

# Towards a Multidisciplinary Vision for Culturally Inclusive Generative AI

Asia Biega<sup>\*1</sup>, Georgina Born<sup>\*2</sup>, Fernando Diaz<sup>\*3</sup>, Mary L. Gray<sup>\*4</sup>,  
and Rida Qadri<sup>\*5</sup>

1 MPI-SP – Bochum, DE. [asia.biega@acm.org](mailto:asia.biega@acm.org)

2 University College London, GB. [g.born@ucl.ac.uk](mailto:g.born@ucl.ac.uk)

3 Carnegie Mellon University – Pittsburgh, US. [diazf@acm.org](mailto:diazf@acm.org)

4 Microsoft New England R&D Center – Cambridge, US. [mlg@microsoft.com](mailto:mlg@microsoft.com)

5 Google – San Francisco, US. [ridaqadri@google.com](mailto:ridaqadri@google.com)

---

## Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 25022 “Towards a Multidisciplinary Vision for Culturally Inclusive Generative AI”. The gathering focused on questions raised by the rapid deployment of Generative AI systems and their integration into global systems of cultural communication, consumption, and production. As these technologies shape our cultures, we urgently need conceptual foundations for investigating the cultural inclusivity of generative AI pipelines (from data collection, to model development and deployment, to evaluation), as well as methods to study the varying societal and cultural impacts of generative AI.

This Dagstuhl Seminar convened scholars and practitioners from computer science, social sciences, the tech industry, and creative industries to discuss the cultural implications of generative AI and find paths toward building generative AI that can be responsive to the diverse needs of individuals, groups, and societies around the world. Together, seminar participants began the challenging but necessary work of building shared language and frameworks for reshaping the technical and social architectures of generative AI.

The seminar was structured along three main dimensions for interdisciplinary discussions:

- Examining the cultural values being currently centered in generative AI.
- Studying the possibilities and risks of encoding cultural knowledge into generative AI technologies.
- Understanding the cultural impact of these technologies.

We succeeded in building an expert network committed to understanding and designing a culturally-attuned generative AI and to lay the foundation for an interdisciplinary research and practice agenda on global inclusion and generative AI.

**Seminar** January 6–9, 2025 – <https://www.dagstuhl.de/25022>

**2012 ACM Subject Classification** Applied computing; Computing methodologies; Human-centered computing; Information systems

**Keywords and phrases** creativity, cultural inclusion, generative artificial intelligence, global south, social impact of ai

**Digital Object Identifier** 10.4230/DagRep.15.1.33

---

\* Editor / Organizer



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Towards a Multidisciplinary Vision for Culturally Inclusive Generative AI, *Dagstuhl Reports*, Vol. 15, Issue 1, pp. 33–49

Editors: Asia Biega, Georgina Born, Fernando Diaz, Mary L. Gray, and Rida Qadri



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

## 1 Executive Summary


*Rida Qadri (Google – San Francisco, US)*

*Asia Biega (MPI-SP – Bochum, DE)*

*Georgina Born (University College London, GB)*

*Fernando Diaz (Carnegie Mellon University – Pittsburgh, US)*

*Mary L. Gray (Microsoft New England R&D Center – Cambridge, US)*

License  Creative Commons BY 4.0 International license

© Rida Qadri, Asia Biega, Georgina Born, Fernando Diaz, and Mary L. Gray

## Motivation

Recent years have seen rapid development and widespread adoption of generative AI systems that algorithmically model human creativity and decision-making. In particular, this technological shift has profound implications for how cultural artifacts like music, news, literature, and film are produced and consumed, raising concerns about the potential cultural implications of this technology. At the same time, these technologies are displaying Western-centrism in AI training and evaluation data, definitions of "success", and evaluation methods. As a result, generative AI technologies, while arguably improving their reliability for basic output of sensible prose and images, have a recognizable pattern of failing to generate norms and values representative and inclusive of non-Western perspectives. For example, recent research and media reports have found that models are less competent at generating culturally significant material outside of a Western point of view, frequently omitting non-Western cultural knowledge from outputs, and perpetuating Western stereotypes in generated output. Addressing these failures and their broader impact is crucial to prevent globally-launched generative AI tools from becoming vehicles for reinforcing Western-centric cultural norms and values, production and distribution methods, and in these ways further exacerbating global inequities.

The urgent need for a seminar on these topics was highlighted in the first-of-its-kind 2022 NeurIPS workshop on "AI and Culture" that brought together researchers from computer science, the humanities, and the social sciences at the premier conference of AI researchers and practitioners. At this workshop, emergent conversations pointed to how building culturally sensitive, responsive, and accountable AI systems will require researchers and engineers to include diverse disciplinary voices, community expertise, and cultural knowledge in AI research and development. Such efforts to recognize and incorporate myriad cultural contexts into AI systems are often siloed within disciplines and, as a result, are disjointed and limited in their impact on technological design. In particular, there are no cohesive frameworks to help researchers fold nuanced cultural analyses and situated knowledge into generative AI models. There is therefore a critical, currently unmet need to break down disciplinary silos and create coherent interdisciplinary conceptual foundations for novel, culturally-sensitive generative AI research and practices. The most promising areas in need of interdisciplinary collaborative research include: 1) new approaches to data collection; 2) interdisciplinary frameworks and methods for model development and deployment; and 3) new techniques that integrate and distinguish the value of qualitative and computational approaches to evaluation. We also see the need to develop new interdisciplinary methods, crossing between qualitative and quantitative approaches, to study the societal impacts of generative AI. As generative AI research is currently confined mainly to industry-academic collaborations, we further aim to broaden the contextual and institutional perspectives brought to these challenges beyond academia to include voices from civil society and impacted communities – and this was also a goal achieved in the seminar.

