., "conference space", "can the conference be accessible by wheelchair or power scooter") and how they can be assisted during the conference. HRI only recently (2023) included instructions for accessible submissions linked to the ACM accessibility regulations, given that the theme of 2023 was "HRI for all". Several conferences (e.g., IDC, CHI, etc.), including HRI, have an accessibility chair who can also be contacted in case of specific accessibility concerns. Including a more transparent description of actions taken to promote accessibility could be the next step in improving accessibility and inclusion and raising awareness.

Again, the RO-MAN website does not report much about accessibility and inclusion. We believe that adding a specific page on the website dedicated to accessibility aspects may be beneficial for authors to improve inclusion in both the publication stage and their research design.

C. Transparency & Reproducibility Requirements

Calls for HRI to be more accountable in its reporting of scientific findings are becoming increasingly common, especially in the wake of the replication crisis [23], and the RRI framework encourages the community to conduct transparent research [15]. To this end, multiple works have been published that focus on quantitative and qualitative HRI research methodologies, with a large focus on how to design, conduct, and analyse (user) studies where at least one human and one robot are involved [24]–[26].

However, it is not always clear how to apply these recommendations within HRI. For example pre-registration, where hypotheses are registered online prior to the creation or collection of data, is now often considered a pre-requisite for publishing within psychology and related fields. On the other hand, HRI encompasses a diverse array of disciplines, not all of which require pre-registration or replication (e.g., the design field is not about generalisation but contextual innovation [27]).

Nonetheless, when a project includes directional, confirmatory, hypothesis-driven research, pre-registration should be encouraged [28]. In addition to preventing questionable statistical practices, pre-registration encourages researchers to clearly articulate their experimental design, number of participants, and planned analyses prior to data collection. These practices are also beneficial, regardless of if or what hypotheses are being tested.

Yet, the publication benefit for pre-registering papers within HRI is as of yet unclear. This could in part be due to lack of knowledge - early career researchers (e.g., first year PhD students) or researchers from other fields might simply be unfamiliar with these practices. However, it is likely also due to lack of incentive - pre-registration entails extra work on the part of researchers with (so far) no perceived benefit within the community.

Only two conferences identified in our review reported guidelines for promoting transparency and reproducibility. The HAI conference promotes open access, allowing the

[3]ACII submission guidelines

[4]https://www.acm.org/publications/policies/research-involving-human-participants-and-subjects

[5]https://chi2023.acm.org/for-attendees/chi2023-accessibility-faq/

uploading of pre-prints on open-access libraries such as arXiv. This further promotes open science and reproducibility by recognising open access publications. Nonetheless, there is tension between encouraging pre-prints whilst maintaining the integrity of the peer review process. This discussion is not unique to HRI but rather part of a larger conversation around publishing and open access in academia.

The CHI conference requires that "papers must include enough detail that the research can be reviewed for rigour and reproducibility". If the authors do not provide enough detail to evaluate the research, the paper can be rejected in the early stage. Transparency is of "utmost importance in a CHI paper" and the CHI website includes a dedicated page for transparency guidelines[6].Within its current guidelines, the RO-MAN conference provides no incentives for authors to engage in Open Science Practices.

D. Plagiarism & LLM Use

Plagiarism and the use of large language models (LLMs) in writing papers and conducting research are very recent increasing sources of concern for the research community in general [29]. LLMs have disrupted the way we interact with the world [30] as well as the future of research. Recent works [29] have highlighted many limitations of using LLMs in academic writing, such as repetitiveness and mediocrity of arguments in the research context. On the other hand, LLMs also have the potential to aid with scientific reporting, for example, proofreading text for spelling and grammar errors, coding, etc. The recent and fast rise of LLMs has not made it feasible for researchers in general and HRI communities in particular to distill a set of safe guardrails to use such tools. The few conferences that require disclosure of the use of LLMs do not specify in detail which information should be provided about its usage. On the contrary, with respect to plagiarism, many efforts have been put into developing tools that can assess whether a piece of writing is original or not.

