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

Jonathan M.

Comment of Academic Researchers on
Federal Priorities for Information Integrity Research and Development

Thank you for the opportunity to inform the Interagency Working Group's strategic plan for a whole-of-government approach to information integrity research and development.

We are academic researchers who study information integrity topics from diverse disciplinary perspectives, including communication, computer science, economics, and law. Our scholarship aims to characterize the scope, scale, and effects of mis- and disinformation.[1] We also seek to mitigate these challenges with our research, by examining methods for identifying and responding to mis- and disinformation.[2]

Before turning to our substantive suggestions for the strategic plan, we would like to offer a recommendation for the plan's framing and the goals it may articulate. While the White House announcement of the Working Group and the Working Group's RFI both discuss "understand[ing] the full information ecosystem," the documents emphasize "information manipulation" such as mis- and disinformation. We urge the Working

[1] *E.g.*, Shan Jiang et al., *Modeling and Measuring Expressed (Dis)belief in (Mis)information*, ICWSM (2020); Miriam J. Metzger et al., *From Dark to Light: The Many Shades of Sharing Misinformation Online*, Media & Communication (2021); Tanushree Mitra & Eric Gilbert, *CREDBANK: A Large-Scale Social Media Corpus With Associated Credibility Annotations*, ICWSM (2015); Katherine Ognyanova et al., *Misinformation in Action*, HKS Misinformation Review (2020); Mattia Samory & Tanushree Mitra, *Conspiracies Online: User discussions in a Conspiracy Community Following Dramatic Events*, ICWSM (2018); Julio C.S. Reis et al., *A Dataset of Fact-checked Images Shared on WhatsApp During the Brazilian and Indian Elections*, ICWSM (2020).
[2] *E.g.*, Austin Hounsel et al., *Identifying Disinformation Websites Using Infrastructure Features*, FOCI (2020); Ben Kaiser et al., *Adapting Security Warnings to Counter Online Disinformation*, Usenix Security (2021); J. Nathan Matias, *Nudging Algorithms by Influencing Human Behavior: Effects of Encouraging Fact-Checking on News Rankings* (2020); Julio C. S. Reis et al., *Can WhatsApp Benefit from Debunked Fact-checked Stories to Reduce Misinformation?*, HKS Misinformation Review (2020).

Group to expressly address a broader class of potentially harmful information[3] and to elevate study of the overall information ecosystem in the strategic plan.[4] There is a growing body of evidence that factually false or misleading information accounts for a comparatively small component of the overall information ecosystem.[5] Politically polarized or emotionally charged information is much more prevalent and is an important contributing factor to the individual and societal harms that prompted the Working Group's formation.[6] Advancing understanding of the overall information ecosystem, including the most effective sources of information and trends in the ecosystem over time,[7] will also be essential for contextualizing information integrity challenges and developing possible responses. While our comment focuses on mis- and disinformation research, consistent with the emphasis in the White House announcement and RFI, our observations are generally applicable to study of other types of potentially harmful information and the overall information ecosystem.

We offer four recommendations for the strategic plan that the Working Group is developing. First, federal government R&D efforts related to information integrity would benefit from greater clarity of each agency's responsibilities and coordination of relevant initiatives across agencies. Second, federal research funding should continue to rapidly scale up in this topic area. Third, federal R&D strategy should prioritize enabling methodological advances in information integrity research, by building reusable technical infrastructure, addressing legal ambiguity, and facilitating access to online platform data. Fourth, federal strategy should advance the interdisciplinary research and teaching infrastructure necessary for this area of R&D.

---

[3] The Working Group could, for example, clarify that "information manipulation" and "manipulated information" include a broader class of potentially harmful information, such as instances of politically polarized or emotionally charged information that could cause individual or societal harm.

[4] We commend the administration for including "understand[ing] the full information ecosystem" within the Working Group's charge. Our recommendation is not a substantive change to the Working Group's scope, but rather, a rebalancing of emphasis and priorities in the strategic plan.

