Towards a Systemic Risk Literacy for Tech-Facilitated Gender-Based Violence

Ashley Marie Walker
Google
New York City, USA
amwalker@google.com

Renee Shelby
Google Research
San Francisco, USA
reneeshelby@google.com

Rosanna Bellini
Cornell Tech
New York, USA
rfb242@cornell.edu

Amelia Hassoun
University of Cambridge
Cambridge, United Kingdom
ah2229@cam.ac.uk

Emily Tseng
Microsoft Research
New York, New York, USA
University of Washington
Seattle, WA, USA
etseng42@gmail.com

Abstract

Regulators worldwide increasingly seek to mitigate the harms associated with online spaces. To comply with these regulations, platforms are being required to perform systemic risk assessments around online content and conduct for the first time. Doing so requires bridging foundational research understanding online harm with practical knowledge of how platforms and policymakers can assess the risk of harm in practice: what we call systemic risk literacy. We argue that systemic risk literacy is under-studied — hampering our ability to create the effective and data-driven risk assessments that we need to ensure digital safety. Tech-facilitated gender-based violence (TFGBV) is a valuable starting point for creating systemic risk literacies: the harms of TFGBV are urgent, and there already exist robust ecosystems of science and legislation around TFGBV worldwide. In this workshop, we will work collaboratively to: (1) highlight the sociotechnical dynamics that contribute to TFGBV; (2) map potential interventions; and (3) suggest evaluations to understand whether these mitigations are effective.

CCS Concepts

• **Security and privacy → Human and societal aspects of security and privacy**; • **Human-centered computing → HCI theory, concepts and models**.

Keywords

risk literacy, digital safety, risks, harms, tech-facilitated gender-based violence

ACM Reference Format:
Ashley Marie Walker, Renee Shelby, Rosanna Bellini, Amelia Hassoun, and Emily Tseng. 2025. Towards a Systemic Risk Literacy for Tech-Facilitated Gender-Based Violence. In *Companion of the Computer-Supported Cooperative Work and Social Computing (CSCW Companion '25), October 18–22, 2025, Bergen, Norway*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3715070.3748298

1 Motivation

Recent global regulations aim to mitigate harms in online spaces [4, 11, 21], requiring platforms to conduct *systemic risk assessments* around online content and conduct for the first time [13]. While the concept of systemic risk assessments is well-established in other regulated sectors, their application to online platforms is still in its nascent stages [6]. As seen with legislation in other other highly technical fields (e.g., automotive, aviation, agriculture [12]), this regulatory attention necessitates a negotiation of core questions regarding these sociotechnical systems. Key among these are: What are the specific sociotechnical dynamics that contribute to harm? What controls are necessary to mitigate this potential for harm, and how can we evaluate whether the mitigations are effective? Ultimately, improving digital safety for end users requires a clear understanding of sociotechnical dynamics that contribute to online harms, the interventions possible to limit these dynamics, and rigorous, evidence-backed methods for stakeholder groups—including industry, academia, civil society and regulators—to assess the risk of online platforms and systems.

Currently, there is a gap between foundational research describing the dynamics underpinning online harms and the practical implementation of risk assessments by platforms and legislators. Developing an evidence-backed understanding of the risk landscape, or **risk literacy**, is a key step for mitigating harms to end users. Risk literacy encompasses the ability to understand, assess, and make informed decisions about risks, their tradeoffs, and their implications [16]. However, previous work in technical risk literacy has identified common problems, including incorrect mental models of how systems function [7], disconnects between knowledge of risks and necessary protective actions [20], differing understandings of technical systems across stakeholder groups [18], and knowledge voids as emerging technologies evolve [17]. To bridge the gap between foundational research and the practical requirements of current digital safety legislation, developing systemic risk literacies for online harms is essential. Without these risk literacies, accurately assessing the risk of harmful sociotechnical dynamics becomes impossible.

CSCW Companion '25, Bergen, Norway

ACM ISBN 979-8-4007-1480-1/25/10
https://doi.org/10.1145/3715070.3748298

Tech-facilitated gender-based violence (TFGBV) encompasses a wide range of ways technical systems can perpetuate harms associated with systemic inequality [1, 8]. TFGBV manifests in forms such as sextortion, image-based sexual abuse, hate speech, impersonation, cyberstalking and more [2]. TFGBV is a comparatively well-developed area with a substantial body of published evidence in recent years (e.g., [3, 5, 9, 10, 15, 22]). Given this breadth of existing research, its specific mention in legislation (e.g. the Digital Services Act in the EU [19]), and the established ecosystem of support from scholars, civil society organizations, industry, and regulators, TFGBV serves as a valuable test case for iterating on rigorous, evidence-backed risk evaluations. As a primary step, the community must develop a systemic risk literacy for TFGBV, including identifying current constraints limiting progress towards this goal.

