Federal Register Notice 87 FR 15274, https://www.federalregister.gov/documents/2022/03/17/2022-05683/request-for-information-on-federal-priorities-for-information-integrity-research-and-development, May 15, 2022

Request for Information on Federal Priorities for Information Integrity Research and Development

Indiana University Observatory on Social Media (IU-OSoMe)

Response to NITRD/NSF Request for Information on Information Integrity

This note is a response by the Indiana University Observatory on Social Media to "Request for Information on Federal Priorities for Information Integrity Research and Development". Based on our research, we respond to the questions (highlighted in a blue italic font) from the request below.

1. *Understanding the information ecosystem:* There are many components, interactions, incentives, social, psychological, physiological, and technological aspects, and other considerations that can be used to effectively characterize the information ecosystem. What are the key research challenges in providing a common foundation for understanding information manipulation within this complex information ecosystem?

The key challenges are...

- Understanding and quantifying the key vulnerabilities of the information ecosystem. We currently know about the widespread sharing of misinformation, coordinated posting of support for minority opinions, the bias of algorithms toward popular posts, the limited attention of users, and the division of the ecosystem into echo-chamber communities or epistemic bubbles. These vulnerabilities may be due to malicious behavior but can also emerge from organic behavior on social networks and corresponding issues of platform algorithms. Further ongoing research is required to monitor these current vulnerabilities and identify new vulnerabilities.
- Understanding and quantifying how the information ecosystem is manipulated. Studies have identified how inauthentic accounts, commonly known as bots, are used to manipulate social media. They have been used to amplify specific agendas and issues, and create echo chambers. Other studies have shown how a small number of accounts are responsible for a large proportion of misinformation.
- Understanding and quantifying the real-world impacts of information manipulation. This challenge is perhaps the most important, but yet the least studied. Early work has linked online misinformation to vaccine hesitancy and impacts on democratic systems. However, causal links between misinformation at scale and offline behavior remain largely understudied. The gold standard for these causal studies involve creating interventions and studying behavioral changes.

- Developing conceptual and empirical understandings of the characteristics and mechanisms of information ecosystems. There are many disciplines working on information systems, and an organizing framework that integrates these would be of value. This would enable more holistic understanding of the vulnerabilities of the systems.

2. *Preserving information integrity and mitigating the effects of information manipulation:* Strategies for protecting information integrity must integrate the best technical, social, behavioral, cultural, and equitable approaches. These strategies should accomplish a range of objectives including to detect information manipulation, discern the influence mechanisms and the targets of the influence activities, mitigate information manipulation, assess how individuals and organizations are likely to respond, and build resiliency against information manipulation.

Q1. What are the key gaps in knowledge or capabilities that research should focus on, in order to advance these objectives?

The key gaps in our knowledge are as follows...

- Future-proof strategies for detection and monitoring covert manipulation of social media platforms. Covert strategies tend to use systemic flaws in social media platforms. There is an ongoing arms-race between researchers and increasingly-advanced automated or cyborg (mixed automated and manually operated) accounts. These types of accounts can use coordination [1] to give the illusion of support for a perspective, manipulate the way accounts interact with one another [2], or delete large amounts of their posts to hide the evidence [3]. Another growing concern is the manipulation of detection systems themselves.
- Strategies for detection and monitoring of overt manipulation of social media platforms. These kinds of manipulation are generally done by well-known 'superspreader' individuals who have large numbers of followers. Studies have linked superspreaders to widespread health-misinformation on Twitter [4].
- Developing strategies for mitigating and inoculating against the effects of manipulation. These strategies require experimental studies to understand and measure the efficacy of specific interventions in controlled and field environments. Interventions could include tools to teach misinformation literacy [5], to unfollow bad actors, to highlight quality sources in social feed ranking algorithms [6], and to increase the friction of online interactions to make it harder to flood the network.

What are the gaps in knowledge regarding the differential impact of information manipulation and mitigations on different demographic groups?

There is a need to develop research to understand relationships between misinformation and certain demographics at scale. Large-scale observational studies have demonstrated links between misinformation and demographic or political groups [7,8] and there are likely to be many more associations to uncover. Studies should investigate likely disproportionate harms of misinformation on disadvantaged communities. Studies can target specific groups using surveys, controlled experiments, and interventions.

3. *Information awareness and education:* A key element of information integrity is to foster resilient and empowered individuals and institutions that can identify and abate manipulated information and create and utilize trustworthy information. What issues should research focus on to understand the barriers to greater public awareness of information manipulation? What challenges should research focus on to support the development of effective educational pathways?

Our early research [5] points to the benefits of educational tools that simulate social media environments and teach participants better social media literacy. These simulated environments allow teachers to present false information and bad actors to participants in a safe controlled way and develop their skills to recognise manipulation. However, preliminary results on this topic are sparse and suggest small effects. Challenges remain in large-scale assessments of different kinds of literacy interventions as well as in the development of realistic, age-appropriate, classroom-ready and ethically sound environments.

