



Venn diagram illustrating the partial overlaps between misinformation, bisinformation, and inappropriate language use, along with medical information and the BRIMI project focus.

Motivation:

- Research is motivated by the disproportionate harms that biased medical information wreaks on minoritized populations.
- Critical gap to be addressed is the dearth of knowledge on how *bisinformation* (biased information) spreads and manifests online.
- Vertically advances the field because disparities in medical care are exacerbated by bias; addressing at scale reduces inequality
- Transformative because our novel Fairness **via** AI framework reveals that AI can root out societal inequities from their source

Objective:

Our proposal, Bias Reduction In Medical Information (BRIMI), focuses on using AI to study, detect, mitigate and remedy biased, harmful, and/or false health information that disproportionately hurts marginalized and minoritized groups in society.

Technical Approach:

BRIMI develops novel artificial intelligence approaches to both establish health information inequities empirically and reduce them.

Outreach and Broader Impacts Plan:

- BRIMI will embed our research into practice through outreach to patients, integration into existing systems, and disseminating reports and tools to practitioners.
- BRIMI offers outsized promise for improved equity in health for minoritized communities and patients.
- Research and Education are integrated by leveraging advances in reducing bias in medical education and

Prior Results:

- Montenegro et al. developed and piloted bias reduction process in med schools.
- Dataset with 10,594 pages of instructional material, with 3,500 annotated examples of bias and related constructs, which will serve as seed data for BRIMI.

Deliverables

- Triage guidelines for bis- and misinformation disseminated to public health officials and journalists; Bias reduction tools integrated in medical schools; Twitter bot that flags biased or false medical claims.

Schedule:

- Y1-3:** Ethical social media & web data collection, particularly with respect to protected and sensitive characteristics and/or group membership
- Y1:** metrics defining bisinformation; extending prior work to lay language.
- Y2:** novel computational approaches to triage and prioritize types of bis- and misinformation, for the purposes of mitigation and countermessaging efforts.
- Y3:** Utilizing novel computational techniques in order to tailor and adapt public health interventions to specific individuals and/or population groups.

Contact: Shiri Dori-Hacohen, shiridh@uconn.edu