Our overarching goal is to foster a community of TFGBV experts dedicated to bridging the gap between the rich body of TFGBV research and the practice of risk assessments, which are now integral to legislative practice. In this workshop, we will collaboratively:

(1) Highlight the key dynamics in sociotechnical systems that contribute to tech-facilitated gender-based violence.
(2) Develop a map of potential interventions and best practices to address these dynamics
(3) Suggest key measures and metrics to assess the effectiveness of these mitigations.

As such, we propose a one day, in-person workshop with approximately 40 participants. CSCW 2025 presents an ideal opportunity for this workshop due to its proximity to European stakeholders who have experience with regulations already in effect, and CSCW's history of supporting digital safety research, particularly concerning at-risk users.

2 Organizers

Our organizing team represents expertise gained through experience within and across various TFGBV stakeholder groups, including academia, civil society organizations, and industry. Additionally, the organizers have successfully managed previous workshops on sensitive topics, including at CSCW 2024 [14, 23].

- **Rosanna Bellini, New York University**: Dr. Bellini is an Assistant Professor. She specializes in mitigating technology-enabled abuse in intimate partner violence contexts by working directly with abusive partners and around instances of financial abuse.
- **Amelia Hassoun, Google**: Dr. Hassoun is a research scientist working in security and anti-abuse research, particularly focused on on supporting help-seeking for technology-facilitated abuse in the era of generative AI.
- **Ashley Marie Walker, Google**: Dr. Walker is a researcher on the Trust & Safety team within Google. Their work focuses on understanding the systems of marginalization and subsequent risks in sociotechnical systems and how empirically-informed policy design can help with mitigation.
- **Renee Shelby, Google Research:** Dr. Shelby studies the social impacts of technology, particularly the relationships between social inequality, technology design, and governance. Her work identifies potential harms and modes of harm reduction from algorithmic systems by prioritizing the experiences of historically marginalized communities.
- **Emily Tseng, Microsoft Research & University of Washington**: Dr. Tseng's research examines how computing comes to mediate harm, how to intervene, and what it means to do so. Trained as a scientist-advocate for survivors of gender-based violence, her current interests include psychological safety in generative AI and AI red-teaming, and building participatory LLMs with and for journalists.

3 Workshop Schedule

We workshop schedule is designed to achieve two primary goals: (1) prioritizing the conversation that feels most congruent with the set of people in the room; and (2) working towards a concrete artifact from the day's conversations. Given the significant harms associated with TFGBV and its increasing regulatory attention, any progress towards concrete tools can inform ongoing policy discussions. This event aims to establish the design space for systemic risk literacy in TFGBV and foster building community.

Morning: Overview of TFGBV and Agenda Setting

The morning session will build a shared understanding of the problem space of TFGBV risk assessments, and facilitate relationship-building among participants for the afternoon discussion sections.

- *Session 1: (30 minutes) Introductions and Agenda Setting*
  This session will be dedicated to day-of room logistics, context-setting for the workshop, and goals for the day's work. Organizers will provide a high-level overview of the TFGBV landscape, the regulatory environment, and existing systemic risk assessment practices. This overview will serve to ground and guide discussions throughout the rest of the day.
- *Session 2: (45 minutes) 1:1 Relationship Building*
  One of the major benefits of workshops at conferences is the ability to build communities of practice around shared area of expertise. To facilitate building relationships and making connections beyond existing networks, we will have a session where participants have 8 minutes to introduce themselves to each other, discussing research interests and goals for the workshop, before rotating on to a new conversational pairing. In previous workshops on sensitive topics [23], these sessions have been called out by attendees as valuable for establishing relationships with a wide range of colleagues before diving into difficult conversations.
- *Break (15 minutes)*
- *Session 3: (60 min) Large-group Discussion*
  Given the broad range of TFGBV harms and the novelty of related legislation, this session will prioritize discussing key affordances facilitating TFGBV, eliciting best practices for interventions encountered by participants, and surfacing technical constraints affecting intervention effectiveness. The aim is to map the design space for systemic risk literacy in TFGBV, ensuring afternoon sessions focus on practical considerations.
- *Session 4: (30 min) Discussion Elicitation*
  Building on the morning's context setting and group discussion, this session will elicit proposed discussion topics.

Attendees will use sticky notes to suggest questions for collaborative iteration in the afternoon. Organizers will then thematically group these questions to form the foundation for afternoon discussions.

Lunch Break

Afternoon: Design and Problem Solving

The afternoon session will center small-group discussions to dive deep into design spaces raised by the morning discussion elicitation session. Based on prior workshop experiences, likely topics include: evaluating aspects of current risk assessment processes, identifying opportunities for improvement, recommending evidence-backed best practices, and developing a research agenda for what knowledge is missing to make those recommendations.

