



MAY 2018

WHAT'S NEXT FOR DIGITAL DISINFORMATION? A RESEARCH ROADMAP

Summary report of January
25, 2018 discussion hosted
by the Hewlett Foundation
and the Ford Foundation

Authored by Kelly Born, Hewlett
Foundation and Dr. Natalia Stroud,
Center for Media Engagement at
the University of Texas at Austin

TABLE OF CONTENTS



I.	Introduction	3
II.	Key Research Questions for the Field	5
1.	Understanding the Problem	6
a.	The Supply of Disinformation	6
b.	The Demand for Disinformation	7
c.	The Impact of Disinformation	9
2.	Understanding Potential Solutions	9
a.	Influencing How People Engage with Disinformation	9
b.	Influencing Affective Polarization	10
c.	Elevating Quality Content and Reducing Disinformation	11
III.	Infrastructure Needs for the Disinformation Research Field	13
1.	How do we keep on top of who is doing what?	13
2.	How can we build standard disinformation metrics?	14
3.	How can we share methodological expertise and datasets?	15
4.	How do we improve the impact of research?	18
5.	What scholarly ethics should guide research and data use?	19
IV.	Conclusion	19
V.	Appendix	20
1.	Convening Agenda	20
2.	Convening Participants	21

I. INTRODUCTION

Over the last year, concerns about digital disinformation and rising U.S. political polarization and democratic dysfunction have grown.

In some ways, disinformation and polarization are nothing new. In the early days of the United States, partisan rumors were rampant among politically engaged and polarized elites. Yet many things have changed in the contemporary political landscape. Social media now represent an important source of news for many Americans, and algorithms and social recommendations play an increasing role in how we consume information. And, as a society, America is more polarized in recent years than it has been in the past half century of sustained public opinion data collection. There has been speculation about the links among social media, polarization, and disinformation, and much remains to be learned.

Motivated by these questions and concerns, on January 25, 2018, the [Ford Foundation](#) and the [William and Flora Hewlett Foundation](#) convened a group of more than 50 scholars, think tank leaders, funders, and platform representatives to brainstorm key research questions and steps to address these concerns. They focused on digital disinformation, political polarization, and research.

First, digital disinformation spread via social media is just one component of Americans' political information diets, alongside television, talk radio, and offline networks of friends and family. But social media is the newest entrant, representing the area of most recent and rapid change, and is an area with huge reach and leverage.

And beyond polarization and disinformation, there are many other questions and concerns about the impact of social media on democracy: declining trust in institutions, declining journalistic revenue models, news fragmentation, foreign interference, and more. The choice to focus on political polarization and disinformation is in part a function of the conveners' main interests (the Hewlett Foundation's [democracy work](#) focuses here), and in part because of a belief that political polarization and disinformation are related to some of these other concerns. The two may in fact be circular: Disinformation feeds polarization, and polarization makes disinformation more likely to resonate, and circulate. Concerns about social media, disinformation, and polarization are also an area where many open questions remain.

Finally, there has been much discussion about the need for immediate policy interventions, and related advocacy efforts to support this, but to many, it is not yet clear that the problem is fully understood, much less that appropriate interventions that are capable of doing more good than harm, have been identified. And the community qualified to best assess this has not yet been fully formed. Thus focusing this convening specifically on research at the intersection of social media, disinformation, and polarization was viewed as both narrow enough to be manageable and broad enough to be meaningful.

The convening itself was structured around two main questions:

- First, what are the core research questions that must be addressed to enable sound decision making by both platforms and policymakers?
- Second, what core components of “infrastructure” are needed for the research field to succeed? (Essentially “public goods,” such as access to large datasets; shared understanding

of projects currently underway by key researchers [and closer to real-time translations of the policy implications of that research]; ethical codes of conduct for scholars in this field; and more).

This report is divided according to these two main questions, and reflects the authors' best attempts to summarize and explain notes composed by each of the breakout groups. The appendix includes the convening agenda and a list of attendees. While this report is not a reflection of any plans (or endorsements) by the Ford or Hewlett foundations, or the participants themselves, we are sharing the summary of the discussion to be transparent and inspire discussion and move the field forward.

II. KEY RESEARCH QUESTIONS FOR THE FIELD

Much remains to be understood about the impact of digital disinformation and polarization on U.S. democracy. But identifying, at a high-enough level, the most essential questions to inform real-world policy considerations has proven challenging. To help address this, a pre-event survey was fielded, asking “Let us know what you see as the two or three questions with practical, real-world implications that are most essential for us to answer in order to improve the current information environment, and explain why.” More than 100 responses were received from more than 50 experts. We organized these into a short list of questions, which we then compared to similar lists developed from 1) a fall 2017 report, [Philanthropic Opportunities to Mitigate the Disinformation/Propaganda Problem](#), based on interviews with more than 50 experts, developed by the Hewlett Foundation and consulting firm Social Velocity, and 2) [findings from a related discussion](#) at the University of Southern California, hosted by the Knight Foundation, Annenberg School for Communication and Journalism at USC, and Harvard’s Shorenstein Center’s First Draft News. The results of our survey, compared with prior work on identifying key research questions, were summarized into a short list of a half-dozen questions that then formed the breakout groups for this session.

