

# Exploring Large Language Model Interfaces Through Critical and Participatory Design

Helena A. Haxvig

*Dipartimento Di Ingegneria E Scienza Dell'Informazione, Università Di Trento, Via Sommarive 9, 38123 Povo, Italia*

## Abstract

This study presents the proposal for a critical exploration of Large Language Model (LLM) interfaces and their representations of gender equality and LGBTQ+ rights. The research methodology emphasizes the paradigm of Participatory Design, highlighting the importance of including diverse voices in technology development and empowering marginalized communities by collaborating with feminist and queer activists. The outcome of this empirical work is expected to consist of a method for interviewing LLM interfaces and input for a continued research agenda concerned with developing new frameworks for Artificial Intelligence systems that address critical social and ecological challenges evoked by existing technologies.

## Keywords

Human Computer Interaction (HCI), Human-Centered Artificial Intelligence (HCAI), Large Language Model Interface, ChatGPT, Participatory Design, Human Rights

## 1. Introduction

As numerous studies have revealed, the development of Artificial Intelligence (AI) systems in isolation from their contexts, especially their social and ecological realities, has resulted in unforeseen ramifications. Consequently, there is a pressing demand to establish frameworks capable of navigating the intricate and interdependent landscape of our increasingly complex world in which more and more technologies are entangled [1, 2, 3, 4, 5, 6, 7, 8]. Considering this emphasis, an expanding collection of literature [9, 10, 11, 12] offers recommendations for enhancing the considerations of Human-Centered system designs relating to AI, including *Human-Centered AI* [13, 14, 15], *Human-Centered Algorithm Design* [16, 17], *Human-Centered Data Science* [18, 19, 5], and *Human-Centered Machine Learning* [20, 4, 21]. Although there are contradictions among these perspectives, they also share significant similarities, especially in acknowledging the need to consider a broader range of solutions when designing systems for our intricate world. It is insufficient to solely seek technical remedies for problems that are inherently beyond the technical realm. The advancement of technology does not occur independently, and technologies have evolved into more than just passive instruments; they are social interventions that necessitate active participation in social, political, and ethical discussions [10].

---


*CHIItaly 2023: Crossing HCI and AI, Doctoral Consortium, 20–22 September 2023, Turin, Italy*

✉ [helenaamalie.haxvig@unitn.it](mailto:helenaamalie.haxvig@unitn.it) (H. A. Haxvig)

🆔 0009-0003-3858-6617 (H. A. Haxvig)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

## 2. Research Problem

Acknowledging the risk of unintentional bias in automated decision-making has led to the extensive field of Algorithmic Fairness, where multiple definitions and metrics of fairness have been proposed and tested. While some of these approaches are promising, they are limited regarding real-world applications [22]. Thus, it is necessary to move towards a multi-dimensional approach to algorithmic fairness, that goes beyond the quantitatively (mathematically) focused solutions primarily developed so far.

In recent years, the fields of Feminist HCI [23, 24, 25] and Queer HCI [26, 27, 28] have made significant strides in understanding and addressing the issues of gender equality and LGBTQ+<sup>1</sup> rights within the realm of Human-Computer Interaction (HCI), working with applications situated in real-world contexts. However, as societal developments continue to unfold, it becomes increasingly important to investigate the implications of technologies, such as Large Language Models (LLMs)<sup>2</sup>, acknowledging the potential harmful misrepresentation and reinforcement of problematic biases that exist within their use. This is especially a concern now that recent research has demonstrated how LLMs, such as the highly discussed ChatGPT developed by OpenAI, can seemingly exhibit political bias [30]. In extension, the continued dominance of American-led companies in the AI market [31] raises concerns relating to the direction of their politics, particularly regarding women’s rights and LGBTQ+ rights [32, 33, 34]. With restrictions on freedom and access to healthcare, as well as increasing censorship in various states, it is pertinent to question whether these perspectives will also be perpetuated by the LLMs produced in these political environments. Will they reinforce the kind of thinking that primarily serves one societal group, thereby diminishing the power of marginalized groups?