- *Session 4: (45 min) Discussion Section 1*
  For this workshop, we are choosing to prioritize small group discussions to take advantage of the opportunity for in-person engagement across institutions and stakeholder groups. Each discussion group will include a representative from the organizing team to assist with facilitation, note taking, and timekeeping.
- *Break (15 min)*
- *Session 5: (45 min) Discussion Section 2*
  The day-long format allows for two afternoon discussion sessions. Participants can choose to remain with their first discussion topic or rotate to another group. This arrangement allows attendees to either dive deep into their highest priority topic or cover multiple discussions that spark their interest. The facilitator for each group will remain consistent across both sessions to ensure continuity in note-taking and thematic connections.
- *Session 6: (30 min) Report Out*
  As the day draws to a close, this session presents the opportunity for facilitators to report out the major discussion points from the afternoon sessions. Following the format of previous workshops, this allows for large-group connections between discussion threads before transitioning into next steps.
- *Session 7: (30 min) Next Steps*
  The closing session of the day will focus on concrete next steps for the group. Given the workshop's specific intent, this discussion will center on logistical organizing for writing up outcomes and what future scaffolding for community building might be necessary to continue to work towards systemic risk literacy for TFGBV.

3.1 Recruitment Plan

Our goal for this workshop is to cultivate a community of subject matter experts working across the range of contexts that inform TFGBV decision-making, including academia, industry and civil society. Having a range of expertise represented in our workshop will ensure that the day-of discussions are grounded in actionable levers of change and limitations of how sociotechnical systems are beginning to be regulated. To reach this range of potential participants, we will recruit from the organizers' professional networks (including social media followings) and relevant topic-area list-servs. We will take advantage of CSCW's international location this year to proactively reach out to researchers and civil society organizations based in the EU.

3.2 Submission Details

Potential workshop participants will be asked to submit a short position paper (or submission in an alternative format demonstrating comparable effort) addressing one of the following guiding questions:

- Based on your experience, what are the key aspects of sociotechnical systems contributing to tech-facilitated gender-based violence? What appropriate controls could mitigate this dynamic? What tradeoffs might arise from these controls, including their influence on other harms beyond TFGBV?
- What examples exist of effective interventions for TFGBV? How were these interventions designed? How do we know they were impactful? Were there any unintended consequences?
- Thinking at a system level, what metrics and indicators are necessary to understand whether a system is operating within what we would consider "safe bounds" when it comes to the risk of TFGBV? For a given sociotechnical system, what data would be necessary to evaluate the relative safety of system in the context of TFGBV?
- What research and/or data are we missing for developing a systemic risk literacy for TFGBV? What future research directions are of highest priority? Where do we feel like we have sufficient evidence? What is necessary to make the translation between foundational research and effective real world implementation of protections?

Submissions will be evaluated on how well they demonstrate potential participant's familiarity and ability to contribute to discussions on the topic area. We welcome submissions up to 750 words, a five-minute slide presentation, an infographic, design fiction, or other formats representing a similar level of effort. This range of options aims to encourage submissions from participants who might otherwise find standard workshop submissions prohibitive and to encourage potential participants to experiment with new communication formats.

Due to the sensitive nature of the topic area and anticipated discussions, we will not host participant submissions publicly, and discussions will adhere to Chatham House Rules. After the event, we will circulate a high-level overview of major discussion points without individual attribution. At least one author of accepted submissions must attend the workshop, and all participants will register for the workshop and attend at least one day of the conference.

References

[1] Kristine Baekgaard. 2024. *Technology-Facilitated Gender-Based Violence: An Emerging Issue in Women, Peace and Security.* Technical Report. Georgetown Institute for Women, Peace and Security.

[2] Jane Bailey and Jacquie Burkell. 2021. Tech-facilitated violence: thinking structurally and intersectionally. *Journal of Gender-Based Violence* 5, 3 (2021), 531–542.

[3] Christine Barter and Sanna Koulu. 2021. Digital technologies and gender-based violence–mechanisms for oppression, activism and recovery. *Journal of gender-based violence* 5, 3 (2021), 367–375.

[4] European Commission. 2021. The Artificial Intelligence Act. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206

[5] Molly Dragiewicz, Jean Burgess, Ariadna Matamoros-Fernández, Michael Salter, Nicolas P. Suzor, Delanie Woodlock, and Bridget Harris. 2018. Technology facilitated coercive control: domestic violence and the competing roles of digital media platforms. *Feminist Media Studies* 18, 4 (2018), 609–625. doi:10.1080/14680777.2018.1447341

[6] Niklas Eder. 2024. Making systemic risk assessments work: how the DSA creates a virtuous loop to address the societal harms of content moderation. *German Law Journal* 25, 7 (2024), 1197–1218.

[7] Andrew J. Flanagin and Miriam J Metzger. 2000. Perceptions of Internet information credibility. *Journalism & mass communication quarterly* 77, 3 (2000), 515–540.

