Political Speech on Social Media Platforms:
An Initiative to Create Better Governance Mechanisms

The Pocantico Center
February 1-2, 2019
Meeting Summary

Introduction

On February 1-2, 2019, the Consensus Building Institute (CBI) convened a group of academic, policy and technical experts at The Pocantico Center in Tarrytown, New York, to discuss the governance of political speech on social media platforms. The meeting was made possible with generous support from the Rockefeller Brothers Fund (RBF). See Appendix A for a list of participants, and Appendix B for a copy of the meeting agenda.

Prior to the meeting, CBI articulated the following goals for the group:
- A shared understanding of how the platforms are functioning today as forums for political speech, with a focus on problems of disinformation and dangerous speech
- A shared vision of what would make social media platforms valuable forums for political speech
- Shared criteria for good governance of political speech on social media platforms, recognizing the wide range of contexts and issues which governance must address
- Strong, testable proposals for improving the governance of political speech on social media platforms in a variety of contexts
- An action plan for moving those proposals forward, and for the future of this group

CBI also distributed a Framing Paper before the meeting, which presented an initial set of ideas, based on pre-meeting interviews with each of the participants and CBI's own research, to help focus the conversation. Specifically, the Framing Paper articulated ideas around the strengths and limitations of the platforms as forums for political speech, existing governance mechanisms for political speech on social media, and options and proposals for improving governance.

Day 1

Strengths & Limitations of Social Media Platforms as Forums for Political Speech

Following a lunchtime welcome from Stephen Heintz, President of RBF, the first full group session focused on the strengths and limitations of the social media platforms as forums for political speech. The discussion addressed the following questions:
- *What are the strengths and limitations of the current platforms?*
- *What are the underlying drivers responsible for these strengths and limitations?*
- *What would characterize a good forum for political speech on social media?*

Two participants provided opening remarks on the platforms as forums for political speech. They noted that the platforms facilitate more two-way communication between the government and the governed, which has the potential, at least, to reinforce values of democratic dialogue. With regard to users, social media may make it easier for users to participate in practices that are destructive rather than constructive towards democracy, because the preference for fast and cheap communication selects for certain types of social action. We need to improve our understanding of what "good" social engagement looks like on social media by identifying the kinds of productive, boundary-crossing conversations we want to support.

Strengths and limitations: Social media as a "double-edged sword"

In the ensuing group discussion, participants confirmed and amplified the Framing Paper's framing of the impact of social media on political speech and action as a "double-edged sword." On the one hand, social media has facilitated an unprecedented level of transparency regarding the actions of government officials, provided access and a voice in the political process to more individuals than at any time in history, and has allowed people to speak from a position of relative safety through the shield of anonymity. On the other hand, over time we have learned that social media has a variety of more pernicious impacts as well:

- Social media may effectively silence "ordinary" people because of the dominance of "loud" voices — both political leaders and other political influencers with large followings.
- Social media algorithms can give rise to tribalism by limiting exposure to cross-cutting views, and have allowed us to be more fragmented digitally than we ever could have become physically.
- Social media promotes engagement based on emotions such as outrage, which can result in emotional contagion and drive users towards more extreme content.
- Paradoxically, social media can be asocial — the nature of the medium encourages behavior that would be naturally checked by in-person interactions. Norms of good and bad behavior online, like "it is bad to re-tweet false news," have yet to be truly established.
- In developing countries in particular, platforms like WhatsApp provide users with large amounts of unfiltered information without providing them the means to appropriately verify or contextualize this information.

Underlying drivers: User psychology, platform design and market incentives

The group also refined its understanding of the underlying drivers responsible for these strengths and limitations. Participants suggested that challenges emerge from the interactions among user psychology and behavior, platform design, and platform business models. Some problems result from manipulation of the platforms by "bad actors," like professional disinformation brokers, while others result from "known bugs" in the system (such as clickbait, polarization, echo chambers, conspiracy theories, extreme speech, viral emotions, etc.). There is uncertainty and debate about the platforms' incentives and capacities to address these

known problems. Though many call for more "effective" content moderation by the platforms, there are also dangers if the platforms are incentivized by public pressure or required by law to moderate speech in an excessively intrusive or heavy-handed manner.

