THE DISINFORMATION BLACK BOX: RESEARCHING SOCIAL MEDIA DATA

Testimony before the

House Science, Space and Technology Committee

Subcommittee on Investigations and Oversight

September 14th, 2021, 2 PM (EST)

Kevin T. Leicht

Professor of Sociology

University of Illinois Urbana-Champaign

Thank you for the opportunity to present our thoughts on social media data, social media misinformation, and the promises and pitfalls of researching social media data in our current economic and political landscape. I have spent most of the 1990s and 2000s studying the political and social consequences of social inequality and cultural fragmentation. Much of this work has focused on the changing landscape and growing skepticism confronting experts in most scientific fields. Much of this new skepticism is shared and spread via social media and attacks established, scientific knowledge across the board.

In my recent and on-going research (with colleagues Joseph Y. Yun, Geis College of Business, University of Illinois, Brant Houston, John and James L. Knight Foundation Professor of Investigative Journalism at the University of Illinois, Loretta Auvil, Senior Project Director at the National Center of Supercomputing Applications, University of Illinois, Peter Ondish, Research Scientist at the Center for Social and Behavioral Sciences, University of Illinois, Peter Evans, Professor of Sociology and Director of the Computational Social Science Program at the University of Chicago, and Prassana Balprakash, Senior Project Director, Argonne National Laboratory), I have examined how misinformation spreads via social media and whether

attempts by social media platforms to label such information has been effective in reducing sharing practices by users.

Our preliminary findings suggest that reliable and consistent labelling of social media misinformation by Facebook regarding COVID-19 severity, means of transmission, efficacy of vaccines, and potential miracle cures has been somewhat effective at preventing the spread of suspect posts. We have also discovered that Twitter does relatively little labelling of any kind, which is contrary to what we believed was their stated practice. However, because of the way Facebook's algorithms work, and our general lack of access to them as researchers, we are not able to tell whether the reduction in sharing of posts labelled as misinformation results from changes in algorithms (where posts labelled as misinformation are less likely to be prominently placed, less likely to be seen by end users, and thus less likely to be shared), or whether the reduction in sharing is due to actual changes in user's sharing behavior.

This unsatisfying outcome of our research so far highlights some of the problems I will highlight as I answer the committee's questions. While my research group has been helpful in constructing these answers, this testimony should be viewed as mine alone.

1. What patterns have you observed in how misinformation and disinformation spreads on social media platforms and the effectiveness of platforms' moderation techniques?

Apart from our research, there is quite a lot of research investigating how social media misinformation spreads, along with methods for debunking or flagging misinformation and mitigating its spread.

The proliferation of misinformation across Facebook and Twitter is generally quite similar and is often coordinated. The same posts or highly similar posts appear in both places. Contrary to what one might think, much of the low-credibility information comes from high-

profile, official, and verified accounts rather than muddy sources on the “dark web” (see Yang et al., 2021). As researchers and experts begin to fight misinformation contained in social media posts, there may be an uptick in creative ways to embed misinformation in new posts (by, for example, embedding texts in photos).

The spread of social media misinformation is also driven by differences in the actual people consuming the information. Personality type often predicts receptivity to misinformation (Axt et al, 2020) as well as demographic factors like age (older individuals tent to be more susceptible) and technological literacy (those will less tech literacy tend to be more susceptible, see Nagler et al., 2019). Laboratory studies suggest that one leading cause of misinformation spread is being overwhelmed with information – people cannot deliberately process and make accurate assessments when there is so much information being conveyed (Pennycook et al. 2020). There is also evidence that misinformation spreads not because it is fake, but because it is attractively packaged and unconstrained by reality (Acerbi, 2019). Laboratory studies also suggest that emphasizing publisher quality (e.g., The New York Times vs. Breirbart) may not reduce susceptibility to social media misinformation either (Dias et al., 2020). Clearly some people are more vulnerable to spreading misinformation than others and the sheer amount of information overwhelms many people.

The pervasiveness of social media misinformation has led to quite a bit of research on ways of combatting it. Repeatedly seeing fake news increases believability, even when the stories are labelled as “disputed”, except in extremely false cases (e.g. the Earth is a perfect square, see Pennycook et al., 2019). Flagging suspect social media posts works in many situations but not all of them (Swire-Thompson et al., 2020). More detailed debunking methods work better than cursory methods – citations to more credible information tend to work better

than a simple label (Chan et al., 2017). Finally, there is some evidence that who flags the information and who is debunking it has some effect, especially if the flagging is done by another person or by an AI algorithm (Yaqub et al., 2020.