[5] *E.g.*, Jennifer Allen et al., *Evaluating the Fake News Problem at the Scale of the Information Ecosystem*, Science Advances (2020); Andrew Guess et al., *Less Than You Think: Prevalence and Predictors of Fake News Dissemination on Facebook*, Science Advances (2019)

[6] *See generally* Joshua A. Tucker et al., *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature* (2018).

[7] Trends that may be particularly relevant include the transitions from local to national news coverage, from institutional journalism to social media, from print and TV to online information sources, and from desktop to mobile engagement with information.

1. Federal research and development efforts related to information integrity would benefit from greater clarity of agency responsibility and interagency coordination.

Information integrity research defies easy categorization. Within academia, these topics typically span academic disciplines and fields within disciplines. Similarly, within the federal government, information integrity R&D crosses agency authorities, agency equities, and programs within agencies. There is currently limited clarity of what each agency's responsibilities are and which information integrity topics fall within particular agency programs. There is also limited coordination of information integrity R&D across the federal government.

This lack of clarity and coordination has significant consequences. Researchers and other experts external to the federal government may fall into silos, interacting with agencies, programs, and staff based predominantly on familiarity and preexisting professional networks rather than relevant expertise, experience, and needs. The status quo deprives the field of valuable bidirectional information exchange with the federal government and limits opportunities for federal collaboration with and support of external research.

Consider, as an example, the developing area of research on interventions to address mis- and disinformation related to COVID-19 on online platforms. Which agencies should be responsible for research efforts in the area? NSF is a logical candidate, because of its broad responsibilities related to research support. NIH and CDC also play vital roles, because the topic relates to medicine and public health. But the list continues: this type of mis- and disinformation can have implications for armed forces readiness, creating a role for the Department of Defense and its components, including DARPA. This type of content can also have foreign policy implications, or be associated with foreign information operations, such that the State Department and the Intelligence Community have relevant equities. Because this type of content can impact critical infrastructure, especially in the healthcare and emergency services sectors, the Department of Homeland Security has a role. When there is a commercial aspect to false or misleading information, or research examines online platform practices, the Federal Trade Commission may have a role. The list could go on.

These complexities continue within agencies. At NSF, as an example, this type of research may fit within the Directorate for Computer and Information Science and Engineering (CISE) or the Directorate for Social, Behavioral, and Economic Sciences (SBE), among other agency components. And within CISE, the Secure and Trustworthy Cyberspace, Human-Centered Computing, and the Designing Accountable Software Systems programs are all possible points of connection.

This fragmented R&D landscape leads to missed opportunities. Researchers who approach the topic from a computer science perspective predominantly interact with one directorate at NSF (CISE), those who come from social science perspectives interact with another directorate (SBE), and those from medical and public health backgrounds engage with entirely different agencies (NIH and CDC). Meanwhile, interaction with other relevant components of the federal government—especially those with operational responsibilities, which may have valuable experience and expertise in the problem area—is generally limited. The result is missed opportunities for information exchange, research collaboration, research support, and better informed policymaking.

As a starting point, we encourage the Working Group to develop a public directory of federal agencies and programs relevant to information integrity R&D, with a summary of responsibilities and point of contact for each.[8] This simple step could have significant benefits. Earlier this year, NSF division directors published a brief Dear Colleague letter clarifying that the Secure and Trustworthy Cyberspace program supports information integrity research and describing the topics, disciplines, and methods that are within scope.[9] The letter has been invaluable to the information integrity research community, highlighting a point of connection with the federal government that was unfamiliar to some researchers and prompting new interdisciplinary collaborations. We recommend that the Working Group consider replicating that model across federal R&D.

---

[8] We do not take a position on which component of the federal government should be responsible for developing the directory or hosting the convening that we propose. There are a range of possible models, such as the recent National Artificial Intelligence Initiative and the National Quantum Coordination Office. We also take no position on the right allocation of responsibility among agencies or programs—which may benefit from overlapping portfolios, because different components of the federal government have different equities, resources, and types of expertise.

[9] Sylvia Butterfield et al., NSF 22-050, *Dear Colleague Letter: Inviting Proposals Related to Information Integrity to the Secure and Trustworthy Cyberspace Program* (Feb. 24, 2022).