## Program

The seminar lasted 2.5 days. As our goal was to create an interdisciplinary space for discussion, we had 28 participants with backgrounds in multiple disciplines and sectors. Participants included experts in computer science, data science, machine learning (ML), information retrieval (IR), natural language processing (NLP), human-computer interaction (HCI), responsible artificial intelligence (AI), social computing, critical data studies, music and ethnomusicology, anthropology, history, political philosophy, science and technology studies (STS), media studies, communication, and architecture. The seminar also included contributions from filmmakers and the creative industry. The participant pool reflected the broad spectrum of perspectives and expertise on language, culture, and cultural production, necessary to advance the dialogue on AI and culture.

To encourage participants to come to the seminar prepared with preliminary reflections on the topic, we asked them to complete a round of preparatory work two months before the seminar. This consisted of sharing a paper that participants had written or found fruitful in their current work in order to introduce themselves to the rest of the group and explain their way of approaching questions of AI and culture. We also asked each participant to reflect on a series of questions: 1) How are you thinking about the term “culture” in the context of artificial intelligence? 2) What is a provocation or critical question you would like to share regarding the intersection of AI and culture? And, 3) Where would you like to see the field of AI and culture head next?

The first day of the seminar was dedicated to sparking discussion and allowing participants to get to know each other’s perspectives on the seminar topic. Recognizing that most of the invited scholars do not regularly cross paths at a single-disciplinary home conference, the first exercise of the day was a series of “speed dating” rounds. Participants rotated through ten-minute introductory conversations with at least three other participants. We asked participants to share basic information about their disciplinary training and home institutions. Then they added background on what they hoped to gain in terms of a deeper understanding of the interdisciplinary challenges and opportunities in the emerging field of generative AI and cultural diversity. During Round 2, participants shared the next project they were working on or their dream project in this space. In Round 3, they discussed examples of AI failures that illustrated their thinking on the seminar topic. These discussions helped to identify the first key areas for the advancement of the field.

Once participants had a sense of the breadth and depth of expertise in the room, we shifted to the first substantive programming component. This took the form of three panels each with three speakers, with each speaker offering 5-minute “firestarter” provocations, followed by an open group discussion. The firestarter presentations were followed by a short, individual reflective writing session, where participants could document their questions and reactions to the discussions, and contribute to a collective note-taking document. The nine speakers were selected to give firestarter talks on the basis of the submitted preparatory work. The organizers conducted thematic coding of the received documents ahead of the seminar in order to assemble the panels, with the aim of putting participants into multi- and interdisciplinary dialogue early on in the seminar.

On the second day of the seminar, participants were invited to collectively come up with themes they wanted to discuss further. They then broke into small group discussions on the chosen themes. The small group discussions were followed by shareout sessions, followed by the generation of provocations, and the genesis of potential future collaborative projects among participants. As noted above, we used the themes and points of friction from Day 1’s Firestarter discussions and individual reflections to brainstorm and then thematically code

the questions and clusters of discussion that had the most consensus. After spending the morning consolidating themes, we identified and converged on three clusters for small group discussion. The themes were articulated as: Discussion Cluster 1: Power, Future, History; Discussion Cluster 2: Interdisciplinarity in Computer Science Cultures; and Discussion Cluster 3: Culture Encodability. Each Discussion Cluster is described below.

On the third day of the seminar, participants discussed the next steps in small groups.

## Outcomes / Planned outcomes

This seminar fostered a critical reflection on the development of culturally inclusive AI, highlighting the rare opportunity for interdisciplinary learning, and generating a profound sense of urgency and clarity regarding the challenges. Participants formed working groups around three key outcomes: an agenda-setting document, a “meta-metadata” project, and a research project on integrating qualitative and quantitative evaluation methods.

**Agenda-Setting Project** This initiative aims to establish a shared, nuanced understanding of AI, culture, and technology, moving beyond simplistic definitions. The group will produce a document for funders like the NSF, articulating the challenges and relevance of culturally inclusive AI. This document will influence funding priorities and foster interdisciplinary research, serving as a foundation for broader engagement with policymakers and the public.

**Meta-Metadata Project** This research project focuses on developing and implementing new approaches to metadata creation and management, fostering culturally rich datasets through open-source, collaborative models. Key planned outcomes include a course and hackathon exploring nuanced metadata encoding, and the creation of a network of scholars dedicated to this work. The project also explores leveraging existing platforms like Wikimedia to host and manage detailed, culturally diverse metadata, addressing challenges like image metadata and incentivizing scholarly participation.