By critically examining the potential impact of LLMs on these critical issues, researchers can strive to create technology that fosters inclusivity, social sustainability, and the pluralistic coexistence of diverse perspectives, as well as empowers and uplifts marginalized communities.

### 2.1. Research Questions

This paper presents a sub-study that is part of a larger Ph.D. project, with the goal of developing a new framework for the use of Participatory Design [35] to incorporate human domain knowledge in the design of more holistic, context-aware, and sustainable AI systems. The goal therein is to include human domain knowledge into a system’s dataset and learning process, but also and most especially, to potentially include it in a continuous feedback and training loop. This is to secure that the quality and relevance of the dataset and model follow the development of the context in sustainable ways, instead of the context, i.e. humans, needing to adjust their reality to fit the system once it is implemented.

The present sub-study succeeds the preliminary development of a theoretical framework (named Socio-Ecologically Entangled AI, SEE-AI) and precedes future work of exploring how this can be applied in practice, e.g. by co-designing prototypes or looking into possibilities of

---

<sup>1</sup> The acronym stands for Lesbian, Gay, Bisexual, Transgender, Queer/Questioning, with the + representing additional gender identities and sexual orientations. The acronym can be found in other variations.

<sup>2</sup> “A large language model is a trained deep-learning model that understands and generates text in a human-like fashion.” [29]

retraining LLMs based on the SEE-AI framework. As such, the current focus relevant for this paper is on exploring issues with current LLM interfaces, specifically ChatGPT, in relation to the context of gender equality and LGBTQ+ rights. On that account, the research questions defining this proposed sub-study are as follows:

- How could a methodological 'interview' approach be developed to effectively explore Large Language Models (LLMs) – through interactions with the interface – on their portrayal of sensitive or problematic topics without encountering standardized non-answers?
- How can the insights and inputs provided by participants in workshops and collaborations with activist communities shape and inform the continued research agenda in addressing biases and critical issues of LLM interfaces faced by marginalized groups?

These research questions will guide the methodological development and investigation in this research endeavour, ensuring a focused approach contributing towards the broader goals of accountable and equitable artificial intelligence design.

### **3. Proposed approach**

With the positioning of this research being focused on issues of gender and LGBTQ+ representation in LLMs, there is a relevance in including and building upon established perspectives and research presented in the relatively new fields of Feminist HCI [23, 24, 25] and Queer HCI [26, 27, 28]. This study will also be positioned within the paradigm of Participatory Design (PD) [35] to abate some of the weaknesses of HCI regarding genuine user involvement. One particular reason is, that to address societal challenges and achieve a comprehensive perspective on technology development, it is crucial to adopt a human-centered approach that considers a broader concept of "humans" beyond simply the intended users.

#### **3.1. Participatory Design Values**

Participatory Design [35] has emerged as a pivotal research paradigm that primarily goes beyond involving users in the design process merely as a show of tokenism [36].

While early user-centered approaches in HCI emphasized user feedback and testing, PD moves beyond this relatively superficial user involvement and accentuates genuine inclusion of users and other people of interest in all stages of the design process, ensuring their agency and empowerment as active co-creators [37]. In addition, the cognitive approach, once dominant in HCI, has been critiqued by the PD field as narrow, calling for a broader consideration of social, cultural, and contextual factors that influence technology use [38]. By actively involving users, especially marginalized and underrepresented groups, in decision-making and co-design, a PD approach challenges traditional power dynamics and fosters more equitable and inclusive design outcomes [35], similar to the principles of Feminist HCI [23, 24, 25].

##### **3.1.1. Activist Communities and Participant Considerations**

Given the chosen social domain, this study intends to engage with activist communities advocating for gender equality and LGBTQ+ rights, starting locally in Trento, Italy. The heterogeneous

nature of activist communities calls for a pluriversal [39] perspective, recognizing that not all design approaches will work for every participant and emphasizing the importance of equal opportunities for all participants to have their voices heard and included [40, 37].