Participants acknowledged open questions and conflicting evidence regarding the psychological impacts of social media and platform features. For example, there is evidence that social media exposes users to an increased diversity of viewpoints, but this exposure triggers stronger negative reactions to contrary views, which reinforces group norms.

Some participants also expressed the need to prioritize between "first order" and "second order" problems. For example, Americans' diminished trust in established democratic institutions is arguably a "first order" problem, because it will not be possible to address problems such as polarization without first resolving the trust issue. Another way to prioritize is to recognize that bad behavior on social media exists on a spectrum. On this spectrum, problems like organized efforts to intervene and disrupt the U.S. political process are especially worthy of attention. Others noted that different problems can become paramount depending on the context. For example, in fragile democracies, speech on social media may be much more likely to lead to violence in the weeks immediately before an election, in which case tighter regulations could be called for.

Characteristics of a good forum: Support for platform pluralism and user choice

Participants generally agreed that there does not need to be (and arguably should not be) one model or standard for speech across all social media platforms. Some contended that there is value in increasing "platform pluralism" and user choice. It would arguably be better for users and for political discourse if there were a greater variety of platform options available, if users had a greater ability to choose among platforms, and if users' social media networks could be ported from one platform to another.

Participants also worked to refine the original "characteristics of a good forum" offered in the Framing Paper. Overall, participants suggested revisions reflecting the following principles:
- With regard to transparency:
    - Anonymity is important for protecting vulnerable speakers and should be permitted. It is important to distinguish between anonymity and the problem of false identities and bots. While these latter issues are serious problems that need to be addressed, it is not clear that eliminating anonymity would be an effective approach. There are ways to detect inauthentic identities that do not reveal the speaker.
    - There should be transparency around paid advertising, platform policies, and strategies for content moderation.
- Platform policies should support fact-based political discourse. Demonstrably false information/disinformation should be removed.

- There is value in providing users with opportunities for constructive engagement across differences, but this is not a necessary characteristic of a good forum. It is legitimate for platforms to support political discourse within like-minded groups.

See Appendix C for the specific "Characteristics of a Good Forum" that emerged from this discussion. At the end of the discussion, participants expressed willingness to support this revised set of characteristics, with some caveats noted in the language reflecting areas where differences of opinion remained.

Governance Mechanisms and Opportunities for Improvement

During the next session, participants discussed existing governance mechanisms for political speech on social media and how they could be improved. The session sought to address the following questions:
- *How is political speech governed now on the platforms? (Who makes the rules? Who interprets and enforces them? What redress exists for those who disagree with the rules and/or their enforcement?)*
- *What are the strengths and limitations of these existing approaches, considering key contextual factors (e.g. state-society relations, platform penetration, political polarization)?*
- *How would we know if we had good governance of political speech on social media platforms?*

Three participants offered opening remarks. One presented the concept of "dangerous speech," defined as any form of expression that increases the risk that its audience will condone or take part in violence against members of another group, and the challenges of governing it. Thus far, internet platforms have relied almost entirely on just a few methods for responding to harmful content (including but not limited to dangerous speech): taking that content down, deleting the accounts that post it, and making harmful content less visible without eliminating it. Since these methods do not diminish the rate at which new harmful content is posted, they are band-aids, and other methods need to be explored. It is important to distinguish dangerous speech from hate speech: dangerous speech may have nothing to do with hate and is often more about spreading fear than hatred.

A second set of remarks explored social media governance in Sub-Saharan Africa. In many African countries, there are large Facebook and WhatsApp groups. The administrators of these groups play influential governance roles as moderators, using tools such as discussion muting functions, kicking people out, etc. Governments across Africa have sometimes chosen to deal with what they consider subversive or otherwise unacceptable political speech within these groups by mandating telecom companies to shut down the internet, as well as by re-introducing laws that allow legal action against individuals who criticize the government.

