

Anticipating Risks and Identifying Governance Measures for the Use of genAI and FPT: A Participatory Procedure to Elicit Citizens' Perspectives

Chiara Ullstein*
chiara.ullstein@tum.de
Chair of Cyber Trust,
Department of Computer Science,
Technical University of Munich
Munich, Germany

Michel Hohendanner*
michel.hohendanner@tum.de
Chair of Cyber Trust,
Department of Computer Science,
Technical University of Munich
Munich, Germany

Jens Grossklags
jens.grossklags@in.tum.de
Chair of Cyber Trust,
Department of Computer Science,
Technical University of Munich
Munich, Germany

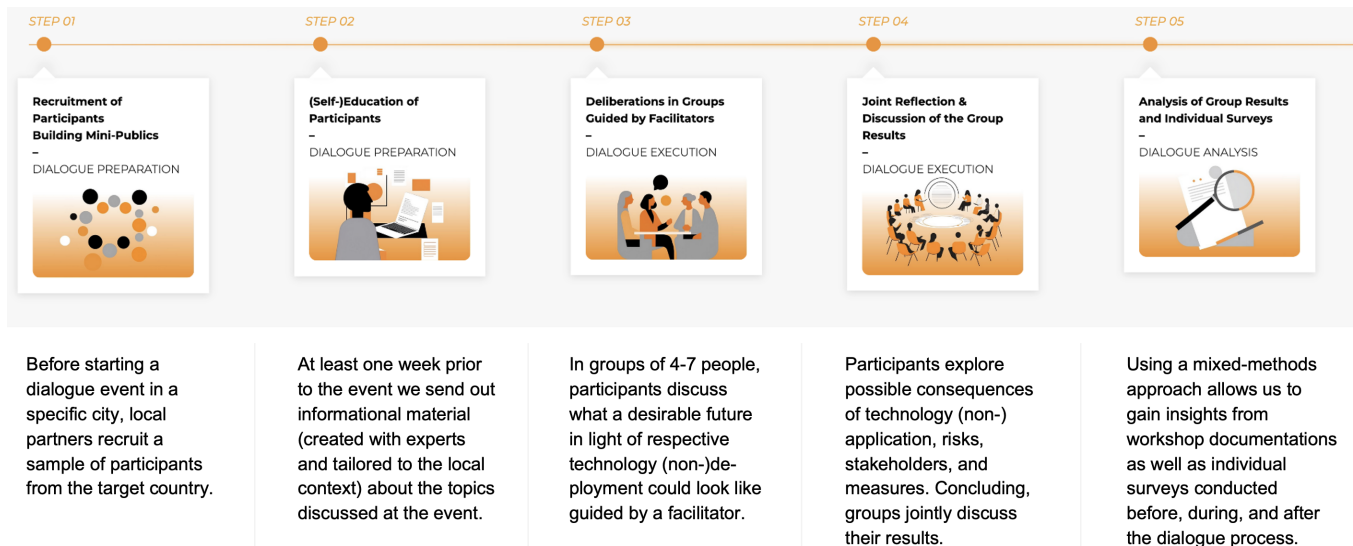


Figure 1: Process of the Global AI Dialogues from Participant Recruitment to Data Analysis | Source: <https://designxdiscourse.com/global-ai-dialogues/>

Abstract

With the increasingly rapid development and release of AI systems, policy discourses primarily take place on an expert level. Aiming to broaden the discourse, we propose the exploration of laypeople's informed opinions as a measure to evaluate the social impact of AI systems, and to inform forward-looking policies. We conceived and organized a dialogue series, the *Global AI Dialogues*, inviting citizens around the world to engage, discuss, and contribute their perspectives on AI. The goal was to better understand how people worldwide evaluate the (social) impact of AI on their everyday lives today and in the future, given the real-world challenges of their local contexts. During the dialogues, 284 participants across six countries (Germany, Nigeria, Japan, India, Mexico, Bolivia) critically engaged with what a desirable future in light of generative AI (genAI) and Facial Processing Technologies (FPT) could look like. They explored the consequences of technology deployment, assessed risks, mapped stakeholders, and derived measures to achieve a desirable goal. We contribute to sociotechnical AI governance by presenting a participatory procedure to anticipate high-priority

*Both authors contributed equally to this research.

risks, from the perspective of citizens. In this workshop paper, we briefly summarize the project by outlining our method, including the citizen dialogue workshop process, and provide first results. We show how our methodological approach leads participants to identify where to focus governance efforts. We present empirical data on citizens' needs and goals in AI governance from our dialogues around the world and two sociotechnical scenarios (genAI and FPT).

CCS Concepts

• **Human-centered computing** → **Empirical studies in HCI**; **Empirical studies in collaborative and social computing**; • **Social and professional topics** → *Computing / technology policy*.

Keywords

citizen dialogue, civic participation, participatory AI, stakeholder involvement, public perception, generative artificial intelligence, facial processing technologies, AI governance

1 Introduction

With the increasingly rapid development and release of AI systems, policy discourses take place primarily on an expert level. Aiming to broaden the discourse, we propose exploring the informed opinions of laypeople as a measure to evaluate the social impact of AI systems and to inform forward-looking policies.

In prior work, we surveyed 1070 laypeople cross-nationally about their perceptions of AI and 61 policymakers and advisors about what they want to know from laypeople to inform AI policymaking [30]. Building on these results, we initiated the *Global AI Dialogues*¹, a dialogue series that invites citizens around the world to engage, discuss, and contribute their perspectives on AI. The goal was to better understand how people around the world evaluate the (social) impact of AI on their daily lives today and in the future, given the real-world challenges of their local contexts. During the dialogues, participants critically engaged with what a desirable future in light of generative AI (genAI) and Facial Processing Technologies (FPT) could look like. These two AI technologies exemplify the diversity of potential impacts on people's everyday lives.

We contribute to sociotechnical AI governance by presenting a participatory procedure for **anticipating high-priority risks**, from the perspectives of citizens. We describe how we **designed the dialogues**. We show how our methodological approach leads participants to **identify where to focus governance efforts**. We present **empirical data of citizens' needs and goals in AI governance** from our dialogues around the world (Nigeria, Japan, Germany, India, Mexico, Bolivia) and **two sociotechnical scenarios** (genAI and FPT).

In the following, we summarize the project based on the paper 'Initiating the Global AI Dialogues: Laypeople Perspectives on the Future Role of genAI in Society from Nigeria, Germany, and Japan' published at CHI'25 [18] and a respective forthcoming working paper on the analysis of the FPT context.

2 Method

Figure 1 provides an overview of the process of the Global AI Dialogues from participant recruitment to data analysis. Figure 2 provides visual impressions from the dialogues in different countries.

2.1 Overview

Participant Recruitment, Selection, and Compensation. We organized one-day laypeople dialogues in Nigeria, Japan, Germany, India, Mexico, and Bolivia from July to December 2024. The team that conducted and analyzed these workshops consisted of researchers from and based in each of the countries. With the aim of providing a space for discourse and comparing laypeople's perspectives across the countries, we decided on a mix of participant selection methods [27]. We recruited 284 participants via targeted recruitment and self-selection, and semi-randomly selected from all registrants via purposive sampling, creating mini-publics [8, 22, 27]. Participants indicated whether they would like to participate in the dialogue on genAI or FPT. We sampled based on the criteria gender, age, and AI knowledge such that each dialogue workshop group had

a gender-balanced, age-diverse, and AI-knowledge-diverse composition. From all registered participants with comparable demographics (age, gender, AI knowledge), we randomly selected as many as required to achieve diversity in each dialogue group, which we achieved for most criteria for all groups. Participants received monetary compensation for participation. The amount and format of monetary compensation were decided on by the country teams after consultation with the initiating researchers to ensure fairness across countries.

Methodological Contextualization and Dialogue Framework. We situate our workshop dialogue approach in the rich conceptual background and methodological history of participatory workshops in HCI [13, 25, 29, 32]. To conduct the dialogue events, we provided a comprehensive workshop structure building on methods from design futuring [4, 5, 21], future studies [5, 9, 23, 24, 28], narrative and digital ethics [10, 16] as well as technology assessment [12]. These methods, stemming from different domains, all integrate a participatory element. In that sense, we build upon and are inspired by participatory speculative design [e.g. 7, 16], participatory futures research [1, 3, 20], narrative [2, 11] and digital (media) ethics [6, 11], and participatory technology assessment [12]. The workshop process also builds upon prior work by the authors conducting risk assessment formats [31], eliciting sociotechnical discourse through (participatory) speculative design [15–17, 19], and assessing their value to inform policymaking [14]. One to two weeks before the dialogues, participants received informational material vetted by experts for self-education. At the dialogues, after introductory talks, participants tested the technology, discussed and assessed the potential benefits and risks of specific AI applications, and evaluated whether the technology could be part of a desirable future. If so, participant groups considered stakeholders and their roles and formulated measures they would like to see implemented. Concluding, groups shared their vision for the technology in the future and engaged in joint discussions. The dialogues were accompanied by three surveys to collect individual reflections.

Data Analysis. We performed a mixed-method analysis. Country analysis teams applied initial/open coding, evaluation coding, and focused coding on their qualitative data through multiple rounds of coding [26]. Themes were then jointly discussed across country teams in two analysis workshop sessions. We applied frequency analysis, analyses of variances or Kruskal-Wallis tests, and pairwise comparisons on the quantitative data from the surveys.

2.2 Citizen Dialogue Workshop Process

Workshop Design and Structure. The dialogue workshops were structured as 6-hour and 3.5-hour (applied in Japan) sessions. Participants were welcomed and introduced to the research project, followed by an initial survey assessing their perceptions of AI and genAI or FPT. After receiving foundational knowledge through short talks from the research team, participants completed a second survey regarding their prior experiences with genAI or FPT. They then engaged in group discussions, utilizing pre-designed worksheets for guidance. These materials were collaboratively developed with local partners and external reviewers, ensuring local-cultural relevance and clarity through translation into local languages. Group facilitators explained the worksheets, moderated discussions, and documented the process through field notes. A

¹**Project Website:** <https://designxdiscourse.com/global-ai-dialogues/>

Project Video: <https://youtu.be/FINAjEWo8nw>



Figure 2: Dialogues on genAI in Nigeria, Japan, and Germany with laypeople stratified across gender, age, and AI knowledge levels. Copyright: Authors of the paper; photo consent obtained from participants. | Source: [18]

final survey helped participants to reflect on genAI or FPT and their experiences throughout the day.

In the following, we detail the individual steps of the workshop process (please note that groups were separated according to their selected technology, hence groups either engaged with genAI or FPT during the exercises described below):

Introduction. Each group, consisting mostly of 4-7 participants, began with introductions and an icebreaker to share their experiences with genAI or FPT. A practical task involved testing freely available tools, fostering hands-on engagement with the technology. Groups established a code of conduct to ensure equitable participation and a respectful dialogue.

Topic Exploration and Consequence Mapping. Groups were introduced to three pre-selected topic areas through summaries on data cards, supplemented by existing use cases prevalent in public media (if possible, from the respective local context). For FPT: FPT and public spaces and policing; FPT and corporate surveillance; FPT and social interaction. For genAI: genAI and education; genAI and public service; genAI and arts, culture & creative industries. Groups selected a topic and a corresponding "What-if" question (e.g., for a FPT topic, *What if FPT was established as policing tool?* – or, for a genAI topic, *What if genAI systems were established in public service contexts?*) to guide their discussions. Using consequence mapping, groups brainstormed potential outcomes of their chosen scenario, employing a visual method to explore direct and indirect consequences. Risk impulse cards were introduced to account for critical perspectives derived from academic literature on the implications of genAI or FPT.

Prioritization and Stakeholder Mapping. After mapping consequences, groups voted on the most significant outcomes and assessed their likelihood and impact. This led to stakeholder mapping, where groups identified relevant actors and strategies for achieving an anticipated desired outcome dealing with genAI or FPT in the respective application context, utilizing an adapted narrative analysis model [10].

Backcasting for Future Scenarios. In the final phase, groups engaged in backcasting. They articulated a future desirable status quo and identified necessary measures to achieve it, guided by backcasting impulse cards that prompted reflection on possible

steps. The workshops concluded with a plenary session where groups shared their backcasting.

3 Results

The following summarizes the first analyses. Section 3.1 refers to all countries. The genAI results refer to Nigerian, Japanese and German participants' perspectives. The FPT results additionally cover perspectives from Bolivian, and Indian participants.

3.1 Need for Regulation and Responsible Institutions

After the dialogue, in their individual survey reflection, participants in both dialogues (genAI and FPT) indicated that the technologies could be part of a desirable future (genAI: 96%-100%; FPT: 75%-100%), as long as the technologies satisfy certain requirements and if appropriate measures are implemented. 93% of respondents in the context of genAI and 96% of respondents in the context of FPT across all countries indicated that there should be legally binding regulatory measures for genAI and FPT, partially in combination with voluntary measures. When asked which institutions they would trust the most in establishing measures that make the use of AI in general safe, participants from Nigeria and Bolivia considered that companies developing AI were responsible. Across all countries, participants discussing FPT rated an international AI safety institute followed by intergovernmental organizations as among the most trustworthy institutions. Participants discussing genAI, in particular, highlighted an independent regulator as one of the most trustworthy institutions. Across all countries and both contexts (except Bolivia and Mexico), the government ranked as the second or third most trustworthy institution. Civil society organizations were considered relevant by German and Mexican participants discussing genAI, and German, Japanese, and Nigerian participants discussing FPT.

3.2 Facial Processing Technologies

For FPT, we find that, overall, ideas of FPT leading to increased public safety and increased convenience were dominant despite the

participants acknowledging limitations and risks. Participants from India and Germany were more critical of FPT than participants from Bolivia, Nigeria, and Japan. Within their groups, participants identified risks they perceived as most relevant. We classified these risks into five groups: misuse and lack of governance (misuse of system or data, lack of appropriate governance); inaccuracy of classification or recognition results (risk of inaccurate results, limitations of training datasets, bias and discrimination, concerns of validity and effectiveness); loss of rights and behavior change (loss of freedom of expression, loss of privacy and anonymity, loss of diversity and self-determination, change in behavior); economic risks; and environmental risks. Participants suggested regulatory measures that can be clustered into six groups: technological requirements, governance measures (e.g., limited and justified use, opt-out policies, protection of human rights, or standardization for quality assurance), monitoring and evaluation mechanisms, implementation measures, educational and awareness measures, and dialogue measures. They argued for keeping society in the loop through dialogue measures and transparency structures for governmental decisions as well as corporate developers. We highlight that while citizens attribute high importance to regulation, risks remain that cannot be solved merely through regulation.

3.3 Generative AI

For genAI, despite differences in experiences, exposure, and media coverage, citizens across all countries perceived genAI as an enabling technology that can make – if certain requirements are

met – valuable goods such as knowledge (in the context of education) or public service processes (in the context of public services) more accessible. The requirements for genAI to create value include: accessibility; education and reflective use; safety and robustness; privacy protection; governance structures; genAI awareness; and ecological awareness. Participants highlighted, in particular, five measures for achieving a desirable future with genAI: ensuring education and conscious societies; fostering technical innovation and best practices; ensuring regulation; building infrastructure; and building area expertise for system implementation.