[8] Nicola Henry, Asher Flynn, and Anastasia Powell. 2020. Technology-Facilitated Domestic and Sexual Violence: A Review. *Violence Against Women* 26, 15-16 (2020), 1828–1854. doi:10.1177/1077801219875821 arXiv:https://doi.org/10.1177/1077801219875821 PMID: 32998673.

[9] Nicola Henry and Anastasia Powell. 2015. Embodied Harms: Gender, Shame, and Technology-Facilitated Sexual Violence. *Violence Against Women* 21, 6 (2015), 758–779. doi:10.1177/1077801215576581 arXiv:https://doi.org/10.1177/1077801215576581 PMID: 25827609.

[10] Nicola Henry and Anastasia Powell. 2018. Technology-Facilitated Sexual Violence: A Literature Review of Empirical Research. *Trauma, Violence, & Abuse* 19, 2 (2018), 195–208. doi:10.1177/1524838016650189 arXiv:https://doi.org/10.1177/1524838016650189 PMID: 27311818.

[11] The White House. 2022. Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People. https://www.whitehouse.gov/ostp/ai-bill-of-rights/

[12] Nancy G. Leveson. 2016. *Engineering a Safer World: Systems Thinking Applied to Safety.* The MIT Press, Boston, MA.

[13] Anna Liesenfeld. 2024. The Legal Significance of Independent Research based on Article 40 DSA for the Management of Systemic Risks in the Digital Services Act. *European Journal of Risk Regulation* 16, 1 (2024), 184–196.

[14] Benedetta Lusi, Adrian K. Petterson, Kamala Payyapilly Thiruvenkatanathan, Michaela Krawczyk, Emily Tseng, Lara Reime, Madeline Balaam, Katie A. Siek, and Cristina Zaga. 2024. Caring for Reproductive Justice: Design in Response to Adversity. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing* (San Jose, Costa Rica) *(CSCW Companion '24)*. Association for Computing Machinery, New York, NY, USA, 693–696. doi:10.1145/3678884.3681832

[15] Clare McGlynn, Erika Rackley, and Rosie Houghton. 2017. Beyond 'Revenge Porn': The Continuum of Image-Based Sexual Abuse. *Fem Leg Stud* 25 (2017), 25–46. doi:10.1007/s10691-017-9343-2

[16] Davy Tsz Kit Ng, Jac Ka Lok Leung, Samuel Kai Wah Chu, and Maggie Shen Qiao. 2021. Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence* 2 (2021), 100041.

[17] Maribeth Rauh, Nahema Marchal, Arianna Manzini, Lisa Anne Hendricks, Ramona Comanescu, Canfer Akbulut, Tom Stepleton, Juan Mateos-Garcia, Stevie Bergman, Jackie Kay, Conor Griffin, Ben Bariach, Iason Gabriel, Verena Rieser, William Isaac, and Laura Weidinger. 2024. Gaps in the Safety Evaluation of Generative AI. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* 7, 1 (Oct. 2024), 1200–1217. doi:10.1609/aies.v7i1.31717

[18] Elissa M. Redmiles, Sean Kross, and Michelle L. Mazurek. 2017. Where is the Digital Divide? A Survey of Security, Privacy, and Socioeconomics. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 931–936. doi:10.1145/3025453.3025673

[19] Elisabetta Stringhi. 2024. The due diligence obligations of the Digital Services Act: a new take on tackling cyber-violence in the EU? *International Review of Law, Computers & Technology* 38, 2 (2024), 215–229.

[20] Monika Taddicken. 2014. The 'privacy paradox'in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of computer-mediated communication* 19, 2 (2014), 248–273.

[21] Ioanna Tourkochoriti. 2023. The digital services act and the EU as the global regulator of the internet. *Chi. J. Int'l L.* 24 (2023), 129.

[22] Emily Tseng, Rosanna Bellini, Nora McDonald, Matan Danos, Rachel Greenstadt, Damon McCoy, Nicola Dell, and Thomas Ristenpart. 2020. The Tools and Tactics Used in Intimate Partner Surveillance: An Analysis of Online Infidelity Forums. In *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, Boston, MA, 1893–1909. https://www.usenix.org/conference/usenixsecurity20/presentation/tseng

[23] Ashley Marie Walker, Michael Ann DeVito, Karla Badillo-Urquiola, Rosanna Bellini, Stevie Chancellor, Jessica L. Feuston, Kathryn Henne, Patrick Gage Kelley, Shalaleh Rismani, Renee Shelby, and Renwen Zhang. 2024. "What is Safety?": Building Bridges Across Approaches to Digital Risks and Harms. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing* (San Jose, Costa Rica) *(CSCW Companion '24)*. Association for Computing Machinery, New York, NY, USA, 736–739. doi:10.1145/3678884.3681824