The third set of remarks highlighted the challenges a global social media company like Facebook faces in governing speech in very different contexts all over the world. Nonetheless, it is clear

that a hands-off approach was no longer viable for the major platforms, particularly when foreign states have interfered in domestic politics. Initiatives for fact-checking (such as NewsGuard) and user education are useful, but on their own they are insufficient to deal with the magnitude of the problem.

Shared governance principles, mechanisms and challenges

To kick off the discussion on this topic, CBI presented the group with a draft set of possible principles of good governance from the Framing Paper, and asked for participants' comments and reactions. The draft governance principles are available in Appendix D.

Overall, participants agreed that there is value in articulating principles of effective governance, and that it is appropriate for principles to be aspirational. However, they suggested that the draft governance principles were too aspirational. To be effective, such principles must also be credible as a basis for practical action, and it is important to articulate both values and realistic mechanisms for achieving them. Currently, there is a need to address the incentives (or lack thereof) that the dominant platforms have to support better governance mechanisms, and the risks of ineffective government regulation.

Participants suggested that it is difficult to systematically improve on the governance status quo for at least three reasons:
- The quality of governance only marginally affects the companies' bottom lines;
- Public pressure to date has had limited impacts; and
- A truly effective governance regime has the potential to impinge on the platforms' ability to innovate.

Overall, participants agreed that the goal should be to develop creative strategies and tactics to influence the dominant platforms and/or promote platform pluralism.

Ways to improve platforms as forums for political speech/democratic discourse

Participants offered numerous ideas for improving social media platforms as forums for productive democratic discourse, as well as incentives and leverage points to address barriers to reform.

They suggested the following broad arenas for improving on the status quo:
- Supporting platform pluralism through requiring network portability and interoperability. This would increase user choice and allow for the emergence of new public interest social media platforms.
- Developing industry ethical standards for the platforms. The Forest Stewardship Council could serve as a model and the European Commission's Code of Practice on Disinformation could be a building block.
- Developing and promulgating professional standards for social media engineers, who are a highly influential group of stakeholders within the platforms.

- Working through various fora for inter-governmental action, including the G20 and the UN Secretary General's High-Level Panel on Digital Cooperation.
- Laws, regulations, and self-regulation to improve U.S. electoral processes, for example through increasing transparency in advertising.
- Focusing on non-U.S. country contexts, such as the EU, which may be more willing to regulate large tech companies, and developing countries, where the risks of dangerous speech and/or government censorship may be heightened.

They also noted a number of incentives and leverage points that could help in addressing these challenges. The major platforms could be motivated to improve governance in order to gain or retain market access; address public pressure (whether expressed through traditional advocacy or through viral online campaigns); pre-empt or shape government regulation; and/or to retain engineering talent. The recent walk-out of Google engineers in response to the company's handling of sexual harassment shows how employee perspectives could provoke change in these companies from the inside. In addition, companies might be motivated to change if improved governance is framed as a positive opportunity to innovate, for example as a means to increase and improve user engagement.

For governments, voter concerns about privacy and polarization could incentivize more effective regulations. Lastly, among users and civil society groups, there is interest in having alternatives to the current platforms available for political engagement and organizing.

Day 2

Options & Proposals for Improving Platforms as Contributors to Democratic Discourse and Practice

Following an evening dinner and discussion, participants reconvened on Day 2 of the program to present and discuss concrete options and proposals for improving social media platforms as contributors to democratic discourse and practice. During an initial full group discussion, participants put forward the following proposals:
- Platform pluralism (discussed above and in more detail below).
- Incentivizing more responsible government representation through a certification and amplification (upranking) program for high-quality candidate forums during elections.
- Addressing externalities of the platform business model through regulations to improve competition, e.g. through merger conditions and conditions on data sharing; raising privacy standards; and requiring more transparency of algorithms.
- Changing the authorities of DoD and the NSA to enable effective, real-time information sharing with the platforms on foreign bad actors.
- Presenting specific proposals to Facebook and lobbying for a multi-stakeholder, collaborative governance system with genuine power sharing.
- Establishing standards for researchers to access platform data.

- Encouraging social media companies to adopt new rules to remove provable false disinformation under clear principles and with a robust appeals process.
- Lobbying for new regulations around advertising transparency and limits to the granularