
Digital Intersectionality and Marginalization in the Majority World

Sujata Mukherjee, Mona Elswah, and Crystal Abidin

The concepts of ‘intersectionality’ and ‘marginalization’ are empirical necessities in Majority World trust and safety research because the harms that Majority World users face are constituted by multiple systems of power, not by any one axis alone. In this context, marginalization operates at three registers simultaneously: the marginalization of Majority World *users* by platform infrastructure that are not inherently designed to accommodate their needs; the marginalization of Majority World *communities* when existing frameworks of harm do not acknowledge or recognize their lived experiences; and the marginalization of Global Majority *scholarship* by the field’s own institutional and review structures. In response, we initiated this special issue to explore Digital Intersectionality and Marginalization in the Majority World. Here we reflect on our editorial process, from the initial set of 59 submissions we received, to the 8 peer-reviewed papers that we ultimately accepted. We examine these published papers through the lens of intersectionality and marginalization before introducing them across three thematic clusters: moderation systems and AI infrastructure; platform governance; and gendered harm and queer safety. Across these, we corroborate a unifying structural insight: Global North safety frameworks often reproduce structural inequality when deployed in Majority World contexts. We also name an additional dimension of intersectionality that structures the issue itself: the intersection of academic knowledge production and industry practice that defines trust and safety research as a field, and that this journal was designed to sustain. We close with implications for journals, funders, and platforms, and with a reflection on our positionality as editors who are simultaneously Global North-based, Majority World-committed, and situated across the academic-industry divide. This special issue is a call to action for advancing research and practice on trust and safety issues in under-examined regions.

1 The Problem this Issue was Designed to Address

Our July 2025 call for papers opened with a straightforward observation: trust and safety¹ practices have largely been shaped by Western perspectives, rarely addressing the unique requirements of the Majority World.² Guiding policy principles at major online platforms—free speech, individual agency, and Western-centric harm taxonomies—frequently do not align with the diverse experiences of people in Africa, Asia, Latin America, the Middle East and the Pacific (Díaz and Hecht-Felella 2021; Sambasivan et al. 2021). This is not a new critique, but it is one that is still in the fledgling stages of generating a body of peer-reviewed trust and safety scholarship proportional to its urgency.

We chose to title this special issue “Digital Intersectionality and Marginalization in the Majority World.” Intersectionality is a concept developed by Kimberlé Crenshaw to describe how multiple systems of inequality compound and interact, originally in the context of anti-discrimination law but with implications that have since reshaped feminist theory, critical race studies, and social science methodology (Crenshaw 1989). Harms experienced by users are rarely singular; they reflect gender compounded by language, geography, and politics. As the contributions in this issue repeatedly demonstrate, understanding digital harm in these contexts requires an intersectional lens.

Marginalization names the other half of the problem. Majority World users are marginalized in the design of platform infrastructure. Majority World researchers face “data scarcity, safety concerns, methodological limitations, language barriers, and restricted access to global discourse forums” (Mohamed, Png, and Isaac 2020; Nigatu et al. 2026). This special issue is an attempt to address that marginalization in a small way—by creating a dedicated space and designing a process that was honest about the structural advantages and disadvantages that researchers bring to it.

Throughout this editorial we use the terms “Global Majority” and “Majority World” interchangeably.³ Neither term implies homogeneity. The Global Majority encompasses radically different political economies, linguistic ecologies, legal systems, and platform governance histories. The intersectionality and marginalization this issue names operates within Global Majority contexts, not only between them and the Global North: class, language, religion, nationality, and political context compound differently across and within regions. This special issue is thus intended to surface the specificity of particular struggles in particular places. For platform trust and safety teams, this specificity carries a direct implication: the concerns of Global Majority users, while shaped by shared structural conditions, are not identical across regions, and solutions designed with one

1. We follow the Trust & Safety Professional Association (TSPA) T&S Curriculum style guide: lowercase “trust and safety” as an adjective for the field, capitalized “Trust and Safety” sparingly as a proper noun, and avoid “T&S” except where space requires.

2. “Majority World” (Alam 2008) and “Global Majority” (Campbell-Stephens 2021) were coined to reframe power dynamics against Eurocentric perspectives, pointing out that decision-making nations represent a small fraction of global population.

3. To refer to non-Western/post-colonial regions (e.g., Southeast Asia, Latin America, the Middle East, Africa), replacing terms like “Global South” or “developing countries.”

context in mind are not always portable to another. We use “Global Majority” and “Majority World” as strategic coalitions—a way of naming a shared structural position relative to dominant platform governance—not as descriptions of shared experience.

This special issue builds on the momentum of recent institutional recognition within the field, including dedicated sessions at TrustCon,⁴ the *Journal of Online Trust and Safety*’s pilot grant program,⁵ and the launch of the Global Majority Research Committee (GMRC) by the Research Coalition of the Trust and Safety Foundation (TSF).⁶ It also complements growing literature on content moderation in low-resource languages (Elswah, Bhatia, and Thakur 2025; Nicholas and Bhatia 2023), language equity in AI-driven moderation (Mukherjee and Mathew 2026), and localized harms (Houtti et al. 2024; Mukherjee 2026; Townsend and Gwagwa 2023; Umbach, Singh, and Walker 2023).

Finally, this issue reflects another critical dimension of intersectionality in the field: the intersection of academic knowledge production and industry practice. Research in trust and safety as a field does not sit neatly on either side of the academic-industry divide, it emerged and continues to evolve as a hybrid necessity. The platform practitioners who develop and enforce content policy need empirical research to guide them in ambiguous or novel situations, particularly in the Majority World contexts they increasingly govern but rarely understand from the inside. The academic researchers who study platform affordances and governance need access to operational context, institutional knowledge, and the lived experience of practitioners to produce findings that are meaningful rather than merely publishable. The *Journal of Online Trust and Safety* and the Trust and Safety Research Conference, both now hosted by the Stanford Tech Impact and Policy Center, were designed to sustain exactly this bridge: to create a venue where research from across the academic-practitioner spectrum is held to shared standards of rigor and relevance, and where neither community can ignore the other. This special issue inherits and extends that institutional intersectionality. It brings Majority World research into the trust and safety field, and it does so through an editorial team that itself crosses the academic-industry boundary.

The 59 submissions we received and the 8 papers we ultimately published taught us far more than we anticipated. In this editorial, we share those lessons across three thematic clusters, closing with reflections on our own positionality as editors and its bearing on the strengths and limitations of this issue.