We feel fairly comfortable that the resulting list elevates the research priorities the field must address in the coming years. These have been divided into questions focused on two main areas: understanding the problem and understanding potential solutions:

Understanding the Problem

1. **The supply of disinformation:** This included questions about the sources of disinformation and their incentives, the overall volume/prevalence of disinformation, the role that the media and politicians play in distributing disinformation, or the channels that are used most frequently for distribution.
2. **The demand for disinformation:** This included questions about people’s information environment and diets, behavioral/psychological drivers of polarized information consumption, factors that determine virality, and identification of segments most likely to consume or share disinformation.
3. **The impact of disinformation:** This included questions about the extent to which disinformation is shaping political beliefs, actions, and levels of polarization, as well as influencing other media sources, elected officials/constituent relationships, or democratic norms and values.

Understanding Potential Solutions

1. **Influencing how individuals engage with disinformation:** This included affecting the demand for quality information, correcting misperceptions, or reducing sharing tendencies.
2. **Influencing affective polarization among members of the public:** This included understanding how to encourage people to engage with diverse perspectives in democratically beneficial ways, or how to improve empathy or civility online.
3. **How to elevate quality content and/or reduce disinformation:** This included both the “how” (technical/design features that affect how people are exposed to or engage with content), as well as normative questions about whether to do so, and explanatory questions about the

impact of recent platform changes. Disinformation was defined broadly to include fabricated or manipulated content, as well as highly uncivil, inflammatory, and/or polarizing content.

At the meeting, participants self-selected into groups based on the six topics above. Following is a summary of these conversations.

Overall, perhaps as illuminating as the breakout conversations themselves was the pre-conference survey that, as mentioned, helped us to identify the main categories of questions requiring greater scholarly attention. For us, this has served as a useful starting point for beginning to organize the field of scholarly research.

1. Understanding the Problem

a. The Supply of Disinformation

The task of understanding the supply of *information*, much less the supply of *disinformation*, is challenging, particularly in the contemporary media environment. In earlier times, one could search newspaper coverage and major television news transcripts from LexisNexis and have a decent understanding of the available information.

Today, many things have changed. News organizations are not the gatekeepers they once were. Individuals produce and distribute information, state actors manipulate content, and algorithms affect what the public sees on social media. Add to this the presence of hyperlocals, individual conversations, and microtargeting, and the task of understanding the supply of information becomes even more daunting. And all of this, without even getting into the explosion of options for political information on cable TV and radio.

Moreover, parsing out the supply of disinformation requires first defining disinformation, a task that has yet to come anything close to achieving consensus. Some define disinformation, and distinguish it from other related concepts, on the basis of intentionality:¹

- Disinformation: Intentionally false or inaccurate information that is spread deliberately.
- Misinformation: Inaccurate information that is unintentionally promulgated.

Others have categorized types of mis- and dis-information, such as *fabricated content* (“news content that is 100% false, designed to deceive and do harm”) and *manipulated content* (“when genuine information or imagery is manipulated to deceive”).² And finally, sometimes information from what

¹ <https://www.hewlett.org/future-truth-can-philanthropy-help-mitigate-misinformation/>

² <https://medium.com/1st-draft/fake-news-its-complicated-dof773766c79>

many deem as “reliable” sources is taken out of context by political actors, so while the underlying stories have legitimacy, they are used to drive (and legitimize) heavily biased political narratives. Any research looking at disinformation will need to decide on, and justify, an operational definition.

When looking at the supply of information, it also is key to understand who produces, and who amplifies, disinformation. Many interested parties are involved in its production and distribution, but the extent to which each does so and how these behaviors vary over time and circumstance remains unknown. Several were identified as actors important for additional exploration:

- Foreign state actors and governments.
- Politicians and political elites.
- Profit-motivated actors.
- Organized groups of mass publics (e.g., loosely connected groups of individuals coming together for ideological reasons around particular issues).
- Partisan media.
- Social media accounts with lots of followers.

To capture **the supply of disinformation, how it varies by platform, and who is likely to share and amplify it**, extensive data are required. This might include gathering and archiving data from traditional media sources, in addition to data from other places where information and disinformation circulate, such as posts on Facebook, Reddit, and Twitter, and search results and auto-completes from Google. These data would allow for an examination of the amount of disinformation, and how disinformation spreads differently by platform. (Ideas for partially addressing this are discussed in Part 3 of the “Infrastructure” section below.)

b. The Demand for Disinformation

Just because disinformation is available does not mean that people will engage with it. Examining what factors motivate demand for disinformation represents a next critical research need.

