

DIGITAL OTHERING: COMMUNICATION, ALGORITHMS, AND IDENTITY HARM

STATEMENT OF RESEARCH PROJECT

This project examines how digital platforms exacerbate “digital othering,” understood as the systematic process of marginalizing and dehumanizing social groups, through communicative and algorithmic dynamics. It will create a cross-cultural, multimodal dataset that captures overt, covert, and intersectional forms of othering across formats.

The core research question driving this project is: How can we diagnose, understand, and intervene in processes of digital othering across cultural and technological contexts without exacerbating polarization? To answer this, the project pursues three objectives: (1) to develop robust methods for detecting multimodal and intersectional forms of othering, (2) to examine the psychological and communicative effects of exposure to othering on users, and (3) to evaluate intervention strategies that mitigate harm while maintaining democratic discourse.

Phase 1 constructs the multimodal dataset using a rigorous annotation framework to identify rhetorical strategies such as dog whistles, irony, and metaphor. This phase asks: How does digital othering manifest across cultural contexts and communicative forms? The dataset will enable fine-grained detection of identity-based harms while capturing the communicative tactics used to frame groups as “outsiders.”

Phase 2 examines how exposure to digital othering affects users, integrating communication and psychological theories. It will measure perceived identity threat, stereotype activation, and affective polarization, while also analyzing how message framing, source credibility, and networked amplification shape user perceptions and behaviors. This phase specifically addresses the question: What are the cognitive, emotional, and relational consequences of encountering othering content, and how do these consequences differ across cultural and political contexts? By situating digital othering not only as content but also as a relational process, this phase shows how online harms affect trust, belonging, and democratic participation.

Phase 3 evaluates interventions. Using different communicative formats, we will test how it affects polarization, perceptions of credibility, and willingness to share content. This phase asks: Which communicative strategies are most effective in countering digital othering while minimizing backfire effects and fostering healthier public discourse?

By bridging computer science, communication, and psychology, this research develops a theoretically informed, empirically grounded framework for diagnosing, understanding, and reducing digital othering. Its contribution lies in connecting micro-level user effects with macro-level platform dynamics, offering interdisciplinary insights for platform governance, policy debates, and the design of ethical interventions that strengthen democratic resilience in the digital age.

SCOPE OF WORK FOR SELECTED PHD STUDENT

The selected PhD student will play a central role in designing, executing, and advancing this interdisciplinary project on digital othering. The work is structured across three integrated phases, each offering opportunities to contribute original knowledge at the intersection of communication, psychology, and computational social science.

In Phase 1 (Detection), the student will construct a cross-cultural, multimodal dataset capturing overt, covert, and intersectional forms of othering. This will involve developing and refining an annotation framework grounded in communication theory, coding textual

and visual content, and collaborating with computational experts to train and evaluate classification models.

In Phase 2 (User Effects), the student will design and conduct empirical studies to examine how users respond to othering content. This includes measuring perceived identity threat, stereotype activation, polarization, and trust, while analyzing how framing and networked amplification shape these effects. The student will synthesize psychological and communication literature to produce new insights into how digital harms influence cognition, emotion, and relational dynamics.

In Phase 3 (Interventions), the student will evaluate communicative strategies for mitigating digital othering. Using network analysis and controlled audits, they will test formats (such as fact-checks, counter-narratives, and dialogic responses), assessing their effects on credibility, polarization, and sharing behavior.

Beyond these phases, the PhD student will be encouraged to develop comparative or thematic extensions to broaden the project's regional relevance.

Overall, the student will gain training in multimodal data analysis, experimental design, and theory-driven intervention research, positioning them to make significant interdisciplinary contributions to communication, psychology, and platform governance.