4. TrustCon is the TSPA’s annual global conference for industry professionals, engineers, policymakers, and researchers dedicated to online safety.

5. Initiated by the Stanford Internet Observatory, the original home of the *Journal of Online Trust and Safety*, in December 2023 with support from the Omidyar Network. This program offered up to \$15,000 per award to support researchers studying online trust and safety in the Majority World.

6. The TSF’s Research Coalition was created to bridge gaps between industry researchers, academia, civil society, and tech practitioners across the trust and safety ecosystem. As part of this effort, the GMRC was established in 2024 to foster a more equitable ecosystem by elevating Majority World perspectives and expertise in product design, policy development, and organizational practices. This special issue was initiated as a GMRC goal in 2025, with publication planned for 2026.

2 The State of Global Majority Trust and Safety Research

2.1 The Geography of Voice: Who Submitted, and from Where

The call generated 59 abstract submissions from researchers studying 27 countries—a response that surprised us in its volume and moved us in its reach. Submissions examined contexts in Kenya, Palestine, Mexico, Sri Lanka, Indonesia, Peru, Iran, Bangladesh, India, Ghana, Nigeria, and many others. About 86% focused on Global Majority contexts, and a further 5% studied Global Majority diaspora communities in Northern settings. Among the remaining submissions, 5% focused on non-Global Majority or Northern contexts, and 3% were explicitly cross-cutting or global in scope.

These numbers are, on one reading, encouraging. They suggest that an active community of Global Majority trust and safety researchers exists, that there is significant scholarly activity in these regions, and that a journal willing to prioritize this work can find it. On another reading, they are sobering. These 27 countries represent only a fraction of the Majority World. Across the African continent, the most represented countries were Kenya, Ghana, and Nigeria, the English-speaking hubs with established research university systems, and not coincidentally, the countries where major platforms have established trust and safety operations. East Asia was nearly absent from the submission pool. Central Asia, the Caucasus, and large parts of South East Asia and the Pacific did not appear at all. French-speaking and Arabic-speaking Africa were present primarily through diaspora-based researchers.

This geographic clustering has a straightforward explanation: we published our call in English, in an English-language journal, disseminating the opportunity across personal and academic networks that skew toward English-language institutions. The researchers who submitted to this issue are, overwhelmingly, either native English speakers, educated in English-medium institutions, or working in contexts where English-language publication is professionally incentivized. The scholars whose work we never saw are, in many cases, working in the very contexts our issue was designed to illuminate. There is a compounding structural factor, documented across global science (Nakamura et al. 2023): research from the Global South is routinely required to name its geography—to position itself as a study of Kenya or Bangladesh or Colombia—while research based in Western contexts claims universal applicability by default. The same asymmetry operates in Trust and Safety: “Global Majority Trust and Safety research” is a subfield that must declare its geographic specificity, while “Trust and Safety research” signifies the unmarked WEIRD (Western, Educated, Industrialized, Rich, Democratic) norm. Not merely a naming convention, this reflects an hierarchy of epistemic legitimacy embedded in how knowledge is produced, circulated, and cited—one that shapes which submissions arrive at a journal like this one, how their authors frame themselves, and whose situated knowledge must perform additional labor of justification before it can be heard. We note this not as a reason to discount what we found, but as a caution against over-

interpreting it. The 27 countries in this submission pool represent a significant expansion of trust and safety research visibility, but they do not represent Global Majority research comprehensively.

One pattern from our analysis of the submission pool is worth highlighting as speaking directly to the issue's core theme of marginalization. A meaningful proportion of the most methodologically rigorous submissions came from researchers who were either based at Global North institutions or had graduate training there. In-region scholars—those doing research at institutions in the countries they were studying—faced, on average, greater challenges with methodological framing, literature positioning, and the conventions of writing for academic journals. This is not a comment on the intellectual quality of their work; several of the most important empirical insights in the pool came from in-region researchers. It is a comment on what academic publishing rewards, and what it costs to enter it on those terms. Nigatu and colleagues, studying peer review in Human-Computer Interaction—a field with close ties to Trust and Safety—identify this as the “hidden curriculum” of academic writing: the conventions of framing, citation practice, and argument structure that are transmitted through mentorship and informal socialization at Northern institutions and are systematically inaccessible to researchers who have not had that training (Nigatu et al. 2026). Access to this curriculum, not intellectual capacity, is what disproportionately distinguishes submissions that are positioned to succeed in journal review in the Global North from those that are not. The hypothesis that Global Majority scholars are more likely to be heard if they relocate to Global North institutions found more support in this process than we would have liked.

2.2 Quality, Fit and the Real Cost of Participation

Of the 59 abstracts submitted, 31 were invited to submit full manuscripts, of which 23 manuscripts were eventually received. Of those, 10 progressed to double-blind external peer review and all were invited to revise and resubmit. Of these, 8 manuscripts passed a further round of editorial peer review for final publication in this issue.

Our editorial process for this special issue applied the same standards of scholarly rigor per the standards regarded by the *Journal of Online Trust and Safety*. The commitment to including Global Majority voices does not mean accepting work that does not meet those standards; it means investing in the conditions that allow more work to reach them. To this end, we channeled the call for proposals through non-standard and non-academic channels to reach communities of in-region scholars and practitioners that were doing this work. The co-editing team have lived experiences in the Global Majority that are current and ongoing, and work across industry and academia on exactly these topics. For this special issue, we deliberately matched specialist reviewers to specific papers based on the fit of knowledge domains, regional understanding, and positional proximity and reflexivity. These were researchers and practitioners who had themselves worked in, alongside, or in sustained scholarly engagement with the communities and contexts represented in the submitted manuscripts.

Table 1: Thematic clusters and cross-cutting dimensions for the 59 submissions. The two cross-cutting dimensions appeared across submissions, so their counts overlap with the four thematic clusters.

	Theme	Description	Submissions
Thematic clusters	Content Moderation	Labor, practice, & civic use	11
	AI & Algorithmic Systems	Deepfakes, large language models, & automation	11
	Platform Governance	Regulation, policy, & oversight	29
	Gender-Based Violence & Queer Safety	Online harm & feminist resistance	8
Cross-cutting dimensions	Language & Multilingualism	Multilingual & minority language contexts	8
	Colonial & Decolonial Frameworks	Power, knowledge, & digital coloniality	16

As editors, we focused on genuine development of the research quality, offering substantive feedback designed to strengthen work rather than merely evaluate it, extending deadlines when original timelines placed a disproportionate burden on researchers without institutional support structures, and being explicit about what the sequential revisions would need to demonstrate. We do not claim to have resolved all structural tensions in the publishing process. But conflating all Global Majority scholarship as inherently disadvantaged is its own form of condescension, and we have tried to avoid that while taking the concerns of this community of scholars seriously.