4. *Barriers for research:* Information integrity is a complex and multidisciplinary problem with many technical, social, and policy challenges that requires the sharing of expertise, data, and practices across the full spectrum of stakeholders, both domestically and internationally. What are the key barriers for conducting information integrity R&D? How could those barriers be remedied?

A critical and common barrier for research is the availability of data [9]. Many platforms do not release their publicly available data, and non-public data are very difficult to research. Moreover, making available harmful data that has been removed from platforms available to researchers would facilitate direct routes for inquiries that are urgently needed but currently closed [3]. It might be necessary to develop legal frameworks for the sharing of social-media data between platforms and researchers, and between researchers.

[9] Tackling misinformation: What researchers could do with social media data. Pasquetto, I. V.; Swire-Thompson, B.; and others HKS Misinformation Review, 1(8). 2020. http://doi.org/10.37016/mr-2020-49

5. *Transition to practice:* How can the Federal government foster the rapid transfer of information integrity R&D insights and results into practice, for the timely benefit of stakeholders and society?

We envisage two ways the federal government can help. First, they can provide legislation to make sure that all public data from social media platforms is made available to vetted researchers. Second, government can hold platforms accountable to maintaining their own standards for what is and isn't acceptable content to post online, and can back this up by developing national and international strategies for moderation and mitigation of harmful manipulation, and working with the platforms to develop their moderation and mitigation strategies.

6. *Relevant activities:* What other research and development strategies, plans, or activities, domestic or in other countries, including in multi-lateral organizations and within the private sector, should inform the U.S. Federal information integrity R&D strategic plan?

Misinformation knows no national boundaries so international coalition building around mitigation is key. There are growing calls for an international body that monitors and mitigates information integrity–specifically around large predictable events like pandemics, elections and wars. Such a body would be similar to international bodies we already have for conservation, nuclear monitoring, health, peace and security, or monetary cooperation.

7. *Support for technological advancement:* How can the Federal information integrity R&D strategic plan support the White House Office of Science and Technology Policy's mission:

- Ensuring the United States leads the world in technologies that are critical to our economic prosperity and national security; and
- maintaining the core values behind America's scientific leadership, including openness, transparency, honesty, equity, fair competition, objectivity, and democratic values.

The U.S. is a major contributor to the pollution of our information ecosystem. However, it is a concern that reaches all nations around the world. Taking a leadership role in developing a resilient, transparent and open information ecosystem has the potential to bring the U.S. great prestige.

References

[1] Uncovering Coordinated Networks on Social Media: Methods and Case Studies. Pacheco, D.; Hui, P.; Torres-Lugo, C.; Truong, B. T.; Flammini, A.; and Menczer, F. In Proc. International AAAI Conference on Web and Social Media (ICWSM), volume 15, pages 455-466, 2021.
[2] The Manufacture of Political Echo Chambers by Follow Train Abuse on Twitter. Torres-Lugo, C.; Yang, K.; and Menczer, F. In Proc. Intl. AAAI Conf. on Web and Social Media (ICWSM), 2022. Forthcoming. Preprint arXiv 2010.13691
[3] Manipulating Twitter Through Deletions. Torres-Lugo, C.; Pote, M.; Nwala, A.; and Menczer, F. In Proc. 16th Intl. AAAI Conf. on Web and Social Media (ICWSM), 2022. Forthcoming. Preprint arXiv 2203.13893
[4] The COVID-19 Infodemic: Twitter versus Facebook. Yang, K.; Pierri, F.; Hui, P.; Axelrod, D.; Torres-Lugo, C.; Bryden, J.; and Menczer, F. Big Data & Society, 8(1): 1–16. 2021.
[5] Fakey: A Game Intervention to Improve News Literacy on Social Media. Micallef, N.; Avram, M.; Menczer, F.; and Patil, S. Proc. ACM Human-Computer Interaction, 5(CSCW1): 6. 2021.
[6] Political audience diversity and news reliability in algorithmic ranking. Bhadani, S.; Yamaya, S.; Flammini, A.; Menczer, F.; Ciampaglia, G. L.; and Nyhan, B. Nature Human Behaviour. 2022.
http://doi.org/10.1038/s41562-021-01276-5
[7] Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. Pierri, F.; Perry, B.; DeVerna, M. R.; Yang, K.; Flammini, A.; Menczer, F.; and Bryden, J. Nature Scientific Reports, 2022.
https://doi.org/10.1038/s41598-022-10070-w
[8] Neutral Bots Probe Political Bias on Social Media. Chen, W.; Pacheco, D.; Yang, K.; and Menczer, F. Nature Communications, 12: 5580. 2021.