PD emphasizes the role of design in shaping the socio-political landscape through the production of artefacts, and the importance of engaging a diverse group of users and non-users in the process of designing these [35]. These dimensions will guide my engagement with activist communities and the aim to contribute to empowering these communities to navigate the complexities of technology and design while advancing their own goals and causes.

## **3.2. Considered Methods**

With the aim to investigate the gender and LGBTQ+ bias present in linguistic models and its implications, along with new ways of developing alternative models, the study will adopt a critical design [41, 37] approach to facilitate participatory exploration and agenda development.

### **3.2.1. Interviewing non-humans**

First of all, it is already known that systems are biased, as they will have values embedded within them, whether intentionally or not. However, seeing as LLMs are often programmed not to answer specific problematic questions [30], understanding the depth of particular adverse biases requires new approaches. Thus, this sub-study includes the intention to develop a new creative method of 'interviewing' LLMs – through their interfaces – as the subject itself, trying to figure out a strategy for circumventing the blockages of standard responses and accessing the core systematic ideologies encoded in the model. Of course, it is important to note that the concept of interviews normally concerns human subjects, and interactions with LLM interfaces are more considered an exchange of queries or prompts and outputs. However, in the spirit of also adopting more-than-human approaches [42, 43, 44], the intention is to include the non-human entity in a similar (although not identical) manner as a human entity, trying to figure out how it processes and presents information differently from humans, and identify harmful instances.

Using output from focus-group interviews with feminist and LGBTQ+ members and/or activists, the idea is to develop two vignettes [45] grounded in real situations. The vignettes are to be used as prompts, asking ChatGPT to describe the presented situation. Each vignette will be prompted nine times, crossing three different ways of articulating the vignettes with requesting to provide a description from three different viewpoints. The 18 outputs produced by the model will then be analyzed qualitatively with an interpretive lens formed from feminist and LGBTQ+ theories (including Feminist and Queer HCI literature), to interpret the representation of bias in the instances. To evaluate the results of these analyses, the output from the prompts will also be presented to the participants of the study to provide interpretations from perspectives build on the personal experience and contextual expertise of the involved groups.

### **3.2.2. Participation**

An important overarching methodology in the participatory critical design approach for this sub-study is engaging in collaboration with particularly local feminist and queer activists, with

the aim to first establish a shared understanding and identify considerations for developing a specific research agenda that aligns with their perspectives and goals.

Facilitating an 'Open Space Technology' [46] workshop, with variations, will provide a forum for open discussion and identification of important points related to gender bias in linguistic models. The purpose would be to encourage participants to freely contribute, express their concerns, and share their diverse insights to shape the subsequent research agenda.

This can then act as a precursor to an exploratory participatory design process using critical design techniques to provoke reflection and debate among participants, through e.g. 'Thinking Aloud protocols' [47]. The intention is to have the participants engage with LLM interfaces, specifically ChatGPT, while encouraging playfulness and exploration of the models to uncover unexpected uses and responses. This will be followed by facilitating activities that promote critical reflection, dialogue, and debate of the output received from the models.

## 4. Expected outcomes

While the methodological approaches for the study are still being developed and not entirely planned in detail, it is important to acknowledge that predicting future outcomes can be challenging. Nevertheless, I expect the following two general outcomes from this sub-study:

- **A Method for Interviewing Large Language Models:** One of the anticipated outcomes is the development of a method for effectively interviewing LLMs on topics they have been programmed not to answer. For instance, investigating how LLMs portray gender biases and uncovering the underlying mechanisms behind them. This method, translated into a practical tool (of which the form has not been decided yet), will contribute to current research on understanding and addressing biases in AI systems and increasing fairness. It must be noted that research in the areas of Interpretable- or Explainable Machine Learning also addresses bias and fairness, by developing models that can provide explanations and justifications about their decisions or predictions in a way that is understandable to humans [48]. However, this sub-study is not directly taking part in these conversations, as the aim is to provide a qualitative counterpart from the perspective of human-machine interaction, focused on how bias manifests in the interactions with the user interface. The reasoning follows a current discussion on how transparent and explainable AI are still algorithm-centered solutions, where a larger and more human-centered solution space should be considered instead to manage the uncertainty present in interactions with AI systems [49].
- **Input for Continued Research Agenda:** The research study will be inductively guided by the insights and inputs provided by the participants through different methods of active involvement. With the participants' voices and perspectives shaping the research agenda, it will ensure that the research remains relevant, addresses community needs, and focuses on the most critical issues faced by marginalized groups.

The idea for the next step after this sub-study, is to train an LLM on a self-curated dataset to experiment on two points: 1) The outcome of training it on selective data curated from the perspective of the theoretical SEE-AI framework mentioned previously. Thus, also joining the

conversation on creating specifically leaning chatbots, exemplified by David Rozado's creation of the RightWingGPT [50]; 2) exploring what could be the minimal sample size required to achieve fairness in relation to the representation of values of gender equality and LGBTQ+ rights.

Through these efforts, I aspire to make a valuable contribution to the advancement of more accountable, inclusive, and ethical technology practices in the realm of AI.

## References

- [1] A. Birhane, Algorithmic injustice: a relational ethics approach, *Patterns* 2 (2021) 100205. doi:10.1016/j.patter.2021.100205.
- [2] K. Crawford, *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*, Yale University Press, New Haven, 2021. OCLC: on1111967630.
- [3] V. Eubanks, *Automating inequality: how high-tech tools profile, police, and punish the poor*, first picador edition ed., Picador St. Martin's Press, New York, 2019.
- [4] M. Miceli, J. Posada, T. Yang, Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?, *Proceedings of the ACM on Human-Computer Interaction* 6 (2022) 34:1–34:14. URL: <https://doi.org/10.1145/3492853>. doi:10.1145/3492853.
- [5] M. Muller, I. Lange, D. Wang, D. Piorkowski, J. Tsay, Q. V. Liao, C. Dugan, T. Erickson, How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation, in: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, Glasgow Scotland Uk, 2019, pp. 1–15. URL: <https://dl.acm.org/doi/10.1145/3290605.3300356>. doi:10.1145/3290605.3300356.
- [6] S. U. Noble, *Algorithms of oppression: how search engines reinforce racism*, New York university press, New York, 2018.
- [7] C. O'Neil, *Weapons of math destruction: how big data increases inequality and threatens democracy*, Penguin Books, London, 2017.
- [8] A. Schmidt, The End of Serendipity: Will Artificial Intelligence Remove Chance and Choice in Everyday Life?, in: *CHIItaly 2021: 14th Biannual Conference of the Italian SIGCHI Chapter, CHIItaly '21*, Association for Computing Machinery, New York, NY, USA, 2021, pp. 1–4. URL: <https://doi.org/10.1145/3464385.3464763>. doi:10.1145/3464385.3464763.
- [9] N. Sambasivan, S. Kapania, H. Highfill, D. Akrong, P. Paritosh, L. M. Aroyo, "Everyone wants to do the model work, not the data work": Data Cascades in High-Stakes AI, in: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21*, Association for Computing Machinery, New York, NY, USA, 2021, pp. 1–15. URL: <https://doi.org/10.1145/3411764.3445518>. doi:10.1145/3411764.3445518.
- [10] C. Frauenberger, P. Purgathofer, Ways of thinking in informatics, *Communications of the ACM* 62 (2019) 58–64. URL: <https://doi.org/10.1145/3329674>. doi:10.1145/3329674.
- [11] C. Frauenberger, Entanglement HCI The Next Wave?, *ACM Transactions on Computer-Human Interaction* 27 (2019) 2:1–2:27. URL: <https://doi.org/10.1145/3364998>. doi:10.1145/3364998.
- [12] N. van Berkel, M. B. Skov, J. Kjeldskov, Human-AI interaction: intermittent, continuous,