Many features of the modern media environment influence the demand for information and disinformation, including that today:³

- Information travels **faster** and **lacks strong geographic boundaries**.
- The **costs of consumption have flattened**, such that most people today can have access to a wide array of news sources without incurring additional cost.

³ <https://embed.kumu.io/5e60a8840e4bb7do251fbbca577c1347#attention-flow-chart>

- **Digital archiving** makes it possible for news stories to reappear and be consumed at various points across time.
- The **emergence of social media** as a way to share and consume news also means that social ties now influence news consumption.⁴
- The **volume of news and array of channels** influence whether people encounter, and distinguish between, information and disinformation.
- **Ideology and partisanship** also can influence the demand for information and disinformation, yielding a process of motivated reasoning, or confirmation bias, whereby people gravitate toward likeminded information, and interpret messages as more or less valid based on their beliefs.

Defining what counts as “demand” requires careful analysis. Clicking, sharing, commenting, or reacting to disinformation could be conceptualized as demand for a particular piece of content. People’s intentions, however, may not always be to endorse the content. One can click for reasons of disdain, for example. Determining what counts as demand requires careful analysis.

There are also a number of factors that could influence demand that should be better understood, including:

- Seasonal and temporal variation:
 - Elections cycles.
 - Natural disasters and tragedies.
- Psychological drivers:
 - Identity and group affiliation, where disinformation is targeted and consumed on the basis of its alignment with one’s beliefs.
 - Motivations (e.g., for information, entertainment), where some actors consuming, or spreading, disinformation may be doing so for the purpose of entertainment.
- Structural influences:
 - How platforms are designed and how people use them (e.g., to what extent do Facebook users “turn off” people with whom they disagree?).
 - Societal factors that affect demand (e.g., how does demand differ across communities, regions, or countries?).

This work will require large-scale data collection. Datasets that combine attitudes, demographics, and consumption data (e.g., matching survey responses to social media activity) are key. As scholars contemplate different data needs and uses, we must also account for potential limitations on the use of such data, based on the terms under which platforms collected these data and applicable privacy laws.

⁴<https://www.americanpressinstitute.org/publications/reports/survey-research/trust-social-media/>

c. The Impact of Disinformation

Just knowing the supply of and demand for disinformation is not sufficient for understanding the problem. We also need to understand the effects of encountering disinformation.

Although there has been research on the effects of disinformation, more is needed to understand:

- How exposure to disinformation **affects attitudes and behaviors** among the public.
 - Is disinformation a distraction that crowds out (or undermines confidence in) legitimate information?
 - Does disinformation change online behavior, such as searching and sharing?
 - Does disinformation change offline views and behavior (levels of political polarization, voting, etc.)?
 - Which members of the public are more likely to be affected by disinformation?
- How **elites react** to the public's interactions, or perceived interactions, with disinformation.
 - Does disinformation produce polarization among the public, which reduces the ability of members of Congress to bargain?
 - Does disinformation influence the views and decisions of mainstream media? Of political leaders?

Data to tackle these sorts of questions can come in many forms, including:

- Aggregate data (e.g., exposure to disinformation by precinct).
- Matched pairs (e.g., those exposed/not exposed to disinformation).
- Search data pre- and post-exposure to disinformation.
- Experimental data (e.g., sharing information and surveying afterward).

At present, much is assumed about how disinformation is impacting democracies around the world, but much remains to be understood.

2. Understanding Potential Solutions

a. Influencing How People Engage with Disinformation

Reducing engagement with disinformation is complicated for many reasons. Scaling disinformation is much easier than scaling anti-disinformation (e.g., fact-checking). Disinformation uses memes and messages that scale quickly and attach to rising key words. Countering and correcting these tactics requires much more work than disseminating disinformation (e.g., some experts suggest eight to 15 hours of work are needed to fact-check content that may take just five minutes to create). In addition, people's partisan proclivities can influence whether they find disinformation credible.

Interventions could be designed to influence how people engage with disinformation. They likely would need to be different when tackling short-term, breaking news situations, as compared to longer-term solutions aimed at reducing the spread of disinformation.

- Short-term interventions:
 - Could alerts put out by platforms about the potential for disinformation during critical times (e.g., the first few hours after a breaking news event) reduce the circulation of disinformation? These alerts could **warn people about potential disinformation** and how to recognize it, such as the Breaking News Consumer’s Handbook from On the Media.⁵ These alerts also could act as news aggregators and describe known disinformation.

- Long-term interventions:
 - Could giving people **more information about sources** on social media affect engagement? One suggestion would be to create something akin to Google Scholar with information about past posts and engagement with them. This could serve to discredit problematic sources and amplify corrections from credible sources. It also would give the platforms data that could be used in algorithms to reduce the visibility of sources with a record of spreading disinformation.
 - Some sources sharing partisan information are not actual partisans. Figuring out ways to **discredit those infiltrating groups** by revealing that they are not group members may be helpful.

Overall, more research is needed to understand the impacts of various approaches to reducing engagement with disinformation.

b. Influencing Affective Polarization

Affective polarization is a feeling of disagreement, or a lack of favorability between groups. It is often measured using feeling thermometers (e.g., from 0 to 100, how do you feel toward Democrats versus Republicans?), or measures of prejudice toward out-groups (e.g., how would you feel if your son or daughter married a Democrat/Republican?). Research shows that affective polarization has increased dramatically in recent years.⁶

⁵ <https://www.wnyc.org/story/breakingnews-consumers-handbook/>

⁶ <https://academic.oup.com/poq/article-abstract/76/3/405/1894274>

Affective polarization has become a focus in conversations on disinformation because many disinformation campaigns try to target partisans, who presumably are less likely to counter-argue or criticize information that confirms their beliefs.

Several research questions emerged on how to reduce affective polarization:

- Some research suggests that **interpersonal contact** can reduce animosity between groups. Yet this is a very high-cost/high-touch model. Are there ways that this could be scaled? Could it be replicated via social media platforms? Can these face-to-face contacts that have positive effects offline translate to an online environment that is at best impersonal, and at worst anonymous and vitriolic?
- What **media architectures** reduce polarization, or give rise to “the civil self”? Can insights be gained by looking at existing spaces where people seem to engage constructively about politics with others that don’t share their views? ⁷ To what extent does anonymity and quality moderation affect polarization resulting from online engagement?
- Should attempts to reduce affective polarization begin in **non-political areas**? If we identify **spaces** where people with different partisan backgrounds interact about non-political topics, perhaps this could be leveraged to reduce polarization. Alternatively, perhaps efforts should begin with less politicized **topics**, and then slowly work up to those that are more politicized.
- Can the **presentation of information** reduce polarization? Can we share information about those with different political perspectives in a way that cuts through polarized reactions, and instead leads people to greater tolerance for difference? Some ideas mentioned by the group included: using first-person perspective narratives, allowing for anonymity, and building relationships on non-political topics before moving to politics.

c. Elevating Quality Content and Reducing Disinformation

Changing whether, and how, people encounter disinformation is another area where research is needed. Interventions might involve identifying pieces of disinformation, or identifying problematic accounts, and either warning people or algorithmically reducing the visibility of disinformation. It is important to figure out how to identify nefarious content quickly, especially in content vacuums (e.g., right after an event), while at the same time avoiding undue censorship and protecting norms and priorities around free speech.

⁷ Examples might include <https://www.reddit.com/r/changemyview/> or <https://www.reddit.com/r/AskTrumpSupporters/> or https://civilservant.io/about_us.html

One area of intervention that has been widely considered would be to look for quality indicators that could be used by platforms to surface quality content. But more research is needed to identify what indicators of quality exist, and how social media platforms are using them. Possible quality indicators include:

- Behavioral indicators: accounts with verified email addresses, low numbers of blocks/reports, longer times the account has existed, number of times the users post, etc.
- Reputation systems: examining how often disinformation has been sent in the past.
- Linguistic or tonal features: use of excessive hyperbole, uncivil terminology, or known partisan language.
- Adoption of particular practices: e.g., newsroom adoption of the Trust Project's Indicators, new Ethics Statements, etc. ⁸

After identifying possible indicators, it will be important to think through how they could be used or be gamed, and how often they are, or should be, updated. Length of time, for instance, may work for some contexts, but accounts that have been around for a long while can be purchased on the black market (generally at a higher price). Questions to assess the quality of the indicators should include:

- How much low-quality information is getting through?
- Should there be norms around what signals can/can't be used?
- Do signals of quality differ by location/audience segment/platform?

Instead of looking for quality indicators, another strategy could be to focus on the platforms' practices. Platforms could:

- Require more information from those placing ads.
- Introduce friction into trending topics to slow down rapidly spreading information that could be false.
- Cooperate to create an organization that works to reduce disinformation, using the Financial Industry Regulatory Authority as an example.
- Cooperate to trace users sharing disinformation across platforms.

⁸ <https://thetrustproject.org/>

Although the actions above target platforms, other direction actions are possible:

- Information could be given to sharers to allow them to assess the quality of content. Research would need to identify what information would be useful and how it could be communicated in a way that would actually inform users' decisions.
- Journalists could think about how they could help to debunk information in real-time, and brainstorm on the signals that they could create.

The research questions reviewed above represent ways to push our understanding forward. By better understanding the supply of, and demand for, disinformation, the field would be much better prepared to reduce the influence of disinformation and curb the detrimental aspects of political polarization. To this end, these research questions hopefully serve as a starting point for projects that can be tackled in the coming years. However, it should be noted that less attention was paid at this conference to issues around free speech, censorship, and potential unintended consequences of addressing challenges of online disinformation. This was in part due to the unavailability of key scholars in these domains. Incorporating these perspectives into future such convenings will be critical.

III. INFRASTRUCTURE NEEDS FOR THE DISINFORMATION RESEARCH FIELD

After discussing key research questions, attendees then turned to discussing a variety of critical supports that the field has yet to develop — essentially “public goods” that no individual scholar or researcher is incentivized to address, but which collectively would help move the field forward. These needs, surfaced (as with the research priorities above) through both a pre-event survey and a prior convening held on December 15th by the Knight Foundation, Annenberg School for Communication and Journalism at USC, and First Draft News, included:

1. How do we keep on top of **who is doing what**?
2. How can we build standard **disinformation metrics**?
3. How can we share **methodological expertise and datasets**?
4. How do we improve the **impact** of research?
5. What **scholarly ethics** should guide research/data use?

As with the Key Research Questions, participants self-selected to participate in sub-groups focused on one of the above needs. The following section includes notes from these five breakout groups.

1. How do we keep on top of who is doing what?

As concerns about digital disinformation and the role of online platforms in influencing political engagement have grown, many foundations, scholars, policymakers, and civil society actors have begun to do work in this space. With all this new activity, it is difficult to stay, in real time, on top of who is funding/researching what. This understanding will prove essential for foundations seeking to best allocate their funds; scholars deciding what to research and with whom to partner; and journalists, platforms, and policymakers determining who to turn to for answers to key questions.

Several *specific needs* emerged, many of which could be crowdsourced, including:

- Mapping who is funding/undertaking what new research.
- An events list or calendar (given the almost 50 events on this topic in 2017 alone).
- A bibliography of key papers and reports.
- Better mechanisms for platforms and policymakers to understand areas of scholarly consensus on key questions.
- Network-building across disciplines.
- A data repository (see Part 3 below).

Moving forward, the field could benefit from one (or a consortium of) institution(s) being on-point for managing this, as housing this information in a single place would improve field efficiency and enable better cross-pollination across fields and sub-disciplines. A need for field-building and a true accumulation of knowledge (not just data) was discussed.

Several *approaches* to addressing these information gaps were discussed, including:

- A database mapping scholars and funders to each of the 10 to 15 “critical” questions that have emerged (e.g., on supply and demand of disinformation, effective interventions, etc.).
- One-pagers/policy memos summarizing research findings.
- A Facebook page to keep the network connected.
- Webinars or two to four meetings per year, potentially around discreet topics.
- A journal for disinformation (which was de-prioritized in the near-term, given the amount of time it can take to create and appropriately incentivize).

Discussion then moved to who could best do this “infrastructure” work, with ideas including:

- The National Science Foundation (which has funded groups to do this in the past, but has not itself provided this kind of infrastructure).
- The Social Science Research Council.
- RAND Corporation.
- Harvard Kennedy School’s Shorenstein Center’s First Draft Project.
- Stanford’s Global Digital Policy Incubator.
- The Tow Center for Digital Journalism at Columbia
- University of Southern California
- A consortium (with different tasks going to different organizations).

Final questions/ideas included:

- The need for a more interdisciplinary framework.
- How to change publication incentives to best aid field building?
- How to bring in new voices, including computer scientists and legal scholars?
- How to bring in international contexts?

- How to keep people coming back (e.g., cross-publishing, incentives for academics, etc.)?

2. How can we build standard disinformation metrics?

Questions in this breakout group focused first around what, specifically, to measure, and then how to help align the field around common definitions.

Looking at what to measure, questions included:

- Should we measure *misinformers* or *misinformation*? The challenge with the latter is that there are no “normal” uses of words or patterns. In prior eras, information could be more easily evaluated at the level of the individual article. Today, platforms are struggling with this approach and instead appear to be focused more on evaluation at the level of specific actors.
- How do you identify and measure “*toxic behavior*” (assuming that toxic behavior does influence social change)?
- How do we identify malicious versus non-malicious coordination?

Perhaps even more challenging than defining metrics, is getting the field aligned around common definitions like misinformation, disinformation, and propaganda. Collaboration on categorization is essential if research findings are to be comparable across the field. Ideas included:

- Utilizing signals from the Trust Project at Santa Clara University.
- Leveraging anti-virus or anti-terrorism coalitions as a model on how to share information.
- Creating standards for collaboration.
- Sharing methodologies.
- Understanding better what signals make sense to platforms.
- Limiting the scope so as to more easily reach agreement.

3. How can we share methodological expertise and datasets?

Data access will prove critical to answering many of the most pressing questions identified. Yet access remains extremely uneven across platforms, with little access to Facebook data, moderate access to Google, and much greater access to Reddit and Twitter data (via its firehose), resulting in a disproportionate amount of research on the latter. Presently, there is also no way to track users and behaviors across various social media platforms, which will prove essential in understanding an individual’s political experience across platforms (including not just Twitter and Facebook, but also 8Chan, Voat, and others).

Two primary questions were surfaced with respect to Facebook data:

- Feed composition: What are users seeing?
- Network behavior: How are they interacting?

