

Anticipating Risks and Identifying Governance Measures for the Use of genAI and FPT: Citizens' Perspectives Across Six Countries

CHIARA ULLSTEIN*, Chair of Cyber Trust, Technical University of Munich, Germany

MICHEL HOHENDANNER*[†], Chair of Cyber Trust, Technical University of Munich, Germany

JENS GROSSKLAGS, Chair of Cyber Trust, Technical University of Munich, Germany

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 the workshop by presenting a participatory procedure to identify high-priority risks and where to focus governance efforts, from the perspective of citizens.

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

Reference Format:

Chiara Ullstein, Michel Hohendanner, and Jens Grossklags. 2025. Anticipating Risks and Identifying Governance Measures for the Use of genAI and FPT: Citizens' Perspectives Across Six Countries. In *Proceedings of Fourth European Workshop on Algorithmic Fairness (EWAf'25)*. Proceedings of Machine Learning Research, 7 pages.

1 Introduction

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.

*Both authors contributed equally to this research.

[†]Michel Hohendanner was also affiliated with the Munich Center for Digital Sciences and AI, Munich University of Applied Sciences, at the time of project planning and data collection as well as with the German Institute for Japanese Studies (DIJ) Tokyo at the time of data collection and analysis.

Authors' Contact Information: Chiara Ullstein, Chair of Cyber Trust, Technical University of Munich, Munich, Germany, chiara.ullstein@tum.de; Michel Hohendanner, Chair of Cyber Trust, Technical University of Munich, Munich, Germany, michel.hohendanner@tum.de; Jens Grossklags, Chair of Cyber Trust, Technical University of Munich, Munich, Germany, jens.grossklags@in.tum.de.

This paper is published under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International (CC-BY-NC-ND 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

EWAf'25, June 30–July 02, 2025, Eindhoven, NL

© 2025 Copyright held by the owner/author(s).

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 [31]. We presented this work at EWAF’24 and suggested dialogue formats for in-depth discussions between citizens [30]. Our prior survey study [31] informed the current research project on several accounts: i.a., to narrow down the selection of countries across continents; to select different application contexts of interest to policymakers and advisors and to survey participants that were presented to participants of the dialogues for discussion; and to address country-specifics when applicable. Building on these results, we initiated the *Global AI Dialogues*, a dialogue series 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, 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 the workshop by presenting a method for **anticipating high-priority risks**, from the citizens’ perspective. We describe how we **designed the dialogues**. We show how our methodological approach leads participants to **identify where to focus governance efforts**. Further, 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 relevant aspects of the project. For further details on the methodological approach and a summary of first results on genAI we refer to our paper ‘Initiating the Global AI Dialogues: Laypeople Perspectives on the Future Role of genAI in Society from Nigeria, Germany, and Japan’ presented at CHI’25 [18].

2 Methods

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 (based on gender, age, and AI knowledge), creating mini-publics [8, 22, 27]. Participants indicated whether they would like to participate in the dialogue on genAI or FPT. 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, 33]. 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 [32], eliciting

socio-technical 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,¹ 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. We describe the citizen dialogue workshop process in detail in our CHI'25 publication [18].

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 multiple analysis workshop sessions. For the quantitative survey data, we applied frequency analysis, analyses of variances or Kruskal-Wallis tests, and pairwise comparisons.

3 Results

In the following, we present results from our first analyses. The results in Section 3.1 refer to the surveys from all collected countries. The analysis in Section 3.2 refers to all countries but Mexico. The analysis in Section 3.3 refers to Nigeria, Japan, and Germany. This is because we collected the data at different points in time in the second half of 2024 and the analysis timelines for different sets of countries differ from each other.

3.1 Need for Regulation and Responsible Institutions

After the group discussions, in their individual survey reflection, participants across all 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. More specifically, 93% (genAI) and 97% (FPT) of participants believed regulation is required. Thereof, 34% (genAI) and 31% (FPT) of participants indicated the need for legally binding regulatory measures for genAI and FPT, respectively. And 59% (genAI) and 65% (FPT) indicated the need for a combination of voluntary measures and legally binding regulation to ensure a desirable future with genAI or FPT can be achieved.

When asked which institutions they would trust the most in establishing measures that make the use of the discussed AI technology safe, across all countries and contexts, participants rated an international AI safety institute (genAI: 64%; FPT: 53%) as the most trustworthy institution. Companies developing AI were considered as second most trustworthy (genAI: 45%; FPT: 46%), however, this perception was, in particular, driven by participants from Nigeria (genAI: 73%; FPT: 82%). The government, an independent regulator, and intergovernmental organizations ranked third to fifth most trustworthy.

¹genAI: text generation using *DuckDuckGo AI Chat*: <https://duckduckgo.com/aichat>; image generation using *Deep Dream Generator*: <https://deepdreamgenerator.com/>

FPT: face detection and verification using *KI menschlich erklärt* (en: AI humanly explained): <https://ki-menschlich.netlify.app/gesichter-erkennen/>; face classification using the interactive video tutorial *How normal am I*: <https://www.hownormalami.eu/>

3.2 Facial Processing Technologies

We find that, overall, ideas of FPT leading to increased public safety and increased convenience were dominant despite the participants from Nigeria, Japan, India, Germany, and Bolivia acknowledging limitations and risks. Participants from India and Germany were more critical of FPT than participants from the other countries. 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.

