'Data Care: A Humanities and Social Sciences Approach to Debiasing Large Language Models'

Guest Editors
Payal Arora, Professor of Inclusive AI Cultures, Utrecht University
p.arora@uu.nl

Toija Cinque, Associate Professor, Communications (Digital Media), Deakin University
toija.cinque@deakin.edu.au

Baohua Zhou, Professor and Associate Dean Director of the New Media Communication Program, Founding Director of Computational and AI Communication Research Center, Fudan University
zhoubaohua@yeah.net

Large Language Models (LLMs) are AI systems trained on vast datasets to generate, understand, and process human language and expression, enabling applications like chatbots, translation, and content creation. Much research on LLMs is led by computer scientists focused on debiasing data to build fairer models, while humanities and social science scholars remain underrepresented in shaping AI's decision-making processes. Computational research often addresses 'data loss' and 'data deficiencies' through data-centric AI approaches, whereas scholars from media studies, anthropology, and political science for instance take a social-centric approach to critique AI's role in reinforcing historical inequalities through data extraction and algorithmic governance. These critiques, however, while important, rarely translate into co-constructive decision-making to build datasets that are transparent, equitable, and representative of the Majority World.

Emergent research from the Global South highlights AI's potential to challenge traditional gatekeepers, oppressive regimes, and patriarchal norms, fostering a more hopeful perspective on LLM-powered innovations. The rise of diverse LLMs—such as OpenAI's GPT (USA), DeepSeek, Qwen (China), Mistral (France), and Matilda (Australia)—demands a cross-cultural approach to AI development. These models reflect different linguistic, ethical, and socio-political contexts, underscoring the need for a humanities and social sciences analysis around what constitutes localized training data, multilingual adaptability, and culturally aware governance that moves beyond resistance toward rational optimism.

This special issue seeks to engage humanities and social science scholars committed to improving the decision-making of AI by focusing on debiasing strategies around notions of authenticity, provenance, representation, and inclusion in data capture and curation. Centred on the concept of Data Care, it promotes ethical, inclusive, and community-driven data ecosystems guided by the CARE principles: Collective Benefit, Authority to Control, Responsibility, and Ethics. Moving beyond critique, this issue fosters interdisciplinary dialogue on equitable AI development and invites contributions on replicable strategies to debias and diversify LLMs from a cross-cultural perspective.

Themes and Topics
This special issue seeks papers that move beyond critique to actively shape the development of Large Language Models (LLMs), through a humanities and social sciences led approach to debiasing and diversification. Contributions can take theoretical, empirical, or cross-cultural approaches, particularly from Global South and Indigenous contexts. The focus can be on:

- **Creative Methods**--case studies, comparative, and experimental methodologies for inclusive dataset training, curation, annotation, and governance.
- **Equitable and Representative Training Data**--strategies for integrating cultural, linguistic, and epistemic diversity in LLMs, addressing biases in dataset construction.

- **Politics of Inclusion and Exclusion**--analyses of geopolitical and corporate-driven data exclusions, as well as creative activist interventions that challenge algorithmic control, and strives to build new forms of inclusive standards.
- **'Rational optimism' Approaches**--collaborative approaches between humanities scholars, social scientists, and computer scientists to shape AI from within, working with stakeholders on the ground striving to optimize AI innovations and address chronic data deficits to build sustainable solutions.
- **Operationalizing the CARE Principles**--theoretical and empirical research on embedding Collective Benefit, Authority to Control, Responsibility, and Ethics into LLM development.

We invite interdisciplinary and cross-cultural comparative papers that propose actionable pathways for fairer, more culturally responsive AI systems.

Submission Guidelines

We invite contributions from humanities and social sciences scholars including but not limited to media and communications studies, anthropology, historians, cultural studies, STS, AI ethics, political sciences, and design. We particularly encourage submissions from researchers and practitioners based in the Global South.

Deadlines & Key Dates

EXTENDED ABSTRACT Submission Deadline: **Friday 28 March 2025** Please send 1000-1200 words (including references) to llmdatacare@gmail.com
- Include the name(s) of the author(s);
- The affiliation(s) and address(es) of the author(s);
- The e-mail address, and telephone number(s) of the corresponding
- Author.

NOTIFICATION of Accepted Abstracts to develop as full papers: **Monday 14 April 2025**

FULL PAPER DRAFT Submission: **Friday 2 May 2025**

ACCEPTED Papers are invited to attend a Workshop and co-read each other's paper for feedback: **Wednesday 11 June 2025**, University of Utrecht and hybrid.

REVISION Deadline: 1 October 2025

PUBLICATION Date: **2026**

All submissions will undergo a **double-blind peer review** process. The editorial team of *Information, Communication & Society* (ICS) has expressed interest in a full Special Issue proposal comprising approximately ten articles. We invite scholars to contribute their work for consideration, with the potential for inclusion in the proposed issue, pending final approval from the journal. For inquiries, please contact **Toija Cinque** llmdatacare@gmail.com

We look forward to your contributions to this important conversation on ensuring AI systems reflect and serve diverse cultural and creative perspectives!

TOIJA CINQUE, BA, MA, PhD, SFHEA
Deakin University | Melbourne Burwood Campus | Burwood VIC 3125 Australia
T + 613 92517470 | E Toija.Cinque@deakin.edu.au