**Project on Integrating Qualitative and Quantitative Methods for AI Evaluations** This future project addresses the critical need for robust methodologies that integrate qualitative and quantitative data in AI evaluation. It aims to develop frameworks that translate qualitative insights into concrete algorithmic interventions without losing critical nuances. The research will explore methods like “fictions” or imagined scenarios to anticipate potential consequences, and guide development, moving beyond the limitations of current practices that rely on small user groups or subjective judgments.

Beyond these tangible projects, the seminar achieved a significant shift in perspectives. Computer scientists gained a deeper appreciation for the complexities of culture, while social scientists and humanities scholars refined their critiques through a clearer understanding of AI’s potential. This cross-disciplinary dialogue led to a richer understanding of the multi-layered relationship between AI and culture, moving beyond simplified encodings and benchmarks. Participants also valued the seminar’s global representation, moving beyond US/EU-centric viewpoints. The seminar generated significant momentum, energizing participants and sparking new collaborative research directions. As one computer scientist noted, “The questions I came in with are very different from the questions I’m leaving with... I find that the questions I leave with are much richer – and harder.” Similarly, a social scientist expressed, “As a non-technical person the seminar was incredibly insightful to better understand what the state of the art currently is, what the possibilities and limitations for culturally sensitive interventions in these systems may be.”

Participants consistently highlighted the need for a second iteration of the seminar, emphasizing the value of continued multidisciplinary spaces for collaboration. They left with a richer set of concerns and vocabularies, anticipating that this assemblage would transform individual disciplinary research and lead to numerous joint collaborations. The seminar was described as “creatively fortifying and vitalizing,” creating meaningful connections and inspiring participants to push forward in the pursuit of culturally inclusive AI.

## 2 Table of Contents

### Executive Summary

*Rida Qadri, Asia Biega, Georgina Born, Fernando Diaz, and Mary L. Gray . . . .* 34

### Overview of Talks

Firestarters: Initial Areas for Exploration (25022)

*Rida Qadri, Asia Biega, Georgina Born, Fernando Diaz, and Mary L. Gray . . . .* 39

Panel Discussion 1: Definitions of Culture (25022)

*Rida Qadri, Hal Daumé III, Tarleton Gillespie, and Molly Steenson . . . . .* 40

Panel Discussion 2: Encoding Culture (25022)

*Rida Qadri, Mary L. Gray, Huma Gupta, Emanuel Moss, and Alice Oh . . . . .* 40

Panel Discussion 3: Institutional Reflections and Collaborations (25022)

*Rida Qadri, Naveen Bagalkot, Catherine d'Ignazio, and Sara Hooker . . . . .* 42

### Working groups

Working Group 1: Power, Future, History (25022)

*Rida Qadri, Virgilio Almeida, Naveen Bagalkot, Georgina Born, Anita Say Chan, Hal Daumé III, Catherine d'Ignazio, Giovanna Fontenelle, Tarleton Gillespie, Darci Sprengel, Molly Steenson, and Harini Suresh . . . . .* 43

Working Group 2: Interdisciplinarity / CS cultures (25022)

*Rida Qadri, Asia Biega, Tobias Blanke, Marc Cheong, and Mary L. Gray . . . . .* 44

Working Group 3: Culture Encodability (25022)

*Rida Qadri, Kalika Bali, Beth Coleman, Fernando Diaz, Huma Gupta, Sara Hooker, Maurice Jones, Emanuel Moss, Maryam Mustafa, Alice Oh, and Moira Weigel . . .* 45

### Open problems

Future directions based on participant feedback (25022)

*Rida Qadri, Asia Biega, Georgina Born, Fernando Diaz, and Mary L. Gray . . . .* 47

**Participants . . . . .** 49

### 3 Overview of Talks

#### 3.1 Firestarters: Initial Areas for Exploration (25022)

*Rida Qadri (Google – San Francisco, US), Asia Biega (MPI-SP – Bochum, DE), Georgina Born (University College London, GB), Fernando Diaz (Carnegie Mellon University – Pittsburgh, US), and Mary L. Gray (Microsoft New England R&D Center – Cambridge, US)*


**License** © Creative Commons BY 4.0 International license  
© Rida Qadri, Asia Biega, Georgina Born, Fernando Diaz, and Mary L. Gray

Three high-level themes emerged after we moved to individual reflection and a last round of open discussion, setting the stage for Day 2's Clustering exercise. Specifically, we ended our day noting the following key areas for exploration:

- The challenges of defining culture are multifaceted, involving different definitions and disciplinary lenses. There are significant gaps in what is being represented, and understanding what culture can achieve beyond Responsible AI ethics framings is crucial.
- Encoding culture within AI systems presents its own set of challenges, including the use of various computational methods to incorporate cultural knowledge and their potential consequences.
- The role of metadata is complex, and integrating both quantitative and qualitative perspectives is essential yet challenging.
- Institutional aspects and interdisciplinarity play a significant role in the cultures of AI production. There is a need for alternative imaginaries of technology that go beyond the corporate inclusion of data.
- Building collaborative teams that include diverse perspectives and experiences is vital, and fostering interdisciplinarity is key to advancing the field.
- The headwinds that work against multidisciplinary approaches to culturally-inclusive AI are exacerbated by the absence of regulatory frameworks and cultural norms that could foster synergies and accountability across academic and industry-based AI research and development settings.
- The challenges of definitions of culture (different definitions and disciplinary lenses, what is not being represented, what culture can get us beyond Responsible AI ethics framings).
- Encoding culture (different computational methods to include cultural knowledge in AI – and their consequences, the complex role of metadata, and challenges of integrating quantitative and qualitative perspectives).
- Institutional aspects and interdisciplinarity (cultures of AI production and alternative imaginaries of tech, AI beyond corporate inclusion of data, building collaborative teams, interdisciplinarity).