3.3 Generative AI

Despite differences in experiences, exposure, and media coverage, citizens from Nigeria, Japan, and Germany 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 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

We contribute to the workshop by presenting our participatory approach and empirical data from six countries around the world on citizens' perspectives for a desirable future across two sociotechnical scenarios: the application or non-application of genAI and FPT.

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. While citizens defined such requirements for genAI and FPT and attached great importance to regulation, risks remain for which technological solutions and regulation alone are no silver bullet. With our work, we contribute to the debates on AI policy by providing recommendations derived from participants' identified requirements and suggested measures for AI to create value and foster a socially 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 Indian team, Cynthia Maria Montaudon-Tomas, Gabriela Sánchez Bazán, and Christin Schüler from the Mexican team, and Naira Paola Arnez-Jordan, Jose Jesus Cabrera-Pantoja, Indiana Karina Jordan-Barros from the Bolivian team. We thank all facilitators for supporting all dialogues. We are also grateful for the experts who have provided feedback, amongst others Emilia Gómez (Joint Research Center, EC), Orestis Papakyriakopoulos (TUM), Benedikt Zönnchen (MUCDAI/HM), and Téó Sanchez (LMU). This extended abstract has previously been presented at the CHI'25 workshops on *Sociotechnical AI Governance: Opportunities and Challenges for HCI* and on *Emerging Practices in Participatory AI Design in Public Sector Innovation*. We thank all participants for their feedback.

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. <https://doi.org/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. https://doi.org/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. <https://doi.org/10.1016/j.futures.2024.103346>
- [4] Anthony Dunne and Fiona Raby. 2013. *Speculative Everything: Design, Fiction, and Social Dreaming*. MIT Press, Cambridge, MA, USA. <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. <https://doi.org/10.1145/3532106.3534570>
- [6] Charles Ess. 2017. Digital media ethics. In *Oxford Research Encyclopedia of Communication*. Oxford University Press, Oxford, UK. <https://doi.org/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 (PDC '22)*. Association for Computing Machinery, New York, NY, USA, 147–154. <https://doi.org/10.1145/3537797.3537826>
- [8] Archon Fung. 2006. Varieties of participation in complex governance. *Public Administration Review* 66, s1 (2006), 66–75. <https://doi.org/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, USA. <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. <https://doi.org/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 (CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article 593, 15 pages. <https://doi.org/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. <https://doi.org/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 (GoodIT '23)*. Association for Computing Machinery, New York, NY, USA, 351–362. <https://doi.org/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. <https://doi.org/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, 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 (CHI '25)*. Association for Computing Machinery, New York, NY, USA. <https://doi.org/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. <https://doi.org/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. <https://doi.org/10.1017/dsi.2019.168>
- [22] OECD. 2020. *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave*. OECD Publications, Paris, France. <https://doi.org/10.1787/339306da-en>
- [23] John Robinson. 1982. Energy backcasting: A proposed method of policy analysis. *Energy Policy* 10, 4 (1982), 337–344. [https://doi.org/10.1016/0301-4215\(82\)90048-9](https://doi.org/10.1016/0301-4215(82)90048-9)
- [24] John Robinson. 2003. Future subjunctive: Backcasting as social learning. *Futures* 35, 8 (2003), 839–856. [https://doi.org/10.1016/S0016-3287\(03\)00039-9](https://doi.org/10.1016/S0016-3287(03)00039-9)
- [25] Daniela K. Rosner, Saba Kawas, 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 (CSCW '16)*. Association for Computing Machinery, New York, NY, USA, 1131–1141. <https://doi.org/10.1145/2818048.2820021>
- [26] Johnny Saldaña. 2013. *The Coding Manual for Qualitative Researchers* (2nd ed.). SAGE, Los Angeles, CA, USA.
- [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. <https://doi.org/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. <https://doi.org/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, Mamare 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 (PDC '22)*. Association for Computing Machinery, New York, NY, USA, 120–132. <https://doi.org/10.1145/3536169.3537786>
- [30] Chiara Ullstein, Michel Hohendanner, and Jens Grossklags. 2024. Mapping policymakers’ and laypeople’s perceptions of genAI and FPT in the context of the EU AI Act. In *Proceedings of the Third European Workshop on Algorithmic Fairness (EWAF '24)*. CEUR Workshop Proceedings, Aachen, Germany. https://ceur-ws.org/Vol-3908/paper_22.pdf

- [31] 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 (COMPASS '25)*. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3715335.3735470>
- [32] 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 (CSCW Companion '24)*. Association for Computing Machinery, New York, NY, USA, 429–436. <https://doi.org/10.1145/3678884.3681887>
- [33] 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 (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 641, 26 pages. <https://doi.org/10.1145/3613904.3642700>