Another valuable step would be coordinating a recurring event where policymakers and R&D practitioners across the federal government can engage with external researchers and other experts. A potential model is the annual data privacy conference (PrivacyCon) organized by the Federal Trade Commission, which for six years has convened federal privacy regulators and leading experts to exchange ideas, inform policy, and advance the research field.[10] A similar recurring event, connecting federal officials with external experts on information integrity, would be an asset to the field.

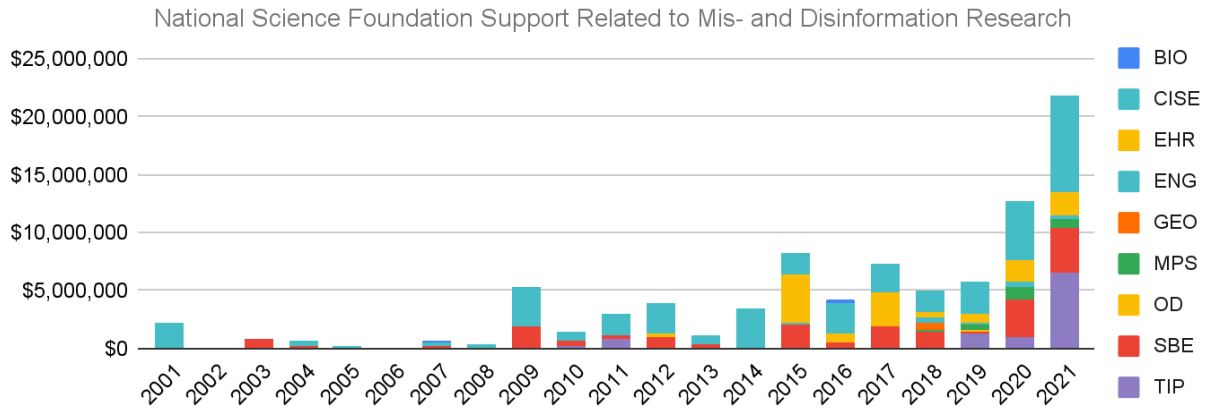
2. The federal government should continue to significantly scale up research support related to information integrity.

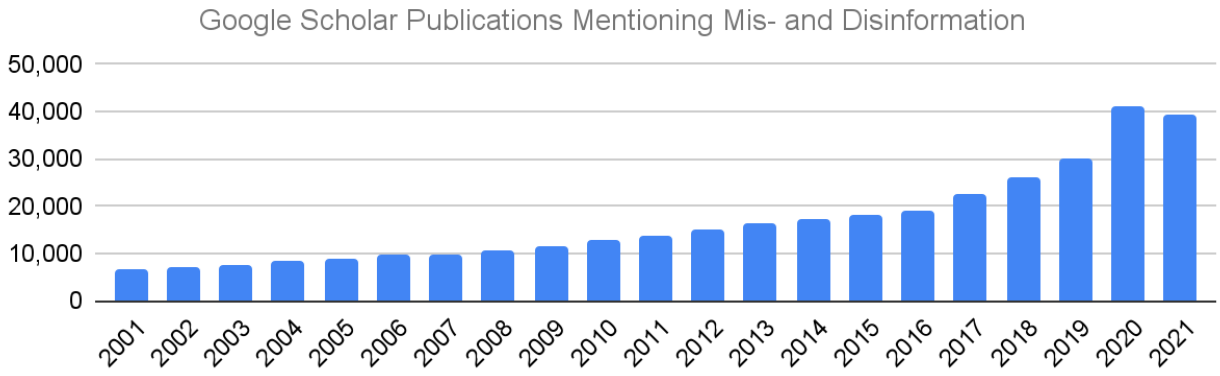
There is, at present, an extreme mismatch between the magnitude of the societal challenge posed by information integrity problems, the level and timeliness of federal support, and the rate of research productivity.