### 3.2 Panel Discussion 1: Definitions of Culture (25022)

*Rida Qadri (Google – San Francisco, US), Hal Daumé III (University of Maryland – College Park, US), Tarleton Gillespie (Microsoft New England R&D Center – Cambridge, US), and Molly Steenson (American Swedish Institute – Minneapolis, US & Carnegie Mellon University – Pittsburgh, US)*

**License**  Creative Commons BY 4.0 International license  
© Rida Qadri, Hal Daumé III, Tarleton Gillespie, and Molly Steenson

Panel Discussion 1 at the Dagstuhl Seminar focused on the theme of “Definitions of Culture” and was presented by Tarleton Gillespie, Hal Daume, and Molly Wright Steenson.

The session began with Tarleton Gillespie opening with a quote from Raymond Williams, highlighting the complexity of defining culture. Gillespie emphasized the importance of representation and the lived practices of stakeholders, including designers and users. He raised questions about how cultural values are inscribed in tools and the biases that emerge over time.

Hal Daume’s presentation focused on the gaps between community knowledge and computational knowledge. He discussed the challenges of measuring culture and the limitations of current AI systems in understanding diverse cultural contexts. Daume highlighted the mismatch between the knowledge of individuals and communities and the knowledge embedded in AI systems, using examples such as sign language and African American linguistic communities. He questioned whether computer science is open to expanding its understanding of culture beyond quantifiable metrics.

Molly Wright Steenson’s contribution focused on the cyclical nature of how industries manage crises related to ethics and safety. She reflected on the ethical crisis in Responsible AI (RAI) in 2018 and what lessons could be learned to rethink the framework. Steenson also discussed the importance of considering cultural practices and norms in the development of AI systems and the potential for cultural imposition by organizational structures. She emphasized the need for a hybrid methodology that integrates qualitative and quantitative approaches to better understand and model cultural norms.

The key takeaways from Firestarter 1 included the recognition of the complexities and tensions in defining and measuring culture within AI systems. The presenters highlighted the importance of expanding the understanding of culture in computer science, moving beyond mere quantification to include qualitative insights. They also highlighted the need for interdisciplinary collaboration and the inclusion of diverse cultural perspectives in AI development.

### 3.3 Panel Discussion 2: Encoding Culture (25022)

*Rida Qadri (Google – San Francisco, US), Mary L. Gray (Microsoft New England R&D Center – Cambridge, US), Huma Gupta (MIT – Cambridge, US), Emanuel Moss (Intel – Santa Clara, US), and Alice Oh (KAIST – Daejeon, KR)*

**License**  Creative Commons BY 4.0 International license  
© Rida Qadri, Mary L. Gray, Huma Gupta, Emanuel Moss, and Alice Oh

Panel Discussion 2 focused on the theme of “Encoding Culture” and was presented by Huma Gupta, Alice Oh, and Emanuel Moss.



**Huma Gupta**

Huma Gupta's presentation centered on the concept of the "Library of Missing Metadata," inspired by Mimi Onuoha's work. Gupta explored the idea of adding meta-metadata to support changing interpretations of artifacts and the complexities of digital architectures. She discussed the aggregation of terms and the legacies of taxonomies, suggesting that meta-metadata could introduce friction and complexity to visually prompt disruption of what counts as culture. Gupta also reflected on the challenges of encoding culture and the potential for cultural imposition by organizational structures.

**Alice Oh**

The session began with Alice Oh discussing her expertise in building large language models (LLMs) and the challenges of creating benchmarks for cultural competence. She emphasized the importance of considering a mix of well-represented and under-represented cultures and highlighted the difficulties in defining what should be included in these benchmarks. Oh also pointed out the presuppositions embedded in questions and the need for application scenarios to create effective LLMs.

Oh specifically discussed the BLEnD Dataset which represents their recent effort to evaluate the cultural commonsense knowledge of large language models (LLMs) across 13 languages and 16 regions. Native speakers collaboratively created a common set of questions, translated them into their languages, and gathered responses from other native speakers. The evaluation of the LLMs on this carefully crafted dataset highlighted serious limitations in LLMs: they struggle to perform well in understanding and representing languages and cultures outside of a few dominant ones.

However, the project also raised deeper questions about methodology and objectives. For example, what exactly is "culture," and how can we ensure that questions posed to annotators avoid embedding cultural presuppositions? Determining the "ground truth" for answers further complicates matters, as cultural identity is complex and multifaceted. A Korean annotator, for instance, might draw on their heritage, personal experiences, and exposure to other cultures, such as life in the U.S. or work in a global field like computer science. LLMs must be designed to navigate such complexities by recognizing the existence of multiple perspectives and acknowledging that some questions or answers can be sensitive or offensive. This means we need careful grounding of the evaluation process, defining cultural knowledge in concrete terms, and considering real-world usage scenarios where LLMs must perform reliably and inclusively.


**Emanuel Moss**

Emanuel Moss contributed to the discussion by emphasizing that culture cannot be encompassed by any individual or described by simple rules. He described AI as an artifact of culture and defined culture as a set of shared conceptions expressed through symbolic forms. Moss raised questions about the possibility of benchmarking and encoding culture, the relational and collective aspects of culture, and the potential harms of trying to benchmark culture. He also highlighted the importance of considering the processes of cultural production and the risks of encoding culture into corporate databases and models.

The key takeaways from Firestarter 2 included the recognition of the complexities and challenges in encoding culture within AI systems. The presenters illuminated the importance of considering diverse cultural perspectives and the potential harms of misrepresentation and exclusion.

### 3.4 Panel Discussion 3: Institutional Reflections and Collaborations (25022)

*Rida Qadri (Google – San Francisco, US), Naveen Bagalkot (Manipal Academy of Higher Education – Bangalore, IN), Catherine d’Ignazio (MIT – Cambridge, US), and Sara Hooker (Cohere For AI – Toronto, CA)*

License  Creative Commons BY 4.0 International license  
© Rida Qadri, Naveen Bagalkot, Catherine d’Ignazio, and Sara Hooker

The seminar then shifted to its third and final panel focused on the theme of “Institutional Reflections and Collaborations” that was presented by Catherine D’Ignazio, Sara Hooker, and Naveen Bagalkot.

#### Catherine D’Ignazio

The session began with Catherine D’Ignazio discussing her multi-year project working with activists to do participatory work at every stage of AI development. She emphasized the importance of privileging subjugated knowledges and questioned the current organization of resources in the AI ecosystem. D’Ignazio highlighted the need for sustainable and just resources, envisioning alternative initiatives that do not rely on corporate structures.

#### Sara Hooker

Sara Hooker, who leads Cohere for AI, discussed the importance of surplus and excess in driving innovation. She reflected on the history of scientific breakthroughs and the need for open science collaboration across many companies. Hooker raised concerns about the marginalization of academia and the underrepresentation of researchers from the Majority World. She emphasized the need for alternative, sustainable, and just approaches to AI development that support participation and inclusivity.

#### Naveen Bagalkot

Naveen Bagalkot contributed the last point of view to the discussion by sharing a narrative about the futures of AI. He highlighted the importance of centering the processes of cultural production and considering how technologies and interactions are result of these processes. Bagalkot emphasized the need for alternative imaginaries of technology and the importance of building collaborative teams that include diverse perspectives. He also discussed the challenges of changing research culture and the need for new funding structures that support interdisciplinary collaboration.

The key takeaways from Firestarter 3 included the recognition of the need for alternative approaches to AI development that prioritize sustainability, justice, and inclusivity. The presenters discussed the importance of participatory methods, open science collaboration, and the inclusion of diverse cultural perspectives. They also highlighted the challenges of the current corporate-dominated AI ecosystem and the need for new funding structures and research cultures that support interdisciplinary and inclusive innovation.

Each Firestarter session concluded with a call to rethink how cultural values are embedded in AI tools and to ensure that these tools are sensitive to the cultural contexts in which they operate.

## 4 Working groups

### 4.1 Working Group 1: Power, Future, History (25022)

*Rida Qadri (Google – San Francisco, US), Virgilio Almeida (Federal University of Minas Gerais – Belo Horizonte, BR), Naveen Bagalkot (Manipal Academy of Higher Education – Bangalore, IN), Georgina Born (University College London, GB), Anita Say Chan (University of Illinois at Urbana Champaign, US), Hal Daumé III (University of Maryland – College Park, US), Catherine d’Ignazio (MIT – Cambridge, US), Giovanna Fontenelle (Wikimedia – Sao Paulo, BR), Tarleton Gillespie (Microsoft New England R&D Center – Cambridge, US), Darci Sprengel (King’s College – London, GB), Molly Steenson (American Swedish Institute – Minneapolis, US & Carnegie Mellon University – Pittsburgh, US), and Harini Suresh (Brown University – Providence, US)*

**License** © Creative Commons BY 4.0 International license

© Rida Qadri, Virgilio Almeida, Naveen Bagalkot, Georgina Born, Anita Say Chan, Hal Daumé III, Catherine d’Ignazio, Giovanna Fontenelle, Tarleton Gillespie, Darci Sprengel, Molly Steenson, and Harini Suresh

Working Group 1 focused on the intricate interplay between power, future, and history in the context of AI and cultural inclusivity. This cluster, involving participants that represented a cross-section of institutional and disciplinary perspectives (Giovanna, Virgilio, Naveen, Harini, Catherine, Tarleton, Georgina, Molly, Darci, Anita, and Hal), delved into the visionary aspects of AI development and its implications for society. The discussions revolved around the vision, architecture, governance, and barriers to creating a network of shared infrastructures that foster alternative imaginations and inclusivity.