Three ways of addressing these problems were proposed:

- First, via survey by building a panel — something like what YouGov does, paying \$5 to users that are willing to contribute data to research, or akin to the American National Elections Studies (discussed below, e.g., a new “National Information Survey”).
- Second, scraping related data by looking for the same handles on different platforms: e.g., building an app that enables scraping people’s news feeds with permission (the challenge here being that, even within Reddit, people use different accounts).
- Third, through browser trackers (essentially compensating people to install a tracker allowing scholars to see all or most of their web behavior).

The most prominent idea discussed was around creating a new ANES (American National Elections Study) for online information — called perhaps the “National Information Survey” (NIS) — funded by foundations, platforms, and/or National Science Foundation. The essential components would include:

- Survey data — with a heavy emphasis on tracking internet and social media usage — linked to social media data, including multiple platforms (e.g., Twitter, Facebook, and other relevant platforms), supplemented by web browsing trackers.
- Consistent survey questions over time, potentially including open-ended questions.
- Some component of publicly available data, potentially with a privileged API for scholars.
- Open, transparent mechanisms for academics without the necessary relationships to propose research projects.
- Immediate data deposit so scholars can go back in time and analyze data to make generalizable conclusions, since problems are changing so quickly.

As an alternative to survey-based data compilations — likely preferable, but less feasible given current platform positions — funders could support a trusted institution to function as a data repository of social media datasets, with several key features:

- Working with platforms to ensure access to data for academic research and appropriate privacy safeguards.
- Providing a means for scholars to access this data, but that would include safeguards to make sure data is used properly.
- Corroborating accuracy of data received from platforms.

This would be valuable to platforms as well, enabling them to conduct research they currently cannot undertake given their lack of access to one another’s data. If run as a new entity, questions explored who “cares a lot about this, less about their own research, and is sociable.”

Additional ideas included:

- Using Amazon Mechanical Turk (though limited because it is not representative).

- Buying Facebook ads (but new transparency policy requires landing page to show all advertisements).
- Backing into high-level summary statistics on platform audiences, ad exposure, etc., through public earnings reports.
- Using public information such as AdWords, Facebook ad purchasing tools, etc.

One current effort, led by Mozilla, was discussed: efforts towards a “**cross-platform insight sharing group**” to share data with each other in privacy. This includes Mozilla (Mark Sherman, who runs Mozilla, and Philip Smith, who recently left and is at Stanford), alongside Microsoft, and several platforms focused on disinformation. They are currently exploring whether or not to involve a subset of scholars.

Operational questions included:

- How to address Internal Review Board limitations?
- How to access mobile data (e.g., from providers like Nielsen or Luth Research)?
- How can we develop tools that can help collect future data, such as browsing data beyond Facebook?
- How to manage the clash between the replication movement and terms of service?

A final need discussed in more detail was for a **tool repository**, listing all the tools that already exist, potentially with funders paying for subscriptions to these tools (e.g., open tool sharing/GitHub social media analysis). To that end, several existing resources were documented for tracking activities across platforms:

- SAM Desk: <https://www.samdesk.io/>
- CrowdTangle: <http://www.crowdtangle.com/>
- SimilarWeb: <https://www.similarweb.com/>
- Radian6: <https://login.radian6.com/>
- Keyhole: <http://keyhole.co/>
- Social Blade: <https://socialblade.com/>
- Klear: <https://klear.com/>
- Crimson Hexagon: <https://www.crimsonhexagon.com/>
- To know what public data already exists on platforms, participants recommended: <https://newslab.withgoogle.com/training/tools>

To improve on the above, attendees suggested compiling a **matrix of what data is available** through various tools.

Questions then turned to how to **enable greater platform engagement**, given platform concerns around responsibly sharing data with researchers in ways that don’t violate terms of service or user

trust, and that help to improve user experience. Ideas included:

- **Research on impact of policy changes:** When considering policy changes, platforms would like to have academics conduct research on what the impact of changes would be.
- **Specific, coordinated asks:** Academics need to approach platforms with specific product ideas, data requests, or research plans/questions. One idea was for academics to put together a coordinated ask of what information they want to see from the platforms. There is a clear need to agree on (and clarify for the platforms) data parameters — frequency, variables, level of aggregation, etc.
- **Establishing a data repository** (as discussed above) so platforms would have one partner to work with when releasing data, greatly reducing transaction costs for the platforms.

Two specific asks emerged, including:

- **Public user data:** Platforms should consider opening up APIs that expose public user data, or aggregate data sets based on private data, but for specific and general use
- **URL tracking:** Platforms should offer a search API allowing scholars to look for a given URL across pages and groups, and to pull impressions data for a given post through the API.

Of course actions by platforms to make data more accessible will be subject to the platforms' user terms and conditions.