2.3 What the Field is Studying, and What It is Not

The 59 submissions could be cleanly organized into four thematic clusters (Table 1): Content Moderation, AI and Algorithmic Systems, Platform Governance, and Gender-Based Violence and Queer Safety. Two cross-cutting dimensions ran through multiple clusters: Language and Multilingualism and Colonial and Decolonial Frameworks. These clusters map closely onto the topic area named in the call for papers: digital marginalization and vulnerability; emerging technologies; postcolonialism and the digital world; cultural and linguistic diversity online; intersectional approaches.

The distribution of the 8 published papers does not mirror the submission pool. Platform Governance was the largest cluster in submissions and is represented in the final selection by only 2 papers (Chala 2026; Pachouri, Raj, and Harshavardhan 2026). In contrast, AI and Algorithmic Systems is represented by 3 papers (Mustafa and Wu 2026; Gamage et al. 2026; Arora, Miranda Mora, and Zarzycka 2026), and Gender-Based Violence and Queer Safety accounts for 3 papers (Thakur 2026; Hooman and Hodson 2026; Aljizawi et al. 2026).

We defined Content Moderation as the study of who does moderation work, at what

cost, and in what languages (cross-cutting theme) and socio-cultural contexts. This cluster comprised 11 submissions but produced no published papers. This is not a finding about the importance of the topic; content moderation remains among the most consequential aspects of Global Majority platform governance and was explicitly named in the call as a priority area. Rather, it is a finding about research infrastructure. The Content Moderation submissions that came through this process were, in many cases, missing the methodological scaffolding they needed, primarily the limited access to platform data. This reflects the need for greater investment in data access, peer mentorship and reviewer networks for the scholars studying this topic, many of whom tend to be practitioners with less experience in academic publishing.

The over-representation of Gender-Based Violence in the final papers reflects something different: the quality and urgency of the work in this area. Researchers studying gendered digital violence in Global Majority contexts are doing some of the most methodologically sophisticated and theoretically grounded work in the field. They are also—and this is not a coincidence—often the researchers whose own communities are most directly affected by what they are studying.

In addition to discussing what was included, several significant absences stand out across the submissions, and they are worth naming because they illustrate the frontier of what this special issue could not do. China-based platforms with substantial global reach remain understudied in this submission pool.⁷ We also found limited coverage of Central Asian digital governance, no coverage of Francophone West Africa, and while a single submission did address Pacific Islander communities, it did not advance to publication. Platform economics and labor—the conditions under which content moderation is performed in the Global Majority—was addressed partially for some centers in Africa, but did not cover those in other parts of the world, like the Philippines for example. User wellbeing and mental health outcomes from platform harms, outside of Gender-Based Violence research, was essentially absent.

2.4 Implications for Journals, Funders and Research Infrastructure

The picture that emerges from analysis of the editorial funnel is one of significant scholarly energy in Global Majority trust and safety research, the existence of structural constraints on that energy's expression, and gaps in coverage that no amount of scholarly enthusiasm can bridge without institutional investment.

For journals: they must build the scaffolding that enables work from under-resourced contexts to reach them. First, this means expanding outreach beyond traditional academic pipelines to engage broader peer networks, grassroots communities, and practitioner spaces across the Global Majority. However, because informal networks can inadvertently replicate existing hierarchies and privilege already-connected researchers, this outreach must intentionally bypass established gatekeepers. We recommend a systematic map-

7. Though such work has been published in a prior special issue (Fang 2024)

ping of where local discourse and practice occur—including non-English, regional, and non-academic forums—and call on journals and institutions to actively maintain and democratize access to these discovery channels. Second, it means dealing with Global Majority peer reviewer pools as an epistemic necessity. Reviewers without experience in a given context are structurally ill-equipped to evaluate empirical claims about it. Third, it means accounting for deep structural support during review timelines, beginning with substantive feedback that strengthens work rather than opaque refusals, hosting paper workshops, and assisting with co-author matching and mentorship. Within this scaffolding, it is also a requirement for reviewers to reflect on whether their requested revisions and citations reinforce Northern scholarship, which is often treated as the theoretical baseline, and under-reference Majority World scholarship.

For funders: the trust and safety research funding landscape is significantly concentrated in the United States and Europe. The major private philanthropic funders of online safety practice and research have, with some exceptions, primarily funded Northern institutions (Atlantic Council 2023). While our special issue did not involve a grant, the GMRC has attempted to channel funds to support Global Majority research being performed directly in the region. We have found the most value in proactive outreach, grant-writing support, and a genuine willingness to accept methodological diversity that comes with funding research not structured by Northern academic conventions. At times, funders must also be willing to adapt new administrative procedures to navigate the legal, banking, and sanction-related hurdles required to get resources to researchers in underbanked or conflict-prone regions. These adaptations often incur additional expense and risk for many North-based organizations. Attention should also be paid to the structural costs that often go ignored, such as the open-access publication fees at many journals, which can average thousands of dollars per article, though the *Journal of Online Trust and Safety* provides its articles as open-access for free. More broadly, funders should recognize what Nakamura and colleagues call “parachute science”—in which Northern-led research projects extract data from Majority World contexts without meaningful engagement or co-authorship for in-region researchers—as a disqualifying practice, not a model to fund (Nakamura et al. 2023).

For platforms: the large-scale data access programs that make quantitative trust and safety research possible—operated by Meta, YouTube, TikTok, and others—are accessible primarily to researchers at accredited US and European institutions. The contraction of platform data access in recent years has hit Global Majority researchers disproportionately because they have less institutional buffer (Freelon et al. 2024; Mimizuka et al. 2025; Santini et al. 2026). The design of data access programs encodes assumptions about who counts as a legitimate researcher that systematically disadvantage the Global Majority. This is a choice, and it can be made differently.

