Niloufar Salehi
Assistant Professor at the School of Information at UC, Berkeley

Title: From content moderation to school assignment: What do theories of justice teach us about design?

Abstract: Computational systems have a complex relationship with justice: they may be designed with the intent to promote justice, tasked to resolve injustices, or actively contribute to injustice itself. In this talk I will take two theories of justice, restorative and distributive justice, as frameworks to analyze and imagine alternatives to two real-world systems. First, I will analyze online harms such as harassment and revenge porn and how they are currently addressed through content moderation. I will use restorative justice to discuss the shortcomings of content moderation to effectively address those harms and discuss what alternatives we might design. Second, I will analyze an attempt at using computational systems to promote distributive justice in public schools in San Francisco that ultimately failed to achieve its theoretical promises of transparency, equity, and efficiency. I will show how incorrect modeling assumptions about families’ priorities, constraints, and goals clashed with the real world causing the algorithm to fail. Through this work I argue for recognizing the limitations of algorithmic solutions, broadening how we evaluate computational socio-technical systems, and ongoing engagement with those affected by those systems.

Bio: Niloufar Salehi is an Assistant Professor at the School of Information at UC, Berkeley, with an affiliated appointment in Electrical Engineering and Computer Science. Her research interests are in social computing, participatory and critical design, human-centered AI, and more broadly, human-computer-interaction (HCI). Her work has been published and received awards in premier venues in HCI including ACM CHI and CSCW. Through building computational social systems in collaboration with existing communities, controlled experiments, and ethnographic fieldwork, her research contributes the design of alternative social configurations online. http://niloufar.org/

Host: Sarah Fox