Finally, a range of specific methodological challenges were surfaced:

- What do we know about the extent to which **people use the same handles** across platforms? Has anyone tried to collect this data by the scraping method?
- Can we develop **protocols to identify people**? This would be challenging because lots of people use multiple handles.
- How do we **normalize for influence** (e.g., Facebook likes versus Twitter favorites, including impressions numbers against each of these metrics)?
- How do we **identify and account for bots** and troll accounts?

4. How do we improve the impact of research?

Given the group's composition, the primary focus of this discussion was on how to improve the impact of academic research on platform (rather than government) policy. Key recommendations included:

- **Clearer points of contact:** Several platforms have acknowledged the need to provide researchers with more clarity around whom to contact in order to improve responsiveness on both ends (e.g., a list of key platform leads by topic).
- **Clearer methodologies:** Scholars need to provide “more rigid methodologies and explanations” so platforms can make sense of findings. For example: How are scholars coding

for false/“fake news” content? If scholars have identified problematic accounts, they could provide these lists to platforms.

- **Advance notice on publications:** Platforms need more time/advance notice to consider research findings/implications. Instead they are often asked to respond to research within hours of its publication. Sharing working papers before publication could help address this. Contacting platforms earlier in the research process could prove even more helpful, allowing the platforms to work with scholars to inform them about platform-specific considerations that could affect their research design.
- **Panel/hub to filter requests:** Platforms “need some type of system to understand how to prioritize our time.” Platforms currently waste a lot of time responding to research inquiries that may or may not be important or useful. Is there a role for an academic research consortia — e.g., an outside body that could help ensure the credibility of the queries coming in, potentially with a templated structure for research inquiries, reviewed by an intermediary panel that could include platform representatives?
- **Journalist/scholar partnerships:** Journalists and scholars could benefit from journalist-academic partnerships/“mentorships” — pairing teams of data-literate researchers with tech/culture journalists.
- **Convenings:** Several expressed the need for regular research-platform meetings with structured themes and agendas (e.g., more opportunities to have academics present recent work at Google, Facebook, etc.; lightning talks and workshops; webinars; etc.)

5. What scholarly ethics should guide research and data use?

Several outstanding problems and questions were identified around:

- The ethics of doing research using social media data.
- The ethics around informing the public about findings.
- Adverse incentives scholars face around 1) attempting to “cut deals” with platforms individually for data access and 2) pressure to rush out with results, given current news cycles and interest.
- The failure to (yet) make collective data requests of the platforms.

Several ideas for how to operationalize this were elevated:

- First, Microsoft Research was identified as an exemplary open/transparent partner that the other platforms could learn from — a case study of how this has benefited their work could be helpful.
- Second, some form of self-enforcement among the scholars at the review/publication stage could be explored to prevent “cutting one-off deals” or publishing without peer review (i.e., research based on private data that others couldn’t study would not be considered legitimate).

- Third, a pledge to uphold a code of conduct could be introduced as a condition of receiving foundation funding.

IV. CONCLUSION

Before the field of philanthropic organizations, academics, issue experts, or social media platforms focused on disinformation can be in a position to adequately address current challenges of digital disinformation growing political polarization, and their related effects, we must first better understand both the supply of and demand for disinformation, its effect on citizen’s beliefs and behaviors, and the actual impact of proposed interventions. Much research is needed to answer the above.

At the same time, for this emerging scholarly field to succeed, several critical components of infrastructure — essentially “public goods” — must be developed. Thus, beyond providing direct support for research, we believe philanthropy can and must play an important role in helping to develop necessary datasets and metrics, tracking scholarly findings and translating them to policymakers, and otherwise supporting this nascent field.

These discussions have informed the Hewlett and Ford foundations’ understanding of how digital disinformation affects democratic norms, values, and processes. And it is helping them, alongside other funders, identify the ways philanthropy can support further research and response. We hope it likewise proves of use to others.

V. APPENDIX

1. Digital Disinformation and Political Polarization Scholars Convening Agenda

School for Communication and Journalism at USC

THURSDAY, JANUARY 25, 2018

9:30-10:30 a.m. Welcome & Introductions

10:30-11 a.m. Paired Conversations

11:15 a.m.-noon Panel: What’s the Most Important Next Step for the Field?

- Joan Donovan, Data & Society
- Alondra Nelson, Social Science Research Council & Columbia University
- Samuel Woolley, Digital Intelligence Laboratory at the Institute for the Future

1:15-2:45 p.m. Research Gaps, Data Needs, and Real-World Application

Brainstorm specific, real-world research questions related to digital disinformation and polarization, and elaborate on the data needed to answer them.

3:15-5:00 p.m. Platform Introductions and Reactions

Hear from and talk with representatives from several platforms on how they are organizing this work internally, collaborating with researchers, etc.

FRIDAY, JANUARY 26

9:15-9:35 a.m. Review and Debrief on Ideas from Previous Day

9:35-10:25 a.m. Resource Sharing Breakout Groups

Identify resources that could help people to solve challenges that they're facing in their work.

10:25-10:50 a.m. Infrastructure Breakout Preview and Break

10:50 a.m.-noon What Infrastructure is Needed to Best Support the Field?