3 Three Clusters, Common Themes

The eight papers in this issue are organized into three clusters. The first, Moderation Systems and AI Infrastructure, brings together three papers examining different layers of the same digital safety infrastructure: AI safety benchmarks (Mustafa and Wu 2026), deepfake governance (Gamage et al. 2026), and feminist counter-infrastructure (Arora, Miranda Mora, and Zarzycka 2026). The second, Platform Oversight, examines the institutional mechanisms by which platforms govern speech and harm, through the lens of formal oversight structures (Chala 2026) and digital public infrastructure (Pachouri, Raj, and Harshavardhan 2026). The third, Gendered Harm and Queer Safety, examines the specific harms that fall through the gaps of that infrastructure, and the communities that bear the cost (Thakur 2026; Aljizawi et al. 2026; Hooman and Hodson 2026).

These are analytically distinct clusters. The consistent critique of platforms' norms is upheld across all papers: the infrastructure of platform safety—its tools, benchmarks, oversight institutions, harm frameworks—was designed in and for the Global North, and its deployment in Majority World contexts reproduces rather than resolves structural inequality. Evidence is provided across the full stack: from AI benchmark design (Mustafa and Wu 2026) to formal governance institutions (Chala 2026) to the targeted harms experienced by Quechua women creators (Thakur 2026) and exiled Iranian queer activists (Hooman and Hodson 2026). They also collectively suggest what a different architecture might look like: rooted in communities, designed for the actual users it serves, and honest about who bears the cost when it fails.

3.1 The Papers Against the Frame: Intersectionality and Marginalization

The first step in our analysis is to assess the eight published papers against the two organizing concepts named in the special issue's title: digital intersectionality and marginalization. This is not a rubric for ranking the papers; it is a way of being honest about what the issue delivers on its own terms, where the coverage is uneven, and what that unevenness reveals about the state of Global Majority trust and safety research.

Intersectionality. The call defined intersectionality expansively—the compounding of gender, race, class, language, and geography in shaping digital experiences and harms. Against this definition, the eight papers fall on a clear spectrum. Three deliver analyses that are fully intersectional in the constitutive sense, where the harm or vulnerability being studied is not just correlated with multiple identity axes but is produced by their crossing. Thakur's (2026) Quechua women influencers face an abuse pattern that is simultaneously misogynistic, racially marked, linguistically targeted, and colonial in its geography; disaggregating any one axis makes the harm unrecognizable. Hooman and Hodson's (2026) Iranian queer activists face transnational repression that is constituted by sexuality, nationality, exile status, and language operating together: being queer is not sufficient to trigger this repression, neither is solely being Iranian or being in exile—it

is the intersection that creates the specific vulnerability. Aljizawi et al.'s (2026) exiled queer women human rights defenders document a protection that exists precisely at the crossing of gender, sexuality, nationality, immigration status, and political visibility. None of these papers could have been written in a single-axis framework.

Arora, Miranda Mora, and Zarzycka's (2026) analysis occupies a productive middle position. Their framing of "gendered digital creative violence" is explicitly intersectional along the axes of gender, language (Spanish-language legal ecosystem), and access to legal remedy (class and geography), though gender as mediated by AI technology is the dominant axis. The feminist counter-infrastructure they document—OlimpIA⁸—is designed for a specific intersectional subject: a Spanish-speaking woman in Mexico facing AI-generated violence with limited institutional recourse. The paper's intersectionality is present in the design of the solution as much as it is in the analysis of the harm.

The remaining four papers address intersectionality as structural rather than identity-level analysis, which is a different, and equally legitimate, register of the concept. Mustafa and Wu (2026) examine language and geography as structural axes shaping AI safety evaluation infrastructure: the argument is that multilingual communities are systematically disadvantaged by benchmarks designed for English models, but the analysis operates at the aggregate level of linguistic communities rather than individual identity intersections. Gamage et al. (2026) attend to gender within their deepfake governance analysis, and the paper's three-country comparative design implicitly addresses linguistic and geographic intersection, while the primary frame remains platform governance failure across South Asian language contexts rather than the compounding of personal identity axes. Chala's (2026) Oversight Board analysis treats language and region as structural variables that shape institutional processing—intersectional in the sense of identifying how multiple structural axes compound, but not in the sense of centering the lived experience of intersectional subjects. Pachouri, Raj, and Harshavardhan's (2026) study of e-KTP, Indonesia's national digital identity system, is the paper farthest from the call's intersectional frame: its primary axes are postcolonial statehood and digital inclusion, with socioeconomic and regional variables included in the structural equation model but personal identity intersection is not its lens.

Marginalization. The call identified three distinct registers of marginalization that this issue was designed to address—the marginalization of Majority World users by platform infrastructure not built for them; the marginalization of Majority World communities by harm frameworks that do not recognize their harms as harms; and the marginalization of Majority World research by the field's institutional structure. All eight papers address the first two registers to varying degrees. None fully addresses the third from within. The distribution across the first two registers is telling.

8. OlimpIA is an artificial-intelligence-driven digital conversational tool developed through a collaboration between the Ley Olimpia movement, Digital Defenders, and the technology company AuraChat. This is described as feminist counter-infrastructure in this issue's article by Arora, Miranda Mora, and Zarzycka (2026).

The Gender-Based Violence and Queer Safety cluster addresses both registers with the greatest explicitness: Thakur (2026) documents platform infrastructure that cannot read Quechua and therefore cannot moderate Quechua-language abuse; Hooman and Hodson (2026) document harm frameworks that classify coordinated transnational repression as ordinary harassment; Aljizawi et al. (2026) document cybersecurity training designed for a threat model that does not correspond to the actual threats faced by queer women human rights defenders. In all three, the marginalization is not incidental but structural: the infrastructure was designed for different users, and its failure to serve these communities is predictable from first principles. Arora, Miranda Mora, and Zarzycka (2026) add a fourth dimension: the marginalization of Mexican women by a legal system that recognizes 76.9% of digital violence cases as insufficient for legal remedy, and the response (OlimpIA) is among the most direct interventions in this mode of marginalization in the published set.

The AI and Moderation cluster addresses user marginalization with the most quantitative precision: Mustafa and Wu's (2026) finding that 87.1% of benign Bengali queries are incorrectly blocked by safety filters is a measurement of infrastructure-level marginalization in basis points. Gamage et al.'s (2026) documentation that South Asian-language deepfake content remains available after reporting while English-language equivalents are removed measures the second register directly: the harm is not recognized as harm because the platform cannot process the language in which it occurs. The Platform Oversight cluster adds a third variant: Chala (2026) shows that the Meta Oversight Board's processing pipeline structurally advantages English-language cases even when outcomes are not significantly different, a finding about marginalization embedded in institutional harm rather than in content decisions. Pachouri, Raj, and Harshavardhan (2026) is the outlier: it is the one paper in this issue documenting marginalization addressed, partially and imperfectly, by a Majority World's state's own digital public infrastructure design.

Read together against both lenses, the eight papers confirm the special issue's central premise: the Global Majority digital context is not a monolith, and intersectionality and marginalization are critical analytical lenses. The communities most subject to intersectional harm are also the communities most marginalized by the infrastructure that is nominally designed to protect them. Also, the infrastructure failures that produce the most quantifiable marginalization are themselves the product of design choices that failed to account for the intersectional reality of the users they would serve. The clusters that follow trace this logic through three domains of platform safety.

3.2 Moderation Systems & AI Infrastructure: Who The Systems are Built For

The most foundational question in platform trust and safety is: what are these systems actually for? Who were they designed to serve, and what harms were they designed to prevent? The three papers in this cluster approach that question from different angles: the benchmarks that evaluate AI safety systems, the community governance of deepfakes, the feminist counter-infrastructure built when platform systems fail. They all reach the

same answer: these systems were built for a particular kind of user, in a particular kind of language, living a particular kind of life. All other user types are afterthought.

Mustafa and Wu’s structured review of 37 multilingual AI safety benchmarks finds that 18 of the 27 safety-focused benchmarks scored zero on the Dialect and Code-Switching dimension—the dimension most directly relevant to how real users in multilingual societies actually speak. The consequences are not theoretical: perplexity-based safety filters incorrectly block 87.1% benign Bengali queries and 97.9% of benign code-switching queries, while intra-sentence code-switching simultaneously increases jailbreak attack success rates by 46.7% over monolingual English (Mustafa and Wu 2026). The safety infrastructure that fails multilingual users also creates the most exploitable vulnerabilities for adversarial ones. The paper uses a six-principle design framework for culturally grounded multilingual safety evaluation, and its central finding is that the existing benchmarks were not designed with Global Majority users in mind, and the absence of that intention is visible in the numbers.

Gamage and colleagues’ three-country participatory study—49 stakeholders across Sri Lanka, India and Bangladesh—documents what platform governance failure looks like on the ground. Workshop participants personally verified content in Sinhala, Bengali, Tamil, and Hindi as synthetic, reported it to platforms, and watched it remain online while comparable English-language content was removed. The paper explicitly names this as a political condition: platform governance “reproduces the power asymmetries of the global digital economy” (Gamage et al. 2026). The community knowledge infrastructure that has emerged in all three countries—journalist verification networks, feminist peer-support systems, civil society coalitions, each rooted in a specific national context—is evidence of what governance looks like when it is built by the affected communities rather than for them.

Arora, Miranda Mora, and Zarzycka analyze deepfake abuse in Mexico and the paper bridges this cluster and the one that follows. It introduces the concept of “gendered digital creative violence” to name something that existing harm taxonomies miss: that AI creative tools have produced a category of harm that is simultaneously technological, gendered, and manufactured—the harm is fabricated, designed, and deployed as a weapon against women’s bodies, reputations, and autonomy. Despite Mexico’s landmark Ley Olimpia⁹ reforms, approximately 95% of digital sexual violence investigations never reached the judicial stage, a gap the paper attributes to institutional failure rather than absence of law. Against this backdrop, the paper centers OlimpIA—a feminist AI chatbot deployed on WhatsApp offering psycho-emotional support, legal orientation, digital safety guidance, and referral to feminist networks—as counter-infrastructure built outside and against the architecture that has failed Mexican women. The paper’s named contribution, “feminist counter-creativity,” captures how the same generative capacities used to manufacture harm can be reoriented towards collective care and survivor support

9. Ley Olimpia is a comprehensive legal framework that criminalizes violations of sexual privacy in digital spaces, discussed in more detail in this issue’s article by Arora, Miranda Mora, and Zarzycka (2026).

(Arora, Miranda Mora, and Zarzycka 2026).

Language and decolonial threads run through all three papers in ways that connect directly to the call for proposals' named topics of "Cultural and Linguistic Diversity Online" and "Postcolonialism and the Digital World." Mustafa and Wu's (2026) benchmarks reveal a hierarchy of language in which safety evaluation resources are distributed in proportion to economic power, not context or vulnerability. Gamage et al.'s (2026) stakeholders in South Asia work in languages that the dominant moderation infrastructure cannot process. Arora, Miranda Mora, and Zarzycka's (2026) OlimpIA works in Spanish because the English-language tools built for Northern survivors do not serve Mexican women. Across all three papers, language is more than a technical parameter, it is a marker of whose experience the infrastructure was designed to serve.

3.3 Platform Governance—Who Governs What, and for Whom

While the first cluster asks who and what platform safety tools were built for, the second asks who built the institutions that govern them—and who those institutions serve. The two papers here take different methodological approaches: a quantitative analysis of all 147 published Meta Oversight Board decisions (Chala 2026) and a large-scale survey study of national digital identity adoption in Indonesia (Pachouri, Raj, and Harshavardhan 2026). They both make a coherent argument: formal platform governance institutions were designed for recognized actors, operating in recognized languages, governed by recognized categories of harm. Everyone else navigates the edges.

Chala's statistical analysis of 147 Meta Oversight Board decisions yields a counterintuitive headline: Global Majority cases form the largest share of the Board's docket (80 of 147), followed by Global North cases (45), and 22 cases spanning both. The Oversight Board is not, by this measure, ignoring the Global Majority. The more consequential inequality lies in procedural routing, not outcomes. English-language cases are significantly more likely to be resolved through "summary reversal"—Meta's self-correction without full panel review—while non-English cases disproportionately proceed to full deliberation. About 60 of 61 summary decisions are reversals, meaning Meta is more likely to self-correct on English-language content without Board intervention. Chala frames this as "institutional legitimacy as structured attention"—the Board distributes legitimacy not through outcome bias but through structural filters that determine which disputes receive what kind of institutional attention (Chala 2026). The finding that overturn rates among fully reviewed cases are not significantly different between English (64%) and non-English (73%) cases might initially suggest language parity—but this tells us only that once a case reaches full review, English and non-English speakers are treated relatively similarly. The prior question—which cases reach that stage, and why—is where the inequality lives.

Pachouri and colleagues' e-KTP study offers a different angle on the same question—and a partial counterpoint. While e-KTP has achieved 97% coverage among eligible

citizens, the paper's central concern is the gap between coverage and trusted use. Using survey data from 1,565 households and a structural equation model that explained 85% of variance in adoption attitude, they found that institutional trust was the dominant predictor ($\beta = 0.66$), far ahead of perceived ease of use or technical features. In addition, 99% of respondents reported that e-KTP improved their sense of social inclusion; among them, 67% attributed this to feeling more recognized and accepted. Last, their finding that 90% use simple show-of-card verification while only 36% have used fingerprint verification (and 28% face recognition) reveals a substantial gap between technical design and social practice (Pachouri, Raj, and Harshavardhan 2026).

The e-KTP study is the only paper in this issue not primarily about harms. It is about what national digital infrastructure can do when it works—and when it is built around the social context it serves rather than imported from elsewhere. As a counterpoint to the papers that document failure, Pachouri's contribution matters: it shows that the question is not whether states should build digital infrastructure in Majority World contexts, but how that infrastructure should be designed and governed. The postcolonial stakes of that question—who designs it, for whom, under what conditions—apply just as much to Indonesia's identity system as to Meta's content moderation infrastructure.

3.4 Gendered Harm and Queer Safety—What Platforms are Not Protecting

The third cluster is where this special issue's core theme of digital intersectionality is most directly illustrated. The harms documented in these papers are not the product of single identities or single systems: they emerge from the intersection of gender, sexuality, language, nationality, immigration status and political visibility. The call for papers named "intersectional approaches" as a priority topic area precisely because single-axis analyses cannot capture what happens to Quechua women on TikTok, to Iranian queer activists on X (formerly Twitter), or to exiled women human rights defenders navigating cybersecurity tools built for different adversaries. These papers provide the empirical answer to why intersectionality is not optional in Majority World trust and safety research.

Thakur's study of Quechua women content creators on TikTok and YouTube is the most granular analysis of intersectional online Gender-Based Violence in this issue. Automated analysis of 28,625 comments on six Quechua women influencer accounts found that 15.3% contained some form of Gender-Based Violence—with misogynistic threats the dominant category at 12.5%. That 30.5% comments were written in Quechua is itself a limitation on the model's completeness: the classifier was trained on Spanish and English, and Quechua has no equivalent commercial moderation infrastructure. The content moderators interviewed for the paper described the structural impossibility of their task: "In many cases, you cannot translate the same word with the term. There is no word for that same term". This is a downstream consequence of a translation hierarchy—policies designed in English, translated to Spanish, then imperfectly adapted into Quechua. Thakur's own conclusion is that "the platforms through their policies replicate colonial and patriarchal systems of power that subjugate women and specifically

Quechua women” (Thakur 2026). The same Quechua identity that inspires women to create content—and draws audiences to watch it—makes them more visible targets for the very abuse the platforms fail to address.

Hooman and Hodson’s analysis of Iranian queer activists on X, examines a different kind of intersectional harm: the use of platform architecture to amplify queerphobic violence rooted in authoritarian gender governance. Through interviews with five Iranian queer activists in the diaspora and a critical discourse analysis of 59 Persian-language hate speech tweets directed at them, the paper documents “platformized violence”—harassment that is structurally amplified through X’s engagement infrastructure, and rendered harder to address by moderation systems not designed for multilingual, politically restrictive contexts. Activists reported repeatedly filing complaints about explicit threats and receiving automated responses stating that the content did not violate guidelines; abusive accounts that were suspended returned under new usernames (Hooman and Hodson 2026). This points to what we might call a visibility paradox: exiled activists must be publicly visible to do their advocacy work, but that same visibility intensifies exposure to harm. The paper’s core argument is that X’s platform architecture functions as a structural extension of Iran’s authoritarian gender governance, not by design but by indifference: harassment that reproduces state-aligned heteronormative norms circulates freely within a system optimized for engagement.

Aljizawi and colleagues’ study of cybersecurity for exiled queer women human rights defenders takes the argument one level further: not just platform moderation failures, but the structural inadequacy of the cybersecurity field itself. Drawing on 18 interviews and seven written responses from exiled queer women activists from Iran, Syria, Azerbaijan, Morocco, China, and other countries, the paper documents a “protection gap”: none of the LGBTIQ+ respondents in their sample had accessed digital security training, compared to almost half of non-queer respondents in the broader dataset. None of the participants reported that available training addressed gender-specific risks—sexualized harassment, disinformation, or reputational attacks. Training was instead centered on account and device security: two-factor authentication, password hygiene, encrypted messaging, virtual private networks (VPNs). “A secure device doesn’t equal a safe person” is the paper’s most quotable formulation, and it captures the core misalignment: mainstream cybersecurity threat models are built around surveillance and account compromise, not state-sponsored campaigns designed to destroy someone’s reputation, relationships, and community standing. The paper’s call to “queer cybersecurity”—to move beyond device protection and address identity-based vulnerabilities shaped by gender, sexuality, immigration status, and political visibility—is a call for a different threat model, one that begins with the person rather than the device (Aljizawi et al. 2026).

The language thread runs through all three papers in a specific register: not just as a barrier to platform governance, but as a dimension of harm itself. Quechua women are targeted partly because Quechua is a language that automated moderation systems

cannot read. Iranian queer activists are exposed partly because Persian-language content is under-moderated. Exiled women human rights defenders need multilingual security training and multilingual platform reporting channels that do not yet exist widely. In each case, language is both a vector of vulnerability and a marker of who the platforms were not designed to protect.

4 A Note on Where We Stand

Our call for papers required abstract submitters to include a positionality statement: a condition we imposed because we believe where scholarship is produced and by whom, shapes what it sees and what it misses. It would be inconsistent for us, as guest editors, to impose that requirement on contributors without meeting it ourselves. We therefore offer the following as the kind of honest self-location that we asked of the researchers who submitted to this issue.

We are a diverse and intersectional group representing technology platforms and academia, united by a commitment to advocating for marginalized digital communities through inclusive and equitable ecosystem practices, and an affiliation to the TSF Research Coalition's GMRC. All three of us are currently based at institutions in what this issue calls the Global North or Northern-proximate contexts, but our origin stories are rooted in Global Majority experiences of platform use and academic training. Additional proximity to Global Majority trust and safety research comes through professional focus, research expertise, and deliberate institutional commitment.

We also collectively embody the academic-industry intersectionality named in Section 1 as a defining feature of the trust and safety field. The *Journal of Online Trust and Safety* and the Trust and Safety Research Conference exist precisely because neither academia nor industry can understand platform safety alone: research without operational grounding produces findings that do not survive contact with platform realities; practice without research produces interventions whose effects are unknown and whose harms can be invisible. Our editorial team was constituted at this crossing. One of us is primarily an industry practitioner with experience conducting user experience research within platform environments, bringing knowledge of how research findings are translated (and often fail to be translated) into product decisions at scale. The other two are primarily academics who have worked on research projects in close collaboration with platforms, bringing methodological rigor and critical independence alongside operational familiarity. None of us occupies only one side of this intersection, and we believe that is why this collaboration produced an editorial process, and a set of published papers, that is accountable to both communities of knowledge and practice.

Sujata Mukherjee is Senior Director of Trust and Safety Product Management at Upwork and co-chairs the GMRC at the TSF. Her vantage point is that of a practitioner:

twenty years of building trusted product experiences at large technology organizations—including Google, IBM and Genpact—inform her perspective on how platform trust and safety infrastructure is actually designed and governed. This includes direct experience conducting user experience research within platform environments, which has shaped her understanding of both the potential and the limits of research as an input to product decision-making. She was born and raised in India, and has been based in the United States for the last six years, a reflection of the same kind of geographical mobility that Global Majority researchers are expected to demonstrate to obtain institutional legitimacy and voice.

Mona Elswah is Lecturer in Digital Media Governance at the University of Exeter and co-chair of the GMRC at the TSF. She is also the Director of the Digital Majority Centre at the University of Exeter. Her research focuses on content moderation, digital repression, and algorithmic bias in non-Western contexts. She has worked specifically on the content moderation of low-resource languages—as a research fellow at the Center for Democracy and Technology, as a research associate at the Oxford Internet Institute, and as a research fellow at Harvard University’s Carr Center for Human Rights Policy. Her expertise bears directly on the structural problems this special issue addresses. Raised in Egypt, she is currently based in the United Kingdom.

Crystal Abidin is Professor of Internet Studies, Founding Director of the Influencer Ethnography Research Lab, and Founder of the TikTok Cultures Research Network at Curtin University. A world-leading ethnographer of influencer cultures, her practice of traditional and digital anthropology focuses on vernacular internet cultures and digital media in Asia Pacific. She brings localized geographic expertise on one of the world’s largest and most complex Majority World digital ecosystems and a methodological grounding in community-enabled ethnographic research that shapes how this editorial reads the work of the researchers in this issue. Crystal has worked closely as consultant and advisor for the social media industry and for governments, with expertise on creator cultures, content regulations, and platform governance—experience that informs her understanding of where academic and industry knowledge converge and where they remain productively in tension.

We acknowledge that throughout the editorial process, all three of us were based at Northern or Northern-proximate institutions: we were editing a special issue about Majority World research from positions shaped by the structural advantages of our institutional affiliations. We tried to take that seriously in the ways described in Section 2.2. We note it here because transparency about editorial positionality is not merely a scholarly convention; it is, as several scholars have argued, an epistemic and ethical requirement.

References

- Alam, Shahidul. 2008. "Majority World: Challenging the West's Rhetoric of Democracy." *Amerasia Journal* 34 (1): 88–98. <https://doi.org/10.17953/amer.34.1.l3176027k4q614v5>.
- Aljizawi, Noura, Shaila Baran, Marcus Michaelson, and Siena Anstis. 2026. "Security without Safety: Queering Cybersecurity in the Age of Digital Transnational Repression." *Journal of Online Trust and Safety* 3 (3). <https://doi.org/10.54501/jots.v3i3.321>.
- Arora, Payal, Ana María Miranda Mora, and Marta Zarzycka. 2026. "Deepfake Abuse and Gendered Digital Creative Violence: Feminist AI Interventions from Mexico." *Journal of Online Trust and Safety* 3 (3). <https://doi.org/10.54501/jots.v3i3.316>.
- Atlantic Council. 2023. "Scaling Trust on the Web." Atlantic Council, June. Accessed March 2, 2025. <https://www.atlanticcouncil.org/in-depth-research-reports/report/scaling-trust/>.
- Campbell-Stephens, Rosemary M. 2021. *Educational Leadership and the Global Majority: Decolonising Narratives*. Cham: Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-88282-2>.
- Chala, Endalkachew H. 2026. "Platform Oversight in Practice: How Language Shapes Procedure, Not Outcome, in Meta's Oversight Board." *Journal of Online Trust and Safety* 3 (3). <https://doi.org/10.54501/jots.v3i3.333>.
- Crenshaw, Kimberlé. 1989. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics." *University of Chicago Legal Forum* 1989 (1). <https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>.
- Díaz, Ángel, and Laura Hecht-Felella. 2021. *Double Standards in Social Media Content Moderation*. Brennan Center for Justice. <https://www.brennancenter.org/our-work/research-reports/double-standards-social-media-content-moderation>.
- Elsawah, Mona, Aliya Bhatia, and Dhanaraj Thakur. 2025. *Content Moderation in the Global South: A Comparative Study of Four Low-Resource Languages*. Center for Democracy & Technology. <https://cdt.org/insights/content-moderation-in-the-global-south-a-comparative-study-of-four-low-resource-languages/>.
- Fang, Kecheng. 2024. "Wangbao (Cyberbullying) and Jubao (Reporting): Collaborative State-Society Online Influence Operations in China." *Journal of Online Trust and Safety* 2 (3). <https://doi.org/10.54501/jots.v2i3.167>.
- Freelon, Deen, Cristina Monzer, Gayoung Jeon, Cameron Moy, and Natasha Williams. 2024. "The Post-API Age of Social Media Data Access: Past, Present, and Future." *The ANNALS of the American Academy of Political and Social Science* 715 (1): 16–37. <https://doi.org/10.1177/00027162251372557>.

- Gamage, Dilrukshi, Dilki Sewwandi, Shreeti Shubham, Sumaia Arefin, Dilaxshini Dunstan, Kartik Joshi, and Rashmi Gunawardana. 2026. "Deepfakes in the Global South: Understanding Stakeholder Perceptions, Harms, and Governance Challenges in Sri Lanka, India, and Bangladesh." *Journal of Online Trust and Safety* 3 (3). <https://doi.org/10.54501/jots.v3i3.319>.
- Hooman, Niloofar, and Jaigris Hodson. 2026. "Shaping Queerphobia: How X Amplifies Hate Speech against Iranian Queer Activists." *Journal of Online Trust and Safety* 3 (3). <https://doi.org/10.54501/jots.v3i3.313>.
- Houtti, Mohamed, Abhishek Roy, Venkata Narsi Reddy Gangula, and Ashley Marie Walker. 2024. "A Survey of Scam Exposure, Victimization, Types, Vectors, and Reporting in 12 Countries." *Journal of Online Trust and Safety* 2 (4). <https://doi.org/10.54501/jots.v2i4.204>.
- Mimizuka, Kayo, Megan A. Brown, Kai-Cheng Yang, and Josephine Lukito. 2025. "Post-Post-API Age: Studying Digital Platforms in Scant Data Access Times." *arXiv*, <https://doi.org/10.48550/arXiv.2505.09877>.
- Mohamed, Shakir, Marie-Therese Png, and William Isaac. 2020. "Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence." *Philosophy & Technology* 33 (4): 659–84. <https://doi.org/10.1007/s13347-020-00405-8>.
- Mukherjee, Sujata. 2026. "T&S and the Majority World." In *Trust, Safety, and the Internet We Share: Multistakeholder Insights*, edited by Maia Levy Daniel, Amanda Menking, Marlyn Savio, and Jean Claffey, 155–66. New York: Routledge. <https://doi.org/10.1201/9781003621072-19>.
- Mukherjee, Sujata, and Sasha Maria Mathew. 2026. "Digital Sovereignty Means Breaking the Western Monopoly on AI Meaning." Tech Policy Press, April 28, 2026. <https://www.techpolicy.press/digital-sovereignty-means-breaking-the-western-monopoly-on-ai-meaning/>.
- Mustafa, Alisar, and Cherry Wu. 2026. "Evaluating Multilingual Safety Benchmarks for Low-Resource Languages in the Majority World." *Journal of Online Trust and Safety* 3 (3). <https://doi.org/10.54501/jots.v3i3.330>.
- Nakamura, Gabriel, Bruno Eleres Soares, Valério D. Pillar, José Alexandre Felizola Diniz-Filho, and Leandro Duarte. 2023. "Three Pathways to Better Recognize the Expertise of Global South Researchers." *npj Biodiversity* 2 (1). <https://doi.org/10.1038/s44185-023-00021-7>.
- Nicholas, Gabriel, and Aliya Bhatia. 2023. "Toward Better Automated Content Moderation in Low-Resource Languages." *Journal of Online Trust and Safety* 2 (1). <https://doi.org/10.54501/jots.v2i1.150>.

- Nigatu, Hellina Hailu, Farhana Shahid, Vishal Sharma, Abigail Oppong, Michaelanne Thomas, and Syed Ishtiaque Ahmed. 2026. "UnWEIRDing Peer Review in Human-Computer Interaction." In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*, 1–17. <https://doi.org/10.1145/3772318.3791553>.
- Pachouri, Anshul, Abhishek Raj, and Kushagra Harshavardhan. 2026. "Acceptance and Use of the National Digital Identity System in Indonesia." *Journal of Online Trust and Safety* 3 (3). <https://doi.org/10.54501/jots.v3i3.323>.
- Sambasivan, Nithya, Erin Arnesen, Ben Hutchinson, Tulsee Doshi, and Vinodkumar Prabhakaran. 2021. "Re-imagining Algorithmic Fairness in India and Beyond." In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 315–28. <https://doi.org/10.1145/3442188.3445896>.
- Santini, Rose Marie, Hugo Leal, Adriano Belisario, and Bruno Mattos. 2026. "Unequal Data Access Regimes in Social Media." *Political Communication Report*, no. 33 (Summer). <https://doi.org/10.25358/openscience-15829>.
- Thakur, Dhanaraj. 2026. "'Our Content Should Be Treated Differently': Race, Gender, Language, and the Experiences of Quechua Women on Social Media." *Journal of Online Trust and Safety* 3 (3). <https://doi.org/10.54501/jots.v3i3.332>.
- Townsend, Beverley, and Arthur Gwagwa. 2023. "Authoritarian Alliances and the Politicking of Data in Africa." *Journal of Online Trust and Safety* 2 (1). <https://doi.org/10.54501/jots.v2i1.111>.
- Umbach, Rebecca, Anubha Singh, and Ashley Marie Walker. 2023. "'Your Protection is in Your Hands Only': User Awareness and Adoption of Privacy and Security Practices in Five Majority World Countries." *Journal of Online Trust and Safety* 2 (1). <https://doi.org/10.54501/jots.v2i1.114>.

Authors

Sujata Mukherjee is Senior Director of Trust and Safety Product Management at Upwork and Co-Chair of the Trust and Safety Foundation's Global Majority Research Committee. Email: tiamukherjee@gmail.com.

Mona Elswah is Lecturer in Digital Media Governance at the University of Exeter and Co-Chair of the Trust and Safety Foundation's Global Majority Research Committee. Email: m.a.i.elswah@exeter.ac.uk.

Crystal Abidin is Professor of Internet Studies and Founding Director of the Influencer Ethnography Research Lab at Curtin University, and a Member of the Trust and Safety Foundation's Global Majority Research Committee. Email: crystal-abidin@gmail.com.

Acknowledgments

We are grateful to the editorial team at the *Journal of Online Trust and Safety* for their support throughout this special issue. We worked most closely with managing editor Rosie Ith and co-editor Ronald Robertson, whose guidance and patience shaped the issue at every stage from call to publication, and thank them for their institutional stewardship of a journal that takes this kind of work seriously.

This issue would not exist without the rigorous and generous work of its peer reviewers. We thank Amanda Menking, Carolina Are, Fadza Madzingira, Hanne Stegeman, Irene Huang, Joseph Seering, Juliet Shen, Marlyn Savio, Nermine Aboulez and Zahraa Badr for their careful and constructive engagement with the manuscripts. We recruited this group specifically for their positional proximity to Global Majority trust and safety research—their assessments were strengthened by contextual knowledge that cannot be imported from outside. Their expertise strengthened every paper in this issue.

We are also grateful to all 59 author teams who submitted abstracts to this call. We could not name them all here, and the constraints of the editorial process mean that most of their work does not appear in these pages—but the ambition, range, and generosity of that submission pool shaped everything that followed. These are the researchers keeping this discourse alive in contexts where it is hardest to sustain, and this issue would have been thinner if they had not shown us where the field is and where it needs to go.

Finally, we thank the TSF Research Coalition for providing the institutional scaffolding for the Global Majority Research Committee, where the idea for this special issue was born. The Coalition's work, and the community it has built, made this issue possible.

Keywords

Majority World; Global Majority; digital intersectionality; marginalization; academic-industry divide.