As rough empirical guides, we have compiled National Science Foundation awards that mention "misinformation" or "disinformation," as well as publications on Google Scholar that mention either of those terms.[11]

[10] FTC staff also contribute to the organization of an annual workshop on consumer protection and technology (ConPro), which is another valuable possible model for the Working Group.

[11] The NSF figure presents the amount awarded to date by award year, relying on data exported from the NSF award search website. The Google Scholar figure presents the count of papers by year of publication. The keyword matching that we use to generate data on NSF awards and Google Scholar publications is inexact. Grants related to information integrity, including a major $15.7 million award to Northeastern University, do not consistently use the terms "misinformation" or "disinformation" in their title or abstract. Publications may also use different terminology for information integrity concepts, and mentions of the two keywords is not conclusive that a grant or publication is closely related to information integrity.





A pair of trends are evident from this rudimentary data. First, the field of information integrity research has been growing rapidly, and federal support has not kept pace. While philanthropies have stepped in to augment federal funding, such as the Knight Foundation, the NetGain Partnership, and the Omidyar Network, these sources of support are not sufficient to sustain and grow an entire area of research. Corporate funding related to information integrity, meanwhile, remains relatively limited—Meta, for instance, awards about $1 million in competitive information integrity grants annually. Technology firms face complex incentives related to mis- and disinformation, and the federal government should not count on the private sector to address funding gaps or prioritize the most societally urgent aspects of information integrity.

A second readily apparent trend is that support for information integrity research began to significantly increase in 2020. We commend NSF's recent prioritization of this area, especially through the CISE, SBE, and EHR directorates and the TIP directorate's Convergence Accelerator program.

We encourage the Working Group to consider how to ensure consistent and continued growth in the level of support for this research area until the federal government reaches a level commensurate with the societal urgency of the topic. The Working Group may find it helpful to more rigorously benchmark public sector, private sector, and philanthropic support for information integrity research in comparison to other R&D priority areas. Examining past and potential barriers to federal information integrity R&D would also be a valuable step.

3. Federal R&D strategy should prioritize enabling methodological advances for information integrity research, by building reusable technical infrastructure, addressing legal ambiguity, and facilitating access to online platform data.

Research related to information integrity often uses a familiar set of methods, such as interviews, surveys, laboratory experiments, crowd tasks, web crawls, and social media archive analysis. These methods are valuable, and we use them in our own research. But these methods have significant limitations: they often are not representative of user experiences and activities, especially when users are interacting with personalized or targeted content. These methods limit the types of studies that researchers can carry out, and results from studies using these methods can have questionable ecological validity.

As an example, there are hundreds of research publications that explore how to design user interface interventions to counteract mis- and disinformation.[12] The overwhelming majority of recent papers use simple survey-like methods—asking participants to scroll a simulated social media feed and self-report perceptions and predicted actions—as well as similar warning and notice labels. Researchers in the information integrity field recognize that these methods are not very realistic and that there is a pressing need to examine alternative interventions beyond labels attached to content. But, absent methodological innovation, implementing more realistic research designs and exploring alternative interventions will continue to be elusive.

[12] *See* Laura Courchesne et al., *Review of Social Science Research on the Impact of Countermeasures Against Influence Operations*, HKS Misinformation Review (2021).

We urge the Working Group to examine how federal R&D strategy might support methodological advances for information integrity research. One important path forward is building reusable technical infrastructure, offering new data sources and capabilities for measurement and intervention experiments.[13] Mozilla Rally, for example, is an initiative that enables running research studies in participant web browsers, observing real online experiences and deploying interventions in real online settings. By making a significant upfront investment in browser instrumentation, a data analysis environment, and participant recruiting, Rally enables research with innovative methods at relatively low marginal cost and effort. Similar promising projects include The Markup's Citizen Browser (which has enabled dozens of projects examining the Facebook information ecosystem), the NYU Ad Observatory (which provides valuable ongoing data about Facebook advertising content), and Northeastern's upcoming NSF-supported Observatory for Online Human and Platform Behavior (which will enable observational study of real experiences and activities online). The federal government could have a transformative impact on the information integrity research field by supporting initiatives like these.