One of the central themes was the importance of situating AI development within the current socio-political context, including the rise of populism and authoritarianism. The participants emphasize the need to link AI initiatives to alternative histories, such as the cybernetic turn in India and socialism in Chile, to draw lessons from public institutions like libraries, universities, and archives. They discussed the concept of “defensive localism,” which involves creating urgent coalitions against AI authoritarian surveillance without requiring absolute political unity. This approach contrasts with “prospective, future-oriented place-based localisms,” which focus on long-term engagement with local politics to achieve justice and inclusivity.

The cluster also explored the idea of “open AI” and the challenges associated with it. While openness is seen as crucial to avoid the concentration of power, there are critiques of the concept, such as the limitations of open systems that require significant compute and technical knowledge. The participants discussed the potential of local communities to build and train their own models, considering the trade-offs of cost and quality. They highlighted the importance of a decentralized and federated structure that links smaller, local models to avoid dependency on global models. Such an approach could create an alternative ecosystem that aligns with the public good and the common good.

Another key takeaway was the need to address the material and affective demands of participation in AI development. The participants emphasized the importance of ensuring that data contributors’ livelihoods and incomes improve as a result of their participation. They discussed the challenges of engaging local communities in AI projects and the need for new kinds of education that speak to needs outside of commercial tech. The discussions also touched on the history of alternative ideologies in countries like Brazil and the need to create conditions for inclusivity that represent pluralism.

The cluster concluded with a call to formulate research questions that address the uncertainties and challenges identified. These questions could include how interdisciplinary collaboration can effectively identify and address the ethical and social risks of language models, how small language models can contribute to responsible innovation, and how to design decentralized infrastructure architectures that enable users to choose how they share and distribute their data and models. The participants also highlighted the need to pluralize the political economy of technology and reimagine futures through diverse cultural imaginaries.

In summary, the thematic cluster on power, future, and history emphasized the importance of situating AI development within a broader socio-political context, addressing the challenges of openness and decentralization, ensuring the material and affective demands of participation, and formulating research questions that guide future interdisciplinary collaboration.

## 4.2 Working Group 2: Interdisciplinarity / CS cultures (25022)

*Rida Qadri (Google – San Francisco, US), Asia Biega (MPI-SP – Bochum, DE), Tobias Blanke (University of Amsterdam, NL), Marc Cheong (The University of Melbourne, AU), and Mary L. Gray (Microsoft New England R&D Center – Cambridge, US)*

**License**  Creative Commons BY 4.0 International license

© Rida Qadri, Asia Biega, Tobias Blanke, Marc Cheong, and Mary L. Gray

The working group on Interdisciplinarity and Computer Science (CS) cultures explored the complexities and nuances of integrating interdisciplinary approaches in computer science research methods and theoretical frameworks. This cluster, involving participants from CS, anthropology and philosophy, with experience conducting mixed methods studies of AI (Asia, Marc, Mary, and Tobias) looked at the tensions, agreements, and common ground that can develop from merging different disciplinary perspectives and methodologies. The discussions highlighted the challenges and opportunities of fostering interdisciplinary collaboration and the need for a shared understanding and language to bridge the gap between computer science and social sciences/humanities.

One of the central themes of the discussion is the challenge of defining and using terms like “good” and “bad” within interdisciplinary contexts when communicating about AI. The participants quickly realized that these terms carry different meanings across disciplines, leading to potential misunderstandings and miscommunications. To address this, they emphasized the importance of specifying what is meant by these terms in different contexts and developing a common basic language for evaluating iterations of AI that do not assume there is a linear or universal path of improving AI for all users, regardless of context. This shared language would help clarify where disciplinary specificity is needed and where interdisciplinary collaboration can be most effective.

The cluster also explored the integration of qualitative and quantitative methods within computer science. Participants discussed the potential for developing and evaluating models using qualitative methods alone and the need for reflexivity from both social sciences/humanities (SSH) and computer science (CS) about the limits and peculiarities of their ways of knowing and forms of evidence. They highlighted the importance of interdisciplinary teaming at specific points in the development pipeline, imagining pairs of experts from technical and qualitative fields working together step-by-step to negotiate approaches that meet shared goals. This collaborative approach would ensure that both qualitative insights and quantitative rigor are incorporated into AI development.

Another key takeaway from the discussion is the role of participatory (re)design, crowdsourcing, and citizen science in interdisciplinary AI development. Participants emphasized the importance of involving diverse stakeholders in the development process and ensuring equitable terms for their participation. They discussed the potential for deliberative development processes that include input from various stakeholders, including those from civil society organizations, industry, and academia. This inclusive approach would help ensure that AI systems are developed with a broader range of perspectives and are more attuned to the needs and values of different communities. The open question was how, exactly, to sustain these multistakeholder codesign efforts, given market pressures and the lack of meaningful connections with diverse community groups of experts available to CS.