- and proactive, *Interactions* 28 (2021) 67–71. URL: <https://doi.org/10.1145/3486941>. doi:10.1145/3486941.
- [13] O. Papakyriakopoulos, E. A. Watkins, A. Winecoff, K. Jaźwińska, T. Chattopadhyay, Qualitative Analysis for Human Centered AI, arXiv preprint arXiv:2112.03784 (2021).
- [14] W. Xu, Toward human-centered AI: A perspective from human-computer interaction, *Interactions* 26 (2019) 42–46. doi:10.1145/3328485.
- [15] M. O. Riedl, Human-centered artificial intelligence and machine learning, *Human Behavior and Emerging Technologies* 1 (2019) 33–36. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hbe2.117>. doi:10.1002/hbe2.117, \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbe2.117>.
- [16] M. Cherrington, D. Airehrour, J. Lu, Q. Xu, D. Cameron-Brown, I. Dunn, Features of Human-Centred Algorithm Design, in: 2020 30th International Telecommunication Networks and Applications Conference (ITNAC), 2020, pp. 1–6. doi:10.1109/ITNAC50341.2020.9315169, journal Abbreviation: 2020 30th International Telecommunication Networks and Applications Conference (ITNAC).
- [17] E. P. Baumer, Toward human-centered algorithm design, *Big Data & Society* 4 (2017) 2053951717718854. URL: <https://doi.org/10.1177/2053951717718854>. doi:10.1177/2053951717718854, publisher: SAGE Publications Ltd.
- [18] C. Aragon, S. Guha, M. Kogan, M. Muller, G. Neff, *Human-centered data science: An introduction*, MIT Press, 2022.
- [19] M. Muller, C. T. Wolf, J. Andres, M. Desmond, N. N. Joshi, Z. Ashktorab, A. Sharma, K. Brimijoin, Q. Pan, E. Duesterwald, C. Dugan, Designing Ground Truth and the Social Life of Labels, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21, Association for Computing Machinery, New York, NY, USA, 2021, pp. 1–16. URL: <https://doi.org/10.1145/3411764.3445402>. doi:10.1145/3411764.3445402.
- [20] S. Chancellor, E. P. S. Baumer, M. De Choudhury, Who is the "Human" in Human-Centered Machine Learning: The Case of Predicting Mental Health from Social Media, *Proceedings of the ACM on Human-Computer Interaction* 3 (2019) 147:1–147:32. URL: <https://doi.org/10.1145/3359249>. doi:10.1145/3359249.
- [21] J. Holbrook, J. Lovejoy, *Human-Centered Machine Learning*, 2018. URL: <https://medium.com/google-design/human-centered-machine-learning-a770d10562cd>.
- [22] X. Wang, Y. Zhang, R. Zhu, A brief review on algorithmic fairness, *Management System Engineering* 1 (2022) 7. URL: <https://link.springer.com/10.1007/s44176-022-00006-z>. doi:10.1007/s44176-022-00006-z.
- [23] S. Bardzell, Feminist HCI: taking stock and outlining an agenda for design, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '10, Association for Computing Machinery, New York, NY, USA, 2010, pp. 1301–1310. URL: <https://doi.org/10.1145/1753326.1753521>. doi:10.1145/1753326.1753521.
- [24] M. K. Scheuerman, J. M. Paul, J. R. Brubaker, How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services, *Proceedings of the ACM on Human-Computer Interaction* 3 (2019) 144:1–144:33. URL: <https://dl.acm.org/doi/10.1145/3359246>. doi:10.1145/3359246.
- [25] G. C. Pereira, M. C. C. Baranauskas, Gender identity and sexual orientation perceived oppressions in digital systems user interfaces: an exploratory study, in: Proceedings