Three conferences have very recently started including recommendations and/or requirements on plagiarism and the use of LLMs (in 2023). The ICDL and ICARM conferences require that the use of generative AI – such as LLMs – is "disclosed in the acknowledgments section", and the conferences check "against (self-)plagiarism" using the iThenticate tool. The ROBIO conference requires a "Similarity Report" that is generated via the CrossCheck platform. Note that we have only checked the conference websites. Some conferences, such as RO-MAN, IROS and ICRA include such similarity report checks in the submission system but without mentioning it on their websites.

Regarding the use of LLMs, we, as HRI researchers, recognize the nuances of attempting to catch up with its novelty. Therefore, HRI conferences, including RO-MAN, cannot (and potentially should not) prevent the use of such tools in academic writing; however, we can regulate them following the examples of university institutions.

[6]https://chi2023.acm.org/submission-guides/guide-to-a-successful-submission

IV. A Call To Action

Our results show that HRI-publishing conference guidelines are highly heterogeneous. Not a single venue provided guidelines across all of the aforementioned categories (Ethics, Accessibility, Transparency and Plagiarism) diverging from the RRI framework [15]. In particular, the RO-MAN website does not include any of the recommendations for conducting and reporting aspects of responsible research. While this is a reflection of the novelty of the HRI field and its highly dynamic evolution, it cannot be seen as an excuse – rather, an opportunity to have an "early start" at developing guidelines for high quality and responsible research and publications that meets RRI requirements.

We believe that conferences like RO-MAN are the perfect interface between the research that is developed in the field of HRI and the publication of this work to address RRI objectives. This is a unique position for offering guidelines for authors that ensure more transparent, ethical, and accessible research. Our work aims to spark discussions within the HRI community in general and RO-MAN in particular. In fact, this should be seen as a *responsibility* of conferences - no other stakeholder (except journals and publishers, for which this Call to Action should also apply) has so much international, inter-institutional, and interdisciplinary reach. In other research fields, such as HCI, the research community has been actively criticising and proposing best practices to improve the ethics of their community for many years [31]. For instance in 2022, [32] addressed publishing venues such as conferences to foster improved ethical practices that complement institutional and organizational ethics. We believe these community-driven efforts should be present also in the HRI research community, starting from RO-MAN.

Ethical and Human Participant Considerations. Following the RRI framework [15], HRI papers could, at a minimum, include an ethical and human participation statement. Considering that, currently, no universally unified ethical approval procedure is in place, and regulations are imposed either on a national or institution-wide level, such a statement could provide a medium that allows authors to present their ethical procedures structurally. As a result, authors could specify how they adhered to national/institutional guidelines as well as the field-related guidelines, as long as they all adhere to the Declaration of Helsinki. Further, a comprehensive ethical statement provides researchers with the chance to disclose potential impacts and/or misuse of their research, as well as what measures can and will be taken to prevent these. Whether such a statement is included within the page limit or not (such as in ACII and IDC) is up to individual venues to decide.

Accessibility and Inclusion Guidelines. We see other opportunities besides appointing a dedicated chair to make manuscripts accessible for accessibility guidelines, including specific guidelines or suggestions to promote inclusive use/design of robots. For example, we believe that a dedicated page on the conference website to guide authors on how to design/develop robots that may be accessible to all

can be beneficial for HRI researchers. Building upon the work by [22], HRI conferences, e.g., RO-MAN, may be the perfect venue to discuss further the implications of such deployment and design on inclusion and accessibility in HRI. Considerations on inclusion and accessibility of the research could also be included in an ethical impact statement that would follow the main body of the paper, as mentioned above. On other aspects of inclusivity, an increasing effort is also being made to include working parents and parents of young children so that they are also able to participate (in-person) in conferences [33]. This can include, for example, dedicated breastfeeding or daycare rooms, as seen in IROS[7], or additional funding to allow parents to bring their children with them. Some conferences (e.g., IROS[8]) or international communities (e.g., SIGCHI [9]) also offer travel grants and financial support to students from developing countries. The HRI community and conferences could promote similar initiatives to provide a more inclusive venue for researchers all around the world.