The cluster also addressed the challenges of forming interdisciplinary projects without a shared definition of what counts as generative AI—and what will be recognized, professionally, as meaningful contributions to the field. STS and social science-oriented participants agreed that their qualitative methods are often misunderstood or misused on the CS side and that there is an underappreciation of multiple methodologies. They emphasized the need to understand where qualitative analysis should fit in the development and evaluation pipeline and the critical importance of data provenance for meaningful evaluation. The discussions also touched on the philosophical models of reality and knowledge that are useful for thinking about the evaluation process, particularly in the context of foundation models that lack a typical ground truth.

In summary, the thematic cluster on Interdisciplinarity and Computer Science Cultures highlighted the complexities and challenges of integrating interdisciplinary approaches within AI development. The discussions centered the importance of developing a shared language and understanding, incorporating both qualitative and quantitative methods, involving diverse stakeholders in the development process, and addressing the philosophical and methodological challenges of evaluating AI systems.

### 4.3 Working Group 3: Culture Encodability (25022)

*Rida Qadri (Google – San Francisco, US), Kalika Bali (Microsoft Research India – Bangalore, IN), Beth Coleman (University of Toronto, CA), Fernando Diaz (Carnegie Mellon University – Pittsburgh, US), Huma Gupta (MIT – Cambridge, US), Sara Hooker (Cohere For AI – Toronto, CA), Maurice Jones (Concordia University – Montreal, CA), Emanuel Moss (Intel – Santa Clara, US), Maryam Mustafa (LUMS – Lahore, PK), Alice Oh (KAIST – Daejeon, KR), and Moira Weigel (Harvard University – Cambridge, US)*

**License** © Creative Commons BY 4.0 International license

© Rida Qadri, Kalika Bali, Beth Coleman, Fernando Diaz, Huma Gupta, Sara Hooker, Maurice Jones, Emanuel Moss, Maryam Mustafa, Alice Oh, and Moira Weigel

Working Group 3 (Fernando Diaz, Moira Weigel, Huma Gupta, Rida Qadri, Sara Hooker, Maurice Jones, Manny Moss, Kalika Bali, Alice Oh, Maryam Mustafa, and Beth Coleman) took up the challenges of encoding cultural nuances into AI, particularly how to develop technical interventions to preserve the richness of cultures but also cultural protocols that consider whether we ought to/should encode culture. Participants debated whether increasing data volume or enriching data with context was more crucial, and whether solutions lay solely in data or also in model design and evaluation. There was general agreement on the need for richer evaluation methods, non-data-centric interventions like model optimization

changes and interface design, and further research into effective encoding strategies. This cluster further explored examples of the complexities of capturing cultural specificity and the technical approaches that might be used to enhance how models represent cultural expression.

One of the central themes of the discussion was the fundamental question of what aspects of culture can and should be encoded and what governance mechanisms could direct these socially consequential decisions. Participants emphasized the importance of being specific about what cultural elements are being targeted for encoding and the limitations of existing technical processes. They discussed, for example, the challenges of condensing culturally nuanced language into text and the loss of cultural variance in machine translation. The conversation highlighted the need for a deeper understanding of the polysemy and thickness of culture, such as the different structures of languages and the epistemic shifts that occur within them. For example, indigenous languages often have a higher proportion of verbs compared to nouns, which presents unique challenges for encoding.

The cluster also explored the disagreements and agreements around the need for more data versus the need for thicker, more contextually rich data. Some participants argued that more data is necessary to capture the full range of cultural expressions, while others contended that the focus should be on developing thicker development pipelines that incorporate expertise and context. They discussed the limitations of current models, which often operate on crude metrics and may not adequately represent the richness of cultural data. The conversation also touched on the potential for non-data-based interventions, such as designing models that indicate their positionality and highlight absences in the data.

Another key takeaway from the discussion was the importance of cultural protocols in the encoding process. Participants emphasized the need for guidelines on what should and should not be encoded and how to ensure that cultural knowledge is represented accurately and respectfully, without placing the burden of identifying harms on those who might be the most likely targets of them. Particularly of concern was how to think about generating more data without further surveilling data contributors and, on the other hand, the limits of using synthetic or existing datasets that will degrade in accuracy and temporal relevance. The working group discussed, for example, the challenges of creating relational databases that link cultural data to archaeological expertise and the limitations of such approaches. The conversation also highlighted the need for research on whether cultural interventions at different points in the development pipeline are effective and how to design user interfaces that are culturally sensitive.

The cluster also addressed the issue of data absences and the challenges of representing missing cultural information. Participants discussed the potential for structuring data in ways that make absences more visible and the importance of acknowledging the partiality of model outputs. They emphasized the need for models to be transparent about their limitations and the gaps in their data. The conversation also touched on the ethical considerations of data collection and the potential harms of hyper-surveillance and extraction.

In summary, the Working Group 3 on Cultural Encodability highlighted the complexities and challenges of encoding cultural knowledge within AI systems. The discussions centered on the importance of being specific about what cultural elements are being targeted for encoding, developing thicker development pipelines, and adhering to cultural protocols. They also emphasize the need for transparency about data absences and the ethical considerations of data collection.

The seminar's second day sent participants away with some homework, asking them to reflect on what artifacts and projects they would want to specifically take forward as outcomes of the Seminar for building a multidisciplinary research agenda.

