The middleware layer (<https://www.thefai.org/posts/shaping-the-future-of-social-media-with-middleware>, model context protocols) and the regulations imposed on downstream developers (http://arxiv.org/abs/2503.11922) are going to play a key role in safe AI systems development for consumers globally. I am interested in scoping out my own project focused on safe and secure AI for healthcare. First, I would like to investigate the standards creation (e.g. by the National Board of Health in India, the Indian Council for Medical Research, where we have collaborators) for AI deployment in healthcare in the global majority and second I would like to demonstrate a clear application of these standards: I will explore how we can improve the "social norms" alignment of AI systems (recent work<https://mbzuai.ac.ae/news/culturally-yours-a-new-tool-for-understanding-cultural-references-in-text>) by building effective AI-led persuasion experiments to advance digital health literacy as a positive externality of safe AI deployments. This project's impact can be amplified by a predefined partnership with the team at [https://sakhi-health.com](https://sakhi-health.com/) through whom I have access to GenAI health products deployed to over 300 families in India and Bangladesh and reaching 5000 by summer 2025. Sakhi has received offers to partner with large LLM companies and FAANG companies to roll out GenAI health technology over WhatsApp at scale, so I have an opportunity to scope out a high-impact, globally relevant AI safety field experiment complementing my proposed AI governance research that would round out my fellowship contributions.