Another important step for enabling methodological advances in information integrity research would be addressing legal ambiguities. When researchers conduct independent study of an online platform, regrettably, sometimes the platform threatens legal action against the research team. A particularly vivid example of this issue occurred last year, when Facebook sent a cease-and-desist demand to the NYU Ad Observatory team—and then shut down the team's access to Facebook services on a privacy pretext, prompting a rebuke from the FTC. While the NYU team chose to stand up to Facebook, legal threats like these are common and create a chilling effect for independent research on online services.

The federal strategy for information integrity R&D is an opportunity to develop a whole-of-government approach toward a safe harbor for this type of independent platform research, which is vital for progress on information integrity challenges. The strategy could, for example, encourage the Department of Justice to reconsider its guidance on the Computer Fraud and Abuse Act, a notoriously ambiguous law that the

[13] *See* Elizabeth Hansen Shapiro et al., *New Approaches to Platform Data Research* (Feb. 2021).

Supreme Court recently significantly narrowed.[14] DOJ's guidance currently only addresses certain types of computer security research, and it has not updated the guidance in light of new Supreme Court precedent. Similarly, the strategy could encourage the Department of Commerce and DOJ to consider advocating for a Digital Millennium Copyright Act safe harbor for this type of research. The DMCA is another broad and vague law that could be interpreted to encompass certain information integrity research, and the Copyright Office conducts an exemption rulemaking every three years where Commerce and DOJ's views have significant weight.

A final potential path forward for methodological innovation in information integrity research is access to data held by online platforms. There have been a number of recent proposals for this approach to platform regulation, including the European Union's Digital Services Act and recent legislation in the House and Senate. We encourage the Working Group to consider how the administration can best engage with these proposals, such as by advocating for U.S. researcher access under the EU DSA or supporting particular legislation in Congress.

4. Federal R&D strategy for information integrity should prioritize building the emerging field of interdisciplinary information integrity research, especially new institutional infrastructure, educational pathways, and ethical frameworks.

The emerging field of information integrity research would benefit from greater coherence. Research and teaching in the area are often fragmented across disciplines, and groundbreaking research often emerges from interdisciplinary and multi-institutional collaborations. We recommend that the Working Group prioritize steps to build the field.

One promising direction is to invest in and facilitate the fundamental institutional infrastructure that is necessary for the field's success: new research centers, convenings, and publication venues that are specific to information integrity, reaching across disciplines and institutions. The Knight Foundation has emphasized this strategy in its recent grantmaking, and we encourage the Working Group to consider a similar model.

[14] Van Buren v. United States, 141 S. Ct. 1648 (2021).

Another important step would be developing educational pathways for students, including mid-career professionals, who may consider a career in information integrity research or practice. Because the field is still taking shape, educational offerings related to information integrity tend to be one-off courses rather than complete sequences of study. The federal R&D strategy could have a significant beneficial impact by supporting the development of new courses, curricula, degree programs, and academic concentrations that will build a pipeline of information integrity expertise.

Our final recommendation is that the Working Group address ethical considerations for the emerging area of information integrity research. Carrying out work in this field can involve collecting and analyzing personal information and can involve examining online platform practices without the platform's agreement. Information integrity research may also implicate speech protected by the First Amendment and may have a political valence. Ethical frameworks for information integrity research are essential, both to provide guidance for the conduct of research and to address foreseeable dilemmas before they occur. The PERVADE multi-institution collaboration, which has explored the ethics of social media research methods and is supported by NSF, is a potential model for how the federal strategy could address ethical considerations.

* * *

Thank you again for the opportunity to provide input to the federal government's strategic plan for information integrity research and development. We would be glad to provide additional detail or discussion as would be helpful to the Working Group.

Sincerely,[15]

Academic Researcher
Rutgers University

Academic Researcher
Center for Information Technology Policy, Princeton University

[15] We offer this comment as individual academic researchers.

Academic Researcher
Cornell University

Academic Researcher[16]
Princeton University

Academic Researcher
University of Washington

Academic Researcher
Stanford University

Academic Researcher
Princeton University

Academic Researcher
Northeastern University

[16] Principal author of this comment.