Participants spent the third and final day of the seminar working in small groups to identify actionable research directions, fueled by group insights from the Firestarter talks and the Day 2 working group discussions. By the end of our last morning together, seminar participants had identified three specific directions to continue from our seminar: 1) development of an agenda-setting document for research on cultural representation and AI; 2) specific projects aimed at large language models for linguistic diversity; and 3) a clever approach to encoding, dubbed “meta-meta data” evaluation and documentation.

## 5 Open problems

### 5.1 Future directions based on participant feedback (25022)

*Rida Qadri (Google – San Francisco, US), Asia Biega (MPI-SP – Bochum, DE), Georgina Born (University College London, GB), Fernando Diaz (Carnegie Mellon University – Pittsburgh, US), and Mary L. Gray (Microsoft New England R&D Center – Cambridge, US)*

**License** © Creative Commons BY 4.0 International license

© Rida Qadri, Asia Biega, Georgina Born, Fernando Diaz, and Mary L. Gray

The Dagstuhl Seminar, “Towards a Multidisciplinary Vision for Culturally Inclusive Generative AI,” received high praise from participants, who appreciated the interdisciplinary nature of the seminar and the diverse range of fields and disciplines represented. Participants found the seminar to be an unprecedented experience that brought together a broad scope of multidisciplinary research and backgrounds not typically found at computing research venues, fostering rich discussions and collaborations that several noted as a first encounter with that discipline. Many participants noted that the seminar inspired new ideas for their research, development, or teaching.

One of the most frequently mentioned positive aspects of the seminar was the high quality of attendees and the organization of the event. Participants appreciated the format, which included firestarter presentations and ample time for informal conversations over meals and coffee. These informal discussions were seen as highly generative, leading to meaningful exchanges and the development of new ideas. The interdisciplinary and cross-cultural focus of the seminar was also highlighted as a significant strength, with participants noting that it allowed for a deeper understanding of the challenges and opportunities in the field of generative AI and cultural diversity.

However, participants also provided several recommendations for changes to improve future seminars. One of the most common suggestions was to extend the duration of the seminar. Many participants felt that the seminar was too short, given the challenges of developing a shared language for key contested concepts like “cultural representation.” They felt an additional day would have allowed for deeper engagement and more thorough exploration of the topics as well as opportunities to establish solid next steps for collaboration. This extension would also provide more time for informal conversations during the day, which participants found to be highly valuable.

Another recommendation was to include a broader range of voices in future seminars. Participants suggested incorporating more representatives from civil society organizations, funders, and philanthropists, as well as increasing the representation of researchers from regions such as Africa, China, and Latin America. Additionally, some participants recommended involving more junior researchers and providing more opportunities for socializing and personal discussions.

Participants also highlighted the importance of including more detailed case studies and examples of interdisciplinary work in future seminars. They felt that this would generate more concrete and detailed ideas about extending this kind of research. Some participants suggested that the seminar could benefit from more explicit links to other participants' work and position statements beforehand, as well as pre-meeting introductions to help participants get to know each other before arriving at the seminar.

The survey results revealed the success of the Dagstuhl Seminar in fostering interdisciplinary collaboration and generating new ideas. Participants appreciated the unique environment provided by Schloss Dagstuhl.



## Participants

- Virgilio Almeida  
Federal University of Minas  
Gerais – Belo Horizonte, BR
- Elisabeth André  
Universität Augsburg, DE
- Naveen Bagalkot  
Manipal Academy of Higher  
Education – Bangalore, IN
- Kalika Bali  
Microsoft Research India –  
Bangalore, IN
- Asia Biega  
MPI-SP – Bochum, DE
- Tobias Blanke  
University of Amsterdam, NL
- Georgina Born  
University College London, GB
- Anita Say Chan  
University of Illinois at Urbana  
Champaign, US
- Marc Cheong  
The University of Melbourne, AU
- Beth Coleman  
University of Toronto, CA

- Hal Daumé III  
University of Maryland –  
College Park, US
- Fernando Diaz  
Carnegie Mellon University –  
Pittsburgh, US
- Catherine d'Ignazio  
MIT – Cambridge, US
- Giovanna Fontenelle  
Wikimedia – Sao Paulo, BR
- Tarleton Gillespie  
Microsoft New England R&D  
Center – Cambridge, US
- Mary L. Gray  
Microsoft New England R&D  
Center – Cambridge, US
- Huma Gupta  
MIT – Cambridge, US
- Sara Hooker  
Cohere For AI – Toronto, CA
- Maurice Jones  
Concordia University –  
Montreal, CA

- Emanuel Moss  
Intel – Santa Clara, US
- Maryam Mustafa  
LUMS – Lahore, PK
- Alice Oh  
KAIST – Daejeon, KR
- Rida Qadri  
Google – San Francisco, US
- Noopur Raval  
University of California at  
Los Angeles, US
- Darci Sprengel  
King's College – London, GB
- Molly Steenson  
American Swedish Institute –  
Minneapolis, US & Carnegie  
Mellon University –  
Pittsburgh, US
- Harini Suresh  
Brown University –  
Providence, US
- Moira Weigel  
Harvard University –  
Cambridge, US

