Global AI Cultures Workshop ICLR 2024

Rida Qadri, Fernando Diaz, Arjun Subramonian, Sunipa Dev, Georgina Born, Mary L. Gray, Jessica Quaye, Rachel Bergmann.

Workshop summary

Problem Statement

Building globally-inclusive generative artificial intelligence (genAI) that encodes, respects, and valorizes cultural sensibilities as well as performs well for users across cultural contexts, is an important goal as we deploy generative AI products globally. However, existing evaluation, design and deployment practices are not oriented towards a diversity of global cultures, betraying these aims. Emerging research already shows that within the current AI development pipeline only certain western cultures get centered, distributed and amplified [1-5]. There is currently limited conversation within AI/ML communities on how to measure complex variables like cultural representation and social impact of emerging technologies. We also do not fully understand the cultural priorities and needs of genAI users and of the human labor that powers genAI. If this relationship between AI and global cultures is not examined we inadvertently could be universalizing western centered AI and create unforeseen impacts on global cultural production, values and consumption.

To build inclusive AI for people, we must thus focus on how AI pipelines can be more globally inclusive, the cultural values and assumptions we center in our development processes and the cultural impact of emerging AI technologies. This understanding becomes urgently important as new generative AI technologies are increasingly used globally to consume and produce cultural artifacts like art, stories, and music, raising questions of cultural quality, inclusion, and impact.

Such a research agenda requires a cross-disciplinary effort between AI researchers considering the technical nuances of generative AI with scholars from the humanities and social sciences that have long thought about the social and cultural impacts of new technologies. We thus urgently need a cross-disciplinary and cross-community framework for understanding this multifaceted relationship between AI and Culture.

Aims and Importance

This workshop aims to begin a conversation between core AI researchers and experts from the social sciences and humanities to collectively develop a research agenda around the impact of genAI on cultures and the impact of cultural values on genAI. Through this focus, the workshop will encourage field building on deepening our understanding for how we can build and deploy globally inclusive genAI and how we can responsibly encode cultural knowledge into our technologies. It will include discussions about:

1. *Conceptual and Theoretical Foundations for Cultural Inclusion in GenAI:* What does global inclusion mean in the context of generative AI and what are the possibilities and challenges of building culturally inclusive generative models?
2. *Scalable Cultural Representation Evaluations:* How do we build evaluation and development pipelines that can test cross-cultural performance via cultural metrics such as representation, quality, impact, and inclusion at scale?
3. *Culturally-Rich Training Datasets:* What are the features of a culturally representative training dataset and what processes and conditions are needed in order to create or curate such training data?
4. *Methods to study cultural values of generativeAI:* How can we recognize and account for the different cultural values that are embedded in our AI pipelines? How do we bring our cultures of development in sync with our cultures of deployment?
5. *User Interactions with GenAI in Support of Cultural Inclusion:* Are there creative strategies that can reparatively promote the inclusion of subjugated cultural values through UI, deployment, or public education/advocacy?
6. *Cultural impacts of Generative AI:*How can we understand immediate and longer-term impacts of these technologies on the culture industries? How does AI support or challenge support existing dynamics in the culture industries? Are there existing norms or principles in non-AI systems of content creation and distribution?

Although the ICLR community plays a vital role in AI development as it pertains to culture (e.g., advancing generative media models and algorithms), so far it has supported only limited engagements with researchers from outside of the machine learning community. Experts from the humanities and social sciences bring key perspectives and methodologies that deepen our understanding of the relationship between AI and society, as well as conceptualize more creative mitigations to the complex socio-cultural questions brought up by this interaction. Given the preeminence of ICLR within the AI field, we believe it is the ideal venue to host these multi-disciplinary conversations.

The machine learning and AI research community has long supported initiatives to understand the broader social context of its work. This includes workshops such as FATML, Resistance AI, and Participatory Approaches to Machine Learning. It also includes mechanisms like broader impact statements and paper checklists. We see this workshop proposal as a continuation of this line of critical reflective inquiry, but specifically focused on developing a cross-disciplinary space for how global cultural values can shape our understanding of fairness and expand our fairness frameworks to think about socio-cultural impacts of our rapidly advancing AI technologies.

Modality (in-person, virtual, hybrid)

To ensure inclusivity in access for participants from lower-resourced contexts or disciplines for whom ICLR is not a primary conference, we will host a hybrid workshop. We will host the virtual portion of the workshop on Zoom and use Discord for conversation/community building.

Tentative schedule

This will be a full day *hybrid* workshop focused on generative conversation and discussion across disciplinary and cultural perspectives to work towards field building at the intersection of Ai and

Culture. We want to create a space for interdisciplinary collaboration in order to develop shared understandings across fields. We will achieve this through a mix of invited multi-disciplinary panels, poster sessions, and teach-ins where paper presenters teach a key concept from their field that they think is important for a multi- disciplinary audience as a basis for understanding the interactions between AI and culture. The organizers will solicit and review short position papers on important problems on the intersection of AI and Culture. These position papers will help us select speakers for the workshop and also be turned into posters. The workshop will close with a cultural AI evaluation exercise to give people a hands-on demonstration of ways in which AI can shape culture and cultural identities can shape AI.

09:00-09:15	Opening community building activity
9:15-10:30	Shared Concepts in AI & Culture *Each paper author shares a problem or concept in their field they consider vital for understanding the interactions between AI and culture, teaching this to all those present*
10:30 - 11:00	Coffee Break & poster session
11:00-12:00	Invited Panel: Cultural Inclusion in Generative AI: Possibilities and Tensions *Doug Eck (Google), Lauren Wilcox (Ebay), Arvind Narayanan (Princeton), Kalika Bali (MSR)*
12:00 -1:00	Lunch
1:00 - 2:15	*Invited Panel* *Cultural Impact of Generative AI* *Fuchsia Hart (Curator at V&A), Huma Gupta (MIT, Architecture) , Mohreshin Allahyari (Artist), Maya Indira Ganesh (University of Cambridge)*
2:15 - 2: 45	Coffee Break & poster session
2:45-3:45	Cultural Evaluation Exercise
3:45-4:00 pm	Closing

**confirmed*

Anticipated audience size
100, including in-person and virtual

Plan to get an audience
We will advertise on social media channels like Twitter, Bluesky Mastodon, LinkedIn. We will also reach out through relevant mailing lists including respective organizational ones that each organizer is a part of.

Organizers and biographies

Fernando Diaz is an Associate Professor in the Language Technologies Institute (LTI) at Carnegie Mellon University and a research scientist at Google Research. Fernando's background is in the design and evaluation of search engines and recommender systems. His current research covers three themes: quantitatively evaluating AI systems, retrieval-enhanced AI, and

understanding the cultural impact of AI in domains like music and literature through interdisciplinary collaborations. Having held both individual contributor and senior leadership positions in industry, Fernando is particularly interested in the deployment of real AI systems. Previously, Fernando was the assistant managing director of Microsoft Research Montréal, where he also led FATE Montréal, and director of research at Spotify, where he built its research organization on recommendation, search, and personalization. Fernando's work has received special recognition and awards at SIGIR, CIKM, CSCW, WSDM, ISCRAM, and ECIR. He is the recipient of the 2017 British Computer Society Karen Spärck Jones Award and holds a CIFAR AI Chair. Fernando has co-organized several NIST TREC tracks, WSDM (2013), Strategic Workshop on Information Retrieval (2018), FAccT (2019), SIGIR (2021), and the CIFAR Workshop on Artificial Intelligence and the Curation of Culture (2019). He received his BS in Computer Science from the University of Michigan Ann Arbor and his MS and PhD from the University of Massachusetts Amherst.

Rida Qadri is a research scientist at Google Research. Her research interrogates the cultural assumptions underpinning the design and deployment of generative AI systems. She specifically focuses on the harms produced by culturally inappropriate AI design choices and documents how communities resist and repair these technologies. She has co-organized workshops at the intersection of AI and Culture at NeurIPS 2022, CHI 2021 and CVPR 2023. She has published and presented her work at conferences like AIES, FAccT, CSCW, SASE and CHI. She has a PhD in Computational Urban Studies from the Massachusetts Institute of Technology.