Transparency and Reproducibility Requirements. While there are positive examples of conferences providing guidelines for reproducibility and transparency (e.g., CHI, HAI), the HRI community is yet to catch up. Given the inter-disciplinarity of the field and high thematic heterogeneity, we recognize the difficulty in providing generalizable guidelines ("checklists") to researchers. Nonetheless, we outline some potential suggestions below.

With respect to statistical reporting, when applicable, standards are rapidly evolving. It is generally now considered best practice to report (or at the very least make available somewhere) all statistical findings, regardless of whether they were significant or not. There are also often debates regarding what statistical decisions are appropriate. Regardless of what statistical methods are chosen, authors should demonstrate a sufficient understanding of the data, explain their reasoning behind the analyses performed, and ensure that their results are appropriately reported.

In addition, HRI reviewers might at times lack the statistical knowledge to appropriately assess the methods and results reported in quantitative user studies. This leads to papers being published with insufficiently reported, if not blatantly incorrect, statistics. One possible solution to this is to have reviewers rate either their own level of statistical expertise or whether the paper requires advanced statistical knowledge to understand. For conferences, a "statistics chair" or committee could also be appointed. In cases where reviewers or associate editors/chairs do not feel confident evaluating the statistics reported, they would then have a point of contact to refer to.

At present, the HRI community places an emphasis on novel works which are intended to encourage discussion. Novelty and experimental rigour are thus currently being presented as mutually exclusive, and/or a difficult standard to obtain, but they are not, as suggested by the RRI framework [15]. Given the fast-paced nature of HRI, whether replication of findings is always realistic or should be aimed for is another discussion. The same study conducted 5 or 10 years apart might yield totally different results. Replication of robotics studies also poses additional difficulties (e.g., outdated code bases or hardware). Rather than avoid these issues, the community could develop their own standards for reporting of HRI methodologies such as including the type of robot used and method of control.

It is preferable to share, where possible, materials that enhance reproducibility such as links to stimuli or materials used, datasets, questionnaires, analyses, and codes. Again, these are things which require extra effort from researchers but are not currently incentivized within the HRI community. Perhaps conferences can include a short statement on open science and/or a link to a pre-registration template such as the Open Science Framework [10] in their call for papers. For example, the HRI conference website contains an "Additional Information" section in their call for papers containing information about human participants, formatting, supplementary materials etc. Information about open science could be easily added here. Another alternative would be to allow reviewers to nominate papers which they believe do a good job of adhering to Open Science Practices or award an open science badge. Although this process is necessarily subjective (as are all nominations), it would allow recognition of works which have put in extra effort towards supporting open science.

Plagiarism and LLM Use. Regarding the use of large-language models, conferences could require a statement related to disclaiming which parts of the manuscript and/or code and other materials have been generated with the help of generative AI systems. Generative AI can be used as a tool in writing but also in many other stages of the research (e.g. coding, image generation for user studies, design tools and brainstorming, etc.). Many institutions have already released a set of guidelines for the use of such generative models in academic assignments (e.g., The University of Cambridge[11]). With the fast-paced evolution of technology and technological tools, this set of guidelines is ever-evolving and requires constant revisiting. Discussions ought to be had within the community to decide how to make the best use of this technology in a way that complements progress, rigour, and ethical use.

V. CONCLUSION

HRI is an inherently heterogeneous field. Many of the issues touched upon here also have deeper complexities - who should be responsible for funding for travel grants, conference childcare or similar? What happens when inclusivity guidelines for different groups conflict? There are also multiple stakeholders involved in the publication cycle, from authors, reviewers, institutions, sponsors, and publishers, each which may have their own competing or conflicting

[7]https://ieee-iros.org/parenting/ and https://ieee-iros.org/child-care/

[8]https://ieee-iros.org/travel-grants/

[9]https://sigchi.org/awards/gary-marsden-travel-awards/

[10]osf.io

[11]https://www.cambridgeinternational.org/support-and-training-for-schools/artificial-intelligence/

interests. We also do not discount the fact that for many of the conferences discussed here, researchers volunteer their time in the interest of the greater community. We do not claim to have answers to all of these issues. Rather, we hope that our work serves as a starting point for further investigation. Our future work will focus on engaging the HRI community to distill a set of guidelines and recommendations for conducting **responsible HRI**.