In summary, we know a few things about misinformation and disinformation spread and some of the attempts to combat it. Despite this research, social media companies keep much of their data internal, so it is challenging to get an accurate assessment of how effective these platforms are at moderating misinformation. The social media platforms conduct many algorithmic and psychological studies internally, but the results are not disclosed. Some social media platforms like Facebook and Twitter do offer places where researchers can download data about posts and tweets (e.g., the CrowdTangle API for Facebook) and one can get an idea about the prevalence of certain types of posts through these tools. But most of the research that attempts to understand the causes and consequences of social media misinformation use lab settings that are much more controlled and very different from where users see misinformation as consumers. In lab settings, we can control the treatment that people are exposed to and then see what they do with that information. However, the black box algorithms the social media platforms have means that end users are exposed to vastly different "treatments" and we are left trying to understand what they do with the information without knowing what the initial stimulus was.

2. What are the limitations of current tools, techniques and data sets used to analyze social media?

There are serious limitations to the current tools and data sets available to analyze social media data. Those limitations fall into three general categories: (1) limited tools, (2) limited data availability, and (3) lack of coordination.

First, there are a limited number of tools available that are a combination of free/low cost and accessible to those that are not computer scientists. I believe we all can recognize that it's important to allow non-computer scientists to analyze social media at a relatively low cost, but such tools are few and far between. Some examples of tools that help in this way are NodeXL, Gephi, SNAP, and Professor Joseph Yun's own open-source tool, the Social Media Macroscope (at the University of Illinois). We need more funding to continue to build these tools to expand the number of researchers who can conduct this valuable work.

Second, there are limited amounts of social media data available due to company restrictions placed on that data. Many researchers fear litigation that may result from analyzing and publishing results from these data.

Third, there is little coordination regarding the analysis of social media data, especially for the sake of national security and societal wellbeing. Our group has advocated for the creation of national think tanks or laboratories to study social media effects on culture, social cohesion, and political life. In effect, each new researcher confronts the social media landscape virtually alone and without a lot of infrastructural help or assistance.

3. What kind of data can and should be made available by social media companies in order to understand the spread of misinformation and disinformation and its impact on society? Why is it important that researchers independent of social media companies have access to social media data?

A good start regarding data availability would be to focus on the data surrounding media platform misinformation and filtering. At present, this type of data and the algorithms used are less than transparent. There would need to be an open and ongoing conversation between social media companies and the research community, preferably a community represented by centralized think tank(s) for national social media research.

This access needs to be independent from the social media companies themselves because they have a conflict of interest with regards to researching and policing their own content. The goal of the social media companies is attention and engagement and, if extreme content produces that attention and engagement, that means more profit. Pursuing profit is not wrong, but one could question whether this model is contrary to overall social well-being.

The same holds true for research investigating misinformation on social media sites where social media platforms have a potential conflict of interest as well. This conflict of interest is most apparent in 1) the types of research questions investigated by social media companies from the onset, 2) how they interpret and understand the results they obtain, 3) how they report those results to public stakeholders (e.g., individuals, government, etc.), and 4) any action social media companies may or may not take based on their findings. This problem is compounded by the general lack of sharing of research results by the social media platforms.

4. How can the Federal government assist researchers in accessing data from social media companies that can help shed light on the spread of misinformation and disinformation?

Ideally the Federal government should work in concert with social media companies, private foundations, and Federal science funders to craft policies that require social media platforms to provide data for third-party groups to investigate specific, public-interest questions about misinformation incidence, prevalence, and consequences. This type of research would allow for independent checks and research on the power and impact social media companies have on misinformation and its spread.

The ways this data could be made available for researchers to use could take many forms. The creation of central data depositories of the kind that support other sorts of science infrastructure (for example, the Interuniversity Consortium of Political and Social Research, a

data depository at the University of Michigan) might be one model. Access to social media posts, platform algorithms, and sharing patterns could be provided to researchers who meet relatively strict human subjects/data confidentiality protocols. Another possibility would be to make social media platform data available in a restricted use data context, much as individual Census records are protected via Census Data Use Centers on university campuses. Other models are possible via cloud computing applications and depositories whose access could be relatively open while adhering to prevailing data confidentiality protocols.

This kind of policy would be another concrete step by the Federal government to formally consider misinformation for what it is: a public health and national security issue. To date, the deliberate spread of misinformation online has disrupted the US's efforts to address climate change, has enabled terrorist recruiting and communication, has eroded trust in political leaders, and continues to thwart efforts to vaccinate populations against COVID-19. More generally, misinformation threatens to untether the US population from reality by undermining our ability to preserve and recall a collective human history. Solving these public health, national security, and existential threats to the nation will require shared data and effort among industry, academic, and government partners.