Arjun Subramonian (they/அவங்க) is a Computer Science PhD student, Cota-Robles fellow, and NSF MENTOR fellow at the University of California, Los Angeles, and a core organizer of Queer in AI. They work on critical approaches to fairness in graph learning that bridge intersectionality, mathematical theory, and queer perspectives. Their research focuses on unveiling and critiquing value-laden design choices made while modeling graph data, towards minimizing harms faced by marginalized communities. They have given guest lectures and talks on bias, discrimination, and power in graph learning and NLP, as well as grassroots activism in AI, at various universities including USC, DCU, and Umeå. They have organized numerous Queer in AI workshops, and were/are the affinity workshops chair for NeurIPS 2022/2023.

Sunipa Dev (she/her), is a Research Scientist at Google Research working towards fair, inclusive, and socio-culturally aware NLP. Her research centers around inclusion of global perspectives in different pipelines in NLP, particularly in model evaluations to better understand and mitigate potential risks and harms. Prior to this, she was an NSF Computing Innovation Fellow at UCLA, before which she was awarded her PhD at the School of Computing at the University of Utah. She has taught guest lectures and given talks centered on inclusive NLP at multiple places including University of Utah (2023), University of Southern California (2023), University of Bocconi (2021), and a keynote at TrustNLP Workshop (ACL 2023). She is currently a program chair for WINLP (organizing across different NLP venues including NAACL, ACL, and EMNLP), and was the

affinity workshop chair at NeurIPS 2022 and a workflow chair for AAAI 2022. She has also co-organized tutorials and workshops at various venues including KDD 2021, NeurIPS 2022, EACL 2023, and FAccT 2023.

Jessica Quaye is a Computer Science PhD student in the EDGE Computing Lab at Harvard University. Prior to joining Harvard, Jessica graduated from MIT with the highest awards for leadership and academic excellence in Electrical Engineering and Computer Science. She also spent a year at Tsinghua University as a Schwarzman Scholar drawing lessons from China's economic rise for developing countries. Her research interests are in building machine learning systems that work effectively in resource-constrained contexts for developing countries.

Georgina Born is Professor of Anthropology and Music at University College London. Previously she held Professorships at the Universities of Oxford (2010-21) and Cambridge (2006-10). She also had a professional life as a musician in experimental rock, jazz and free improvisation. Her work combines ethnographic and theoretical writings on music, sound, television and digital media. Books include *Rationalizing Culture: IRCAM, Boulez, and the Institutionalization of the Musical Avant-Garde* (1995), *Western Music and Its Others* (ed. with D. Hesmondhalgh, 2000), *Uncertain Vision* (2004), *Music, Sound and Space* (ed., 2013), *Interdisciplinarity* (ed. with A. Barry, 2013), *Improvisation and Social Aesthetics* (ed. with E. Lewis and W. Straw, 2017), and *Music and Digital Media: A Planetary Anthropology* (ed., 2022). She directed the ERC-funded research program 'Music, Digitization, Mediation' (2010-15) and in 2021 was awarded an ERC grant for 'Music and Artificial Intelligence: Building Critical Interdisciplinary Studies'. She has held visiting professorships at UC Berkeley, UC Irvine and Aarhus, Oslo, McGill and Princeton Universities.

Rachel Bergmann is a PhD student in Communication at Stanford University and Research Affiliate at the Centre for Media, Technology, and Democracy at McGill University. Her work uses historical and interpretive methods to analyze information technologies and their cultural contexts. She has published her research in academic journals like *International Journal of Communication* and written articles for wider audiences for AI Now and Real Life Magazine. Previously, she worked as the Research Assistant at the Social Media Collective at Microsoft Research New England and was a Visiting Predoctoral Fellow at the Max Planck Institute for the History of Science. She has organized panels and presented her work at Data & Society, International Communication Association (ICA), Association of Internet Researchers (AoIR), Society for Social Studies of Science (4S), CHI, and CSCW.