This paper aims to initiate a conversation and discussion around Responsible HRI. The results from our review highlight a clear need to develop guidelines and language which can help guide HRI researchers towards this goal in compliance with the RRI framework. Such guidelines should neither be overly prescriptive (disallowing new and innovative ideas which may fall outside the bounds of traditional research methodologies) nor too vague (leaving too much to "unwritten rules"). We call all the researchers in the robotics community focusing on HRI related research to action for undertaking and promoting responsible research and defining such guidelines. We hope that, by sparking this discussion at RO-MAN 2024, the whole community will benefit, encouraging the robotics community to continue evolving in a more transparent, ethical, and inclusive way.

ACKNOWLEDGMENT

Funding: M. Spitale and H. Gunes have been supported by the EPSRC/UKRI under grant ref. EP/R030782/1 (ARoEQ) and EP/R511675/1. M. Spitale is supported by PNRR-PE-AI FAIR project funded by the NextGeneration EU program. **Open Access:** For open access purposes, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

REFERENCES

[1] S. Šabanović, C. A. Berry, and C. L. Bethel, "Introduction to the special issue on hri education," *Journal of Human-Robot Interaction*, vol. 6, no. 2, pp. 1–2, 2017.

[2] M. Burget, E. Bardone, and M. Pedaste, "Definitions and conceptual dimensions of responsible research and innovation: A literature review," *Science and engineering ethics*, vol. 23, pp. 1–19, 2017.

[3] T. Hagendorff, "The ethics of ai ethics: An evaluation of guidelines," *Minds and machines*, vol. 30, no. 1, pp. 99–120, 2020.

[4] K. Siau and W. Wang, "Artificial intelligence (ai) ethics: ethics of ai and ethical ai," *Journal of Database Management (JDM)*, vol. 31, no. 2, pp. 74–87, 2020.

[5] B. C. Stahl and D. Wright, "Ethics and privacy in ai and big data: Implementing responsible research and innovation," *IEEE Security & Privacy*, vol. 16, no. 3, pp. 26–33, 2018.

[6] M. Baker, "Reproducibility crisis," *Nature*, vol. 533, no. 26, pp. 353–66, 2016.

[7] M. Hutson, "Artificial intelligence faces reproducibility crisis," 2018.

[8] C. Power, P. Cairns, and M. Barlet, "Inclusion in the third wave: Access to experience," *New Directions in Third Wave Human-Computer Interaction: Volume 1-Technologies*, pp. 163–181, 2018.

[9] K. Mack, E. McDonnell, D. Jain, L. Lu Wang, J. E. Froehlich, and L. Findlater, "What do we mean by "accessibility research"? a literature survey of accessibility papers in chi and assets from 1994 to 2019," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–18.

[10] H. Gunes, F. Broz, C. S. Crawford, A. R.-v. der Pütten, M. Strait, and L. Riek, "Reproducibility in human-robot interaction: Furthering the science of hri," *Current Robotics Reports*, pp. 1–12, 2022.

[11] M. Spitale, R. Stower, E. Yadollahi, M. T. Parreira, N. I. Abbasi, I. Leite, and H. Gunes, "A systematic review on reproducibility in child-robot interaction," *arXiv preprint arXiv:2309.01822*, 2023.

[12] F. Catania, M. Spitale, and F. Garzotto, "Conversational agents in therapeutic interventions for neurodevelopmental disorders: a survey," *ACM Computing Surveys*, vol. 55, no. 10, pp. 1–34, 2023.

[13] S. Parkinson, V. Eatough, J. Holmes, E. Stapley, and N. Midgley, "Framework analysis: a worked example of a study exploring young people's experiences of depression," *Qualitative research in psychology*, vol. 13, no. 2, pp. 109–129, 2016.