Determine what infrastructure would be useful (e.g., convenings, listservs, datasets, ethical guidelines, etc.) to support scholars doing this sort of research.

2. Convening Participants

Jonathan Albright Research Director at the Tow Center for Digital Journalism

Mike Ananny Assistant Professor at University of Southern California Annenberg School for Communication and Journalism

Jessica Ashooh Director of Policy at Reddit

David Askenazi Director of Learning and Impact at the Knight Foundation

Julia Azari Associate Professor and Assistant Chair in the Department of Political Science at Marquette University

Adam Berinsky Mitsui Professor of Political Science at MIT and Director of the MIT Political Experiments Research Lab (PERL)

Kelly Born Program Officer for the Madison Initiative at the William and Flora Hewlett Foundation

Danielle Bowers Research and Development Manager for Google News Lab

Samantha Bradshaw D.Phil. candidate at the Oxford Internet Institute

Katherine Brown Facebook's Global Policy Development team

Nic Dias Senior Fellow at First Draft News

Renée DiResta Head of Policy at Data for Democracy and Staff Associate at Columbia University Data Science Institute

Eileen Donahoe Executive Director of the Global Digital Policy Incubator at Stanford University's Center for Democracy, Development, and the Rule of Law

Joan Donovan Project Lead for the Media Manipulation project at Data & Society

Charlie Firestone Executive Director of the Aspen Institute Communications and Society Program and a Vice President of the Aspen Institute

Michael Gardner Data Scientist on the Special Ops Team at Reddit

Matthew Gentzkow is Professor of Economics at Stanford University

Tom Glaisyer is the Director of the Public Square Program at the Democracy Fund

Daniel Green Director of Program Advocacy and Communications at the Bill & Melinda Gates Foundation

Andrew Guess Assistant Professor of Politics and Public Affairs at Princeton University

Tim Hwang Director of the Ethics and Governance of joint MIT-Harvard AI Initiative

Jay Jennings Postdoctoral Research Fellow for the Center for Media Engagement and the Annette Strauss Institute for Civic Life at the University of Texas Austin

Jennifer Kavanagh Political Scientist at the RAND Corporation and Associate Director of the Strategy, Doctrine, and Resources Program in RAND's Arroyo Center

Young Mie Kim Professor of the School of Journalism and Mass Communication and a Faculty Affiliate of the Department of Political Science at University of Wisconsin, Madison

Raina Kumra is an Entrepreneur in Residence at Omidyar Network

David Lazer Distinguished Professor at Northeastern University and Visiting Scholar at the Institute for Quantitative Social Science, Harvard

Alex Leavitt Computational Social Scientist at Facebook Research

Colin Maclay Executive Director, Professor (Research) and Lead Instigator at the USC Annenberg Innovation Lab

Neil Malhotra Edith M. Cornell Professor of Political Economy in the Graduate School of Business at Stanford University

Lilliana Mason Assistant Professor of Government and Politics at the University of Maryland

Lori McGlinchey Senior Program Officer for Internet Freedom at the Ford Foundation

Rachel Moran Doctoral Student at the USC Annenberg School for Communication and Journalism

Jonathan Nagler Professor of Politics and Director of the Politics Data Center at New York University

Dawn Nakagawa Executive Vice President of the Berggruen Institute

Alondra Nelson of the Social Science Research Council and Professor of Sociology at Columbia University

Nate Persily James B. McClatchy Professor of Law at Stanford Law School

Nick Pickles Head of Public Policy and Government for the U.K. and Israel at Twitter

Kathy Porto Chang Product Manager at Twitter

Jennifer Preston Vice President for Journalism at the John S. and James L. Knight Foundation

David Rothschild Economist at Microsoft Research

Ashkan Soltani Independent researcher and technologist specializing in privacy, security, and behavioral economics

Daniel Stid Director of the Madison Initiative at the Hewlett Foundation

Talia Stroud Associate Professor of Communication Studies and Journalism, Director of the Center for Media Engagement, University of Texas, Austin

Sara Su Product Manager at Facebook, News Feed

Joshua A. Tucker Professor of Politics, affiliated Professor of Russian and Slavic Studies, and Data Science at New York University, Co-Director of the NYU Social Media and Political Participation (SMaPP) laboratory, and Director of NYU's Jordan Center for Advanced Study of Russia

Duncan Watts Principal Researcher at Microsoft Research

Wilson White is a public policy and government relations senior counsel on Google's Mountain View policy team

Diane Winston Associate Professor and Knight Chair in Media and Religion, USC Annenberg School of Communication and Journalism

Magdalena Wojcieszak Associate Professor at the Department of Communication at University of California, Davis

Clément Wolf Global Public Policy Manager at Google

Samuel Woolley Research Director of the Digital Intelligence Laboratory at the Institute for the Future

Ashley Wynn User Experience Researcher at Google and lead researcher on the News Credibility team