

Expression of interest for the CSCW Workshop ‘Responsibly Training Foundation Models’

SILVIA CAZACU, KU Leuven, Belgium

ACM Reference Format:

Silvia Cazacu. 2025. Expression of interest for the CSCW Workshop ‘Responsibly Training Foundation Models’. 1, 1 (September 2025), 2 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

I would like to participate in the CSCW 2025 workshop ‘Responsibly Training Foundation Models’ because I aim to apply participatory design methods to reimagine how equitable AI practices might look like. I consider dataset curation to be one of the most critical sites where decisions about AI take form: whose data is scraped, who decides what counts as valuable, and who undertakes the labour of curation are all questions that reveal embedded hierarchies of power. I believe this workshop will allow me to situate these questions in the specific context of foundation models, where the unprecedented scale of data makes the stakes of these decisions both more diffuse and more urgent.

In my current work, I apply critical and feminist theory to understand how ethical principles such as social equity can be enacted when working with data in participatory ways. My recent CHI article on participatory data physicalisation [1] demonstrates that even when designing transparent and accessible methods of engaging in data practices, power dynamics persist across planning, deployment, construction, and dissemination. These dynamics manifest in how participants are involved, whether their meanings are accurately represented, and how much agency they retain over the data artefacts they help to construct. This finding foregrounds a crucial tension: ethical principles such as inclusiveness, accountability, and privacy are rarely applied in isolation, and when enacted together they can conflict.

Building on this work, I now want to critically examine data processes at the scale of foundation models. The central question I hope to explore is what participatory methods can look like in contexts where sheer scale renders traditional practices unfeasible, and how we might design processes that equitably balance power while remaining attentive to the inevitability of tension among ethical principles. My methodological reflections on data physicalisation offer a micro-scale exemplar: a simplified participatory data pipeline that already revealed the difficulty of adhering to ethical principles without friction. I am eager to test whether such reflections can inform the more complex terrain of foundation model data curation.

I believe this workshop offers an invaluable opportunity to discuss these ideas with an interdisciplinary community of scholars who approach data and AI equity from diverse perspectives. I hope to learn how other disciplines conceptualise data processes, how they confront the challenges of scale and responsibility, and how together we might begin to articulate methodological approaches that make space for pluralism, contestation, and care in dataset curation.

Author’s Contact Information: [Silvia Cazacu](#), silvia.cazacu-bucica@kuleuven.be, KU Leuven, Leuven, Belgium.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Manuscript submitted to ACM

Manuscript submitted to ACM

References

- [1] Silvia Cazacu, Georgia Panagiotidou, Therese Steenberghen, and Andrew Vande Moere. 2025. Disentangling the Power Dynamics in Participatory Data Physicalisation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 390, 19 pages. doi:10.1145/3706598.3713703