[14] R. Wullenkord and F. Eyssel, "Societal and ethical issues in hri," *Current Robotics Reports*, vol. 1, pp. 85–96, 2020.

[15] R. Owen, P. Macnaghten, and J. Stilgoe, "Responsible research and innovation: From science in society to science for society, with society," *Science and public policy*, vol. 39, no. 6, pp. 751–760, 2012.

[16] A. Sharkey and N. Sharkey, "Granny and the robots: ethical issues in robot care for the elderly," *Ethics and information technology*, vol. 14, pp. 27–40, 2012.

[17] P. Asaro, ""hands up, don't shoot!": Hri and the automation of police use of force," *J. Hum.-Robot Interact.*, vol. 5, no. 3, p. 55–69, dec 2016. [Online]. Available: https://doi.org/10.5898/JHRI.5.3.Asaro

[18] M. M. de Graaf, "An ethical evaluation of human–robot relationships," *International journal of social robotics*, vol. 8, pp. 589–598, 2016.

[19] M. Spitale and H. Gunes, "Affective robotics for wellbeing: A scoping review," in *2022 10th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE, 2022, pp. 1–8.

[20] J. Cheong, M. Spitale, and H. Gunes, ""it's not fair!"–fairness for a small dataset of multi-modal dyadic mental well-being coaching," 2023.

[21] A. Sears and V. Hanson, "Representing users in accessibility research," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2011, pp. 2235–2238.

[22] M. Qbilat, A. Iglesias, and T. Belpaeme, "A proposal of accessibility guidelines for human-robot interaction," *Electronics*, vol. 10, no. 5, p. 561, 2021.

[23] C. Jost, B. Le Pévédic, T. Belpaeme, C. Bethel, D. Chrysostomou, N. Crook, M. Grandgeorge, and N. Mirnig, *Human-Robot Interaction*. Springer, 2020.

[24] G. Hoffman and X. Zhao, "A primer for conducting experiments in human–robot interaction," *J. Hum.-Robot Interact.*, vol. 10, no. 1, oct 2020. [Online]. Available: https://doi.org/10.1145/3412374

[25] P. Baxter, J. Kennedy, E. Senft, S. Lemaignan, and T. Belpaeme, "From characterising three years of hri to methodology and reporting recommendations," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 391–398.

[26] L. Veling and C. McGinn, "Qualitative research in hri: A review and taxonomy," *International Journal of Social Robotics*, vol. 13, no. 7, pp. 1689–1709, 2021.

[27] H. G. Nelson and E. Stolterman, *The design way: Intentional change in an unpredictable world*. MIT press, 2014.

[28] E.-J. Wagenmakers, R. Wetzels, D. Borsboom, H. L. van der Maas, and R. A. Kievit, "An agenda for purely confirmatory research," *Perspectives on psychological science*, vol. 7, no. 6, pp. 632–638, 2012.

[29] R. Fok and D. S. Weld, "What can't large language models do? the future of ai-assisted academic writing," in *In2Writing Workshop at CHI*, 2023.

[30] Y. Chang, X. Wang, J. Wang, Y. Wu, K. Zhu, H. Chen, L. Yang, X. Yi, C. Wang, Y. Wang, *et al.*, "A survey on evaluation of large language models," *arXiv preprint arXiv:2307.03109*, 2023.

[31] C. Munteanu, H. Molyneaux, W. Moncur, M. Romero, S. O'Donnell, and J. Vines, "Situational ethics: Re-thinking approaches to formal ethics requirements for human-computer interaction," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 105–114.

[32] N. E. Gold, R. Masu, C. Chevalier, and F. Morreale, "Share your values! community-driven embedding of ethics in research," in *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 2022, pp. 1–7.

[33] R. M. Calisi and a Working Group of Mothers in Science, "How to tackle the childcare–conference conundrum," *Proceedings of the National Academy of Sciences*, vol. 115, no. 12, pp. 2845–2849, 2018. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.1803153115