Mary L. Gray is Senior Principal Researcher at Microsoft Research and Faculty Associate at Harvard University's Berkman Klein Center for Internet and Society. She maintains a faculty position in the Luddy School of Informatics, Computing, and Engineering with affiliations in Anthropology and Gender Studies at Indiana University. Mary, an anthropologist and media scholar by training, focuses on how people's everyday uses of technologies transform labor, identity, and human rights. Her books include *Out in the Country: Youth, Media, and Queer Visibility in Rural America* and *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. She has served on several boards, including *Public Responsibility in Medicine and Research* and the California Governor's Council of Economic Advisors, in addition to chairing

the Microsoft Research Ethics Review Program—the only federally registered institutional review board of its kind in the tech industry. She is the current Board President of the Data Nutrition Project. In 2020, Mary was named a MacArthur Fellow for her contributions to anthropology and the study of technology, digital economies, and society.

Commitment to Diversity and Inclusion

Our workshop is committed to diversity of disciplines and identities. The question of culture itself stimulates discussion on the western-centric biases and values being embedded in and perpetuated by AI-driven tech. Thus, the topic compels us to create a space where different cultural, identity and disciplinary perspectives are brought in, a task our culturally and disciplinarily diverse organizing team is committed to. In a bid to bring in voices that are often not present in ICLR, our speakers do a mix of research in a diversity of geographical contexts and represent disciplines such as core AI/ML, HCI, Urban studies, Communications, Anthropology and Sociology. Our organizing team has also been involved in organizing Widening NLP workshops at *CL conferences and Queer in AI workshops at AI/ML/NLP conferences, to bring in diverse, underrepresented voices into high-impact AI forums. Our organizing team has also held diversity and inclusion leadership roles at NAACL and NeurIPS. The speakers and organizers are also committed to research on justice, fairness and critical perspectives. We have also taken care for our invited speakers to represent a mix of gender and ethnic identities.

Access

We want to make sure this conversation within the workshop influences the direction of our AI field. All panels will be recorded and shared on the website. Position papers will be non-archival and shared on the website.

Previous related workshops

This is the second iteration of the schedule AI in Culture held at NeurIPS 2022 which successfully brought together social scientists, critical humanities scholars, and AI experts to sustain multi-disciplinary conversation. All the organizers have extensive experience introducing and leading discussions on cultural considerations in AI pipelines. Some other venues where we have co-organized and conducted tutorials and workshops with similar goals include FAccT 2023 (Tutorial on Cross Cultural Considerations in AI; 50 attendees), EACL 2023 (Cross Cultural Considerations In NLP Workshop; 75 attendees), CVPR 2023 (Ethical Considerations in Creative Applications of Computer Vision), ICML 2021/NeurIPS 2021/ICML 2022/ACL 2023 (Queer in AI; 30-50 attendees).

References

1. Sunipa Dev, Akshita Jha, Jaya Goyal, Dinesh Tewari, Shachi Dave, and Vinodkumar Prabhakaran. 2023. Building Stereotype Repositories with Complementary Approaches for Scale and Depth. In Proceedings of the First Workshop on Cross-Cultural

Considerations in NLP (C3NLP), pages 84–90, Dubrovnik, Croatia. Association for Computational Linguistics.

2. Abhipsa Basu, R. Venkatesh Babu, Danish Pruthi, "Inspecting the Geographical Representativeness of Images from Text-to-Image Models," 2023.
3. Yong Cao, Li Zhou, Seolhwa Lee, Laura Cabello, Min Chen, & Daniel Hershcovich. (2023). Assessing Cross-Cultural Alignment between ChatGPT and Human Societies: An Empirical Study.
4. Ramya Srinivasan and Kanji Uchino. 2021. Biases in Generative Art: A Causal Look from the Lens of Art History. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 41–51. https://doi.org/10.1145/3442188.3445869
5. Rida Qadri, Renee Shelby, Cynthia L. Bennett, and Emily Denton. 2023. AI's Regimes of Representation: A Community-centered Study of Text-to-Image Models in South Asia. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23). Association for Computing Machinery, New York, NY, USA, 506–517. https://doi.org/10.1145/3593013.3594016