- of the 14th Brazilian Symposium on Human Factors in Computing Systems, IHC '15, Association for Computing Machinery, New York, NY, USA, 2015, pp. 1–10. URL: <https://dl.acm.org/doi/10.1145/3148456.3148466>. doi:10.1145/3148456.3148466.
- [26] M. A. DeVito, C. Lustig, E. Simpson, K. Allison, T. Chuanromanee, K. Spiel, A. Ko, J. Rode, B. Dym, M. Muller, M. Klaus Scheuerman, A. Marie Walker, J. Brubaker, A. Ahmed, *Queer in HCI: Strengthening the Community of LGBTQIA+ Researchers and Research*, in: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, CHI EA '21*, Association for Computing Machinery, New York, NY, USA, 2021, pp. 1–3. URL: <https://dl.acm.org/doi/10.1145/3411763.3450403>. doi:10.1145/3411763.3450403.
- [27] Y. Strengers, L. Qu, Q. Xu, J. Knibbe, *Adhering, Steering, and Queering: Treatment of Gender in Natural Language Generation*, in: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20*, Association for Computing Machinery, New York, NY, USA, 2020, pp. 1–14. URL: <https://dl.acm.org/doi/10.1145/3313831.3376315>. doi:10.1145/3313831.3376315.
- [28] K. Spiel, O. Keyes, A. M. Walker, M. A. DeVito, J. Birnholtz, E. Brulé, A. Light, P. Barlas, J. Hardy, A. Ahmed, J. A. Rode, J. R. Brubaker, G. Kannabiran, *Queer(ing) HCI: Moving Forward in Theory and Practice*, in: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, CHI EA '19*, Association for Computing Machinery, New York, NY, USA, 2019, pp. 1–4. URL: <https://doi.org/10.1145/3290607.3311750>. doi:10.1145/3290607.3311750.
- [29] A. Tam, *What are Large Language Models*, 2023. URL: <https://machinelearningmastery.com/what-are-large-language-models/>.
- [30] D. Rozado, *The Political Biases of ChatGPT*, *Social Sciences* 12 (2023) 148. URL: <https://www.mdpi.com/2076-0760/12/3/148>. doi:10.3390/socsci12030148, number: 3 Publisher: Multidisciplinary Digital Publishing Institute.
- [31] *Artificial Intelligence Index Report 2023*, Technical Report, Stanford University: Human-Centered Artificial Intelligence (HAI), 2023. URL: <https://aiindex.stanford.edu/report/>.
- [32] E. Kilbride, “This Is Why We Became Activists”, *Human Rights Watch* (2023). URL: <https://www.hrw.org/report/2023/02/14/why-we-became-activists/violence-against-lesbian-bisexual-and-queer-women-and-non>.
- [33] *Center for Reproductive Rights, Abortion Laws by State*, 2023. URL: <https://reproductiverights.org/maps/abortion-laws-by-state/>.
- [34] *Human Rights Campaign, Attacks on Gender Affirming Care by State Map*, 2023. URL: [https://www.hrc.org/resources/attacks-on-gender-affirming-care-by-state-map?utm\\_medium=ads&utm\\_source=GoogleSearch&utm\\_content=GACMap-General&utm\\_campaign=GoogleGrant&utm\\_source=GS&utm\\_medium=AD&utm\\_campaign=BPI-HRC-Grant&utm\\_content=657195962808&utm\\_term=anti%20trans%20bills&gclid=CjwKCAjwpayjBhAnEiwA-7ena\\_-FUjeg0tZ4c-9xqHCm0H5tvARHfewr\\_-V1L7LuXS453iWMNYWLIrRoCDP4QAvD\\_BwE](https://www.hrc.org/resources/attacks-on-gender-affirming-care-by-state-map?utm_medium=ads&utm_source=GoogleSearch&utm_content=GACMap-General&utm_campaign=GoogleGrant&utm_source=GS&utm_medium=AD&utm_campaign=BPI-HRC-Grant&utm_content=657195962808&utm_term=anti%20trans%20bills&gclid=CjwKCAjwpayjBhAnEiwA-7ena_-FUjeg0tZ4c-9xqHCm0H5tvARHfewr_-V1L7LuXS453iWMNYWLIrRoCDP4QAvD_BwE).
- [35] J. Simonsen, T. Robertson (Eds.), *Routledge international handbook of participatory design*, Routledge international handbooks, Routledge, London, 2013. OCLC: 818827037.
- [36] S. R. Arnstein, *A Ladder Of Citizen Participation*, *Journal of the American Institute of Planners* 35 (1969) 216–224. URL: <http://www.tandfonline.com/doi/abs/10.1080/01944366908977225>. doi:10.1080/01944366908977225.



