Shan Jiang | Research Statement

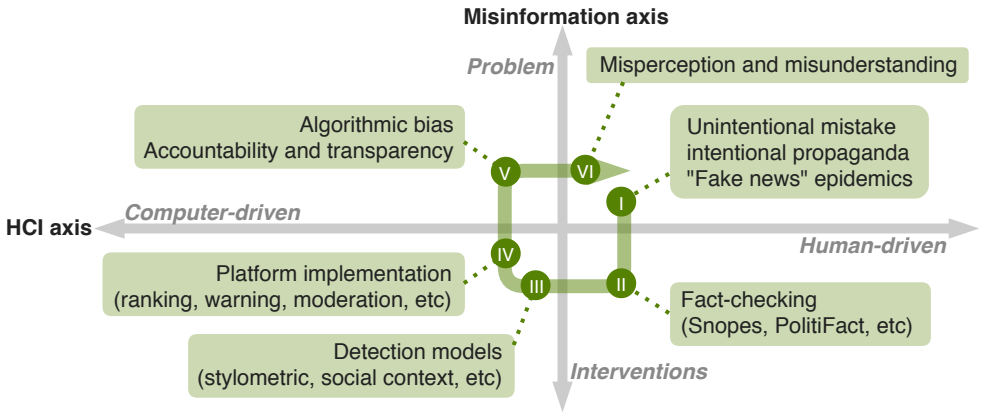
660-674, Interdisciplinary Science & Engineering Complex, 805 Columbus Ave, Boston, MA 02120
(+1) 781-502-8799 | sjiang@ccs.neu.edu | shanjiang.me | printfoo | shan-jiang

Research Interests

My research is focused around **social computing**, **computational social science**, and (more broadly) **human-computer interaction** (HCI). To examine these topics, I use a diverse set of tools, including statistical and causal inference, natural language processing, and machine learning. My recent work investigates HCI factors in the misinformation problem.

HCI factors in the misinformation problem

The spread of misinformation is a problem as old as communication, but has been recently exacerbated due to the emergence of computer-mediation. To help shed light on this, I developed an HCI perspective on the misinformation cycle (see figure) that maps the evolution of misinformation, enabling researchers to pinpoint solutions for it. Using this model, I've examined how misinformation and fact-checking affect social media users' conversational threads, audited algorithmic bias in Google Search, and investigated accusations of biased content moderation against YouTube.



Effect of Human-Driven Misinformation and Fact-Checking. Through computer-mediated communication systems, e.g., social media platforms, misinformation and fact-checking can impact people in many ways. In **CSCW'18a**, my collaborator and I studied how linguistic signals in social media users' conversational threads vary under misinformation and fact-checking. We collected 20K+ fact-check articles from Snopes and PolitiFact, and 2M+ user comments from Facebook, Twitter, and YouTube. Using sentiment analysis, we found that although users can sometimes become aware of misinformation, such misinformation still significantly discourages reasoned conversation by promoting hate speech and aggressive emotional cues in users' discussions. We also investigated the effects of fact-checking services and found that while there are signals indicating positive effects after fact-checking, there are also signals indicating potential "backfire" effects on audiences who view them as biased and unreliable. Leveraging these linguistic signals, we built predictive models for misinformation detections to help implementing more effective social computing systems for misinformation intervention.

Algorithmic Bias in Computer-Driven Interventions. Misinformation is not confined to social media; there is anecdotal that other algorithmically-curated platforms also spread misinformation. In **CSCW'18b**, my collaborators and I looked at this issue using a mixed-methods algorithm audit of partisan audience bias within Google Search. Following Donald Trump's inauguration, we recruited 187 participants to complete a survey and install a browser extension that enabled us to collect search engine result pages from their computers. We quantified partisan audience bias by leveraging the frequency of domain-sharing by a virtual panel of US registered voters with with different party affiliations on Twitter (I built a visualization system, **PolarShare**, for this data). We show that the partisan audience bias of search results depends largely on the queries being conducted, and that the embedded Twitter components have an average right-leaning bias. We also examined the claim that search engines create "filter bubbles" around their users, where its personalization algorithms infer users' preferences and prioritizes results accordingly. However, we found no significant differences in the partisan bias of the Google search results delivered to self-identified Democrats or Republicans, suggesting that concerns about filter bubbles on Google Search may not be grounded.

Misperception and Misunderstanding for Algorithmic Bias. Although some concerns about algorithmic bias are grounded, many claims surrounding this issue are actually due to misperception and misunderstanding about systems by human beings. In our recent work in progress (**WIP'19**, submitted to CHI'19), my collaborators and I investigated one specific claim that "social media is totally discriminating against Republican/Conservative voices" (Donald Trump on Twitter) using YouTube's comment moderation as a lens. We found that although comments under right-leaning videos are moderated more heavily, these comments also contain more hate speech and other inappropriate language. After controlling for confounders using a causal propensity score matching model, we find no evidence to support the claim of biased moderation against YouTube. Instead, we found that comments are more likely to be moderated if the video publisher is extremely partisan in either direction (left or right), if the video content is false, and if the comments were posted after a fact-check.

Besides work centering around misinformation, I am also broadly interested in fairness, accountability and transparency (FAT) issues around the web and the society. For example, in my upcoming **FAT*'19** paper, my collaborators and I measured online A/B/n experiments being run on web users by websites, and found cases of controversial experiments such as personalized news headlines without user awareness or consent. In **WWW'18**, my collaborators and I showed that ridesharing services such as Uber and Lyft provide low-quality services at low-income and high-diversity neighborhoods in San Francisco and New York City.

My work so far has been focused on empirical evaluation of deployed social computing systems using observational data. In the future, I plan to extend my research in the following two directions: 1) conducting experimental studies on user perception for the computer-driven misinformation problem, e.g., if and how can people identify algorithmic bias; and 2) building creative systems to support interventions against misinformation at-scale.

References

CSCW'18a Shan Jiang, Christo Wilson. "Linguistic Signals under Misinformation and Fact-Checking: Evidence from User Comments on Social Media." *Proceedings of the ACM: Human-Computer Interaction (PACMHCI)*, 2(CSCW), Nov 2018

CSCW'18b Ronald Robertson, Shan Jiang, Kenneth Joseph, Lisa Friedland, David Lazer, Christo Wilson. "Auditing Partisan Audience Bias within Google Search." *Proceedings of the ACM: Human-Computer Interaction (PACMHCI)*, 2(CSCW), Nov 2018

WWW'18 Shan Jiang, Le Chen, Alan Mislove, Christo Wilson. "On Ridesharing Competition and Accessibility: Evidence from Uber, Lyft, and Taxi." In *Proceedings of the 27th International Web Conference (WWW 2018)*, Lyon, France, Apr 2018

FAT*'19 Shan Jiang, John Martin, Christo Wilson. "Who's the Guinea Pig? Investigating Online A/B/n Tests in-the-Wild." In *Proceedings of the 2019 ACM Conference on Fairness, Accountability, and Transparency (FAT* 2019)*, Atlanta, GA, USA, Jan 2019

WIP'19 Shan Jiang, Ronald E Robertson, Christo Wilson. A paper investigating whether YouTube comment moderation is biased. Submitted to CHI'19. The title is hidden to compile with double blind policy.

PolarShare Shan Jiang. "PolarShare - Visualization of Media Bias by Polarized Sharing on Twitter." URL: https://polarshare.shanjiang.me