4 Concluding Remarks

Our results highlight the complexity of expectations towards technology, and citizens' needs and aims for AI governance, also taking local factors into account. Citizens could assess the multitude of risks and identify necessary conditions, such as the need for accurate and robust or fair and bias-free technologies – challenges that cannot be solved only through technical means. With our work, we contribute to policy debates on AI by providing recommendations derived from participants' identified requirements and suggested measures for AI to create value and foster a socially desirable future. More specifically, we contribute to the workshop by presenting a participatory procedure for anticipating risks, from the perspective of citizens. Our empirical data highlights citizens' needs and goals in AI governance from six countries around the world and two sociotechnical scenarios: genAI and FPT (non-)application in the light of a desirable future.

Acknowledgments

We thank all participants for joining the dialogues. We appreciate all our project partners' energy and time contributed to this project and thank Arisa Ema, Amelia Katirai, Jun Kuribayashi, Aimi Ozaki, and Tatsuya Daikoku from the team in Japan, Bukola Abimbola Onyekwelu and Olusola Babalola from the team in Nigeria, Nidhi Singh, Srija Naskar, Shashank Mohan, and Nikhil Sharma from the team in India, Cynthia Maria Montaudon-Tomas, Gabriela Sánchez Bazán, and Christin Schüller from the team in Mexico, and Naira Paola Arnez-Jordan, Jose Jesus Cabrera-Pantoja, and Indiana Karina Jordan-Barros from the team in Bolivia. We thank all facilitators for supporting the dialogues. We are also grateful for the experts who have provided feedback, amongst others Emilia Gómez (Joint Research Center, EC), Orestis Papakryiakopoulos (TUM), Benedikt Zönnchen (MUCDAI/HM), and Téó Sanchez (LMU).

References

- [1] Anna Riikka Airiina Armanto. 2024. Futures participation as anticipatory practice – What do futures workshops do? *European Journal of Futures Research* 12, Article 3 (2024), 14 pages. doi:10.1186/s40309-024-00226-4
- [2] Clive Baldwin. 2015. Narrative ethics. In *Encyclopedia of Global Bioethics*, Henk ten Have (Ed.). Springer International Publishing, Cham, Switzerland. doi:10.1007/978-3-319-05544-2_302-1
- [3] Laura Barendregt, Roy Bendor, and Bregje F. van Eekelen. 2024. Public participation in futuring: A systematic literature review. *Futures* 158, Article 103346 (2024), 15 pages. doi:10.1016/j.futures.2024.103346
- [4] Anthony Dunne and Fiona Raby. 2013. *Speculative Everything: Design, Fiction, and Social Dreaming*. MIT Press, Cambridge, MA. <https://mitpress.mit.edu/9780262019842/speculative-everything/>
- [5] Felix Anand Epp, Tim Moesgen, Antti Salovaara, Emmi Pouta, and İdil Gaziulusoy. 2022. Reinventing the wheel: The future ripples method for activating anticipatory capacities in innovation teams. In *Designing Interactive Systems Conference (DIS '22)*. Association for Computing Machinery, New York, NY, USA, 387–399. doi:10.1145/3532106.3534570
- [6] Charles Ess. 2017. Digital media ethics. In *Oxford Research Encyclopedia of Communication*. Oxford University Press, Oxford, UK. doi:10.1093/acrefore/9780190228613.013.508
- [7] Pedro Gil Farias, Roy Bendor, and Bregje F. van Eekelen. 2022. Social dreaming together: A critical exploration of participatory speculative design. In *Proceedings of the Participatory Design Conference 2022 – Volume 2* (Newcastle upon Tyne, United Kingdom) (PDC '22). Association for Computing Machinery, New York, NY, USA, 147–154. doi:10.1145/3537797.3537826
- [8] Archon Fung. 2006. Varieties of participation in complex governance. *Public Administration Review* 66, s1 (2006), 66–75. doi:10.1111/j.1540-6210.2006.00667.x
- [9] Jerome C. Glenn. 2009. The futures wheel. In *Futures Research Methodology – Version 3.0*, Jerome C. Glenn and Theodore J. Gordon (Eds.). The Millennium Project, Washington, D.C. <https://millennium-project.org/publications-2/futures-research-methodology-version-3-0-2/>
- [10] Algirdas Julien Greimas. 1984. *Structural Semantics: An Attempt at a Method*. University of Nebraska Press, Lincoln, NE. <https://www.etera.ee/zoom/50614/view>
- [11] Petra Grimm, Tobias Keber, and Oliver Zöllner. 2019. *Digitale Ethik: Leben in vernetzten Welten*. Reclam Verlag, Ditzingen, Germany. https://www.reclam.de/detail/978-3-15-015240-9/Digitale_Ethik
- [12] Armin Grunwald. 2018. *Technology Assessment in Practice and Theory*. Routledge, London, UK. doi:10.4324/9780429442643
- [13] Christina N. Harrington, Katya Borgos-Rodriguez, and Anne Marie Piper. 2019. Engaging low-income African American older adults in health discussions through community-based design workshops. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, UK) (CHI '19). Association for Computing Machinery, New York, NY, USA, Article 593, 15 pages. doi:10.1145/3290605.3300823
- [14] Michel Hohendanner and Chiara Ullstein. 2024. Speculative Design for Policy? Zur politischen Anschlussfähigkeit von Speculative Design & Design Fiction Ansätzen. In *Wie gestalten wir Gesellschaft? Interdependenzen zwischen Design und dem Feld des Sozialen*, Christoph Rodatz, Iris Ebert, and Sebastian Rahn (Eds.). transcript, Bielefeld, Germany, 201–222. doi:10.14361/9783839472415-010
- [15] Michel Hohendanner, Chiara Ullstein, Yosuke Buchmeier, and Jens Grossklags. 2023. Exploring the reflective space of AI narratives through speculative design in Japan and Germany. In *Proceedings of the 2023 ACM Conference on Information Technology for Social Good* (Lisbon, Portugal) (GoodIT '23). Association for Computing Machinery, New York, NY, USA, 351–362. doi:10.1145/3582515.3609554
- [16] Michel Hohendanner, Chiara Ullstein, Dohjin Miyamoto, Emma Fukuwatari Huffman, Gudrun Socher, Jens Grossklags, and Hirotaka Osawa. 2024. Metaverse perspectives from Japan: A participatory speculative design case study. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2, Article 400 (2024), 51 pages. doi:10.1145/3686939
- [17] Michel Hohendanner, Chiara Ullstein, and Daijiro Mizuno. 2021. Designing the exploration of common good within digital environments: A deliberative speculative design framework and the analysis of resulting narratives. In *Swiss Design Network Symposium 2021 Conference Proceedings*. SUPSI, HSLU, swissdesignnetwork, Virtual Event Swiss, 566–579.
- [18] Michel Hohendanner, Chiara Ullstein, Bukola Abimbola Onyekwelu, Amelia Katirai, Jun Kuribayashi, Olusola Babalola, Arisa Ema, and Jens Grossklags. 2025. Initiating the Global AI Dialogues: Laypeople perspectives on the future role of genAI in society from Nigeria, Germany and Japan. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '25). Association for Computing Machinery, New York, NY, USA. doi:10.1145/3706598.3714322
- [19] Michel Hohendanner, Chiara Ullstein, Gudrun Socher, and Jens Grossklags. 2024. "Good and scary at the same time" – Exploring citizens' perceptions of a prospective metaverse. *IEEE Pervasive Computing* 23, 1 (2024), 27–36. doi:10.1109/MPRV.2024.3366112
- [20] Sohail Inayatullah. 2012. Futures studies: Theories and methods. In *There's a Future: Visions for a Better World*. BBVA, Madrid, Spain, 37–65.
- [21] Leon Karlsen Johannessen, Martina Maria Keitsch, and Ida Nilstad Pettersen. 2019. Speculative and critical design – Features, methods, and practices. *Proceedings of the Design Society: International Conference on Engineering Design* 1, 1 (2019), 1623–1632. doi:10.1017/dsi.2019.168
- [22] OECD. 2020. *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave*. OECD Publications, Paris, France. doi:10.1787/339306da-en
- [23] John Robinson. 1982. Energy backcasting: A proposed method of policy analysis. *Energy Policy* 10, 4 (1982), 337–344. doi:10.1016/0301-4215(82)90048-9
- [24] John Robinson. 2003. Future subjunctive: Backcasting as social learning. *Futures* 35, 8 (2003), 839–856. doi:10.1016/S0016-3287(03)00039-9
- [25] Daniela K. Rosner, Saba Kaway, Wenqi Li, Nicole Tilly, and Yi-Chen Sung. 2016. Out of time, out of place: Reflections on design workshops as a research method. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) (CSCW '16). Association for Computing Machinery, New York, NY, USA, 1131–1141. doi:10.1145/2818048.2820021
- [26] Johnny Saldaña. 2013. *The Coding Manual for Qualitative Researchers* (2nd ed.). SAGE, Los Angeles, CA.
- [27] Daniel Steel, Naseeb Bolduc, Kristina Jenei, and Michael Burgess. 2020. Rethinking representation and diversity in deliberative minipublics. *Journal of Deliberative Democracy* 16, 1 (2020), 46–57. doi:10.16997/jdd.398
- [28] William J. Sutherland and Harry J. Woodroof. 2009. The need for environmental horizon scanning. *Trends in Ecology & Evolution* 24, 10 (2009), 523–527. doi:10.1016/j.tree.2009.04.008
- [29] Sarina Till, Jaydon Farao, Toshka Lauren Coleman, Londiwe Deborah Shandu, Nonkululeko Khuzwayo, Livhuwani Muthelo, Masenyani Oupa Mbombi, Mmare Bopane, Molebogeng Motlathledi, Gugulethu Mabena, Alastair Van Heerden, Tebogo Maria Mothiba, Shane Norris, Nervo Verdezoto Dias, and Melissa Densmore. 2022. Community-based co-design across geographic locations and cultures: Methodological lessons from co-design workshops in South Africa. In *Proceedings of the Participatory Design Conference 2022 - Volume 1* (Newcastle upon Tyne, United Kingdom) (PDC '22). Association for Computing Machinery, New York, NY, USA, 120–132. doi:10.1145/3536169.3537786
- [30] Chiara Ullstein, Michel Hohendanner, Nikhil Sharma, Simeon Emilov Ivanov, Georgi Tsipov, and Jens Grossklags. 2025. Bridging perspectives for socially sustainable AI: What people across 12 countries think about genAI and FPT, and what policymakers want to know their opinions on. In *Proceedings of the 8th ACM SIGCAS Conference on Computing and Sustainable Societies* (Toronto, Canada) (COMPASS '25). Association for Computing Machinery, New York, NY, USA.
- [31] Chiara Ullstein, Julia Katharina Pfeiffer, Michel Hohendanner, and Jens Grossklags. 2024. Mapping the stakeholder debate on facial recognition technologies: Review and stakeholder workshop. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing* (San Jose, Costa Rica) (CSCW Companion '24). Association for Computing Machinery, New York, NY, USA, 429–436. doi:10.1145/3678884.3681887
- [32] Allison Woodruff, Renee Shelby, Patrick Gage Kelley, Steven Rousso-Schindler, Jamila Smith-Loud, and Lauren Wilcox. 2024. How knowledge workers think generative AI will (not) transform their industries. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 641, 26 pages. doi:10.1145/3613904.3642700