- [37] Bannon, Liam J., Ehn, Pelle, Design: design matters in Participatory Design, in: J. Simonsen, T. Robertson (Eds.), *Routledge international handbook of participatory design*, Routledge international handbooks, Routledge, London, 2013, pp. 37–63. OCLC: 818827037.
- [38] Kensing, Finn, Greenbaum, Joan, Heritage: having a say, in: J. Simonsen, T. Robertson (Eds.), *Routledge international handbook of participatory design*, Routledge international handbooks, Routledge, London, 2013, pp. 21–36. OCLC: 818827037.
- [39] A. Kothari, A. Salleh, A. Escobar, F. Demaria, A. Acosta (Eds.), "Pluriverse: A Post-Development Dictionary", 2019.
- [40] DiSalvo, Carl, Clement, Andrew, Pipek, Volkmar, Communities: Participatory Design for, with and by communities, in: J. Simonsen, T. Robertson (Eds.), *Routledge international handbook of participatory design*, Routledge international handbooks, Routledge, London, 2013, pp. 182–210. OCLC: 818827037.
- [41] A. Dunne, F. Raby, *Speculative everything: design, fiction, and social dreaming*, The MIT Press, Cambridge, Massachusetts ; London, 2013.
- [42] P. Coulton, J. Lindley, More-Than Human Centred Design: Considering Other Things, *The Design Journal* 22 (2019) 1–19. doi:10.1080/14606925.2019.1614320.
- [43] D. Loi, C. T. Wolf, J. L. Blomberg, R. Arar, M. Brereton, Co-designing AI Futures: Integrating AI Ethics, Social Computing, and Design, in: *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion*, DIS '19 Companion, Association for Computing Machinery, New York, NY, USA, 2019, pp. 381–384. URL: <https://dl.acm.org/doi/10.1145/3301019.3320000>. doi:10.1145/3301019.3320000.
- [44] I. Nicenboim, E. Giaccardi, M. L. J. Søndergaard, A. V. Reddy, Y. Strengers, J. Pierce, J. Redström, More-Than-Human Design and AI: In Conversation with Agents, in: *Companion Publication of the 2020 ACM Designing Interactive Systems Conference, DIS' 20 Companion*, Association for Computing Machinery, New York, NY, USA, 2020, pp. 397–400. URL: <https://doi.org/10.1145/3393914.3395912>. doi:10.1145/3393914.3395912.
- [45] C. Barter, E. Renold, The Use of Vignettes in Qualitative Research, *Social Research Update* 25 (1999). URL: <https://sru.soc.surrey.ac.uk/SRU25.html>.
- [46] H. Owen, *Open Space Technology: A User's Guide*, third edition ed., Berrett-Koehler Publishers, San Francisco, Calif, 2008.
- [47] K. A. Ericsson, H. A. Simon, *Protocol analysis: Verbal reports as data*, Rev. ed, Protocol analysis: Verbal reports as data, Rev. ed, The MIT Press, Cambridge, MA, US, 1993. Pages: liii, 443.
- [48] F. Doshi-Velez, B. Kim, Towards A Rigorous Science of Interpretable Machine Learning, 2017. URL: <http://arxiv.org/abs/1702.08608>. doi:10.48550/arXiv.1702.08608, arXiv:1702.08608 [cs, stat].
- [49] Z. T. Zhang, H. Hußmann, How to Manage Output Uncertainty: Targeting the Actual End User Problem in Interactions with AI, volume 2903, 2021. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85110534161&partnerID=40&md5=45ef43fd6796cf3a8c88c4f6295d9464>.
- [50] D. Rozado, RightWingGPT – An AI Manifesting the Opposite Political Biases of ChatGPT, 2023. URL: <https://davidrozado.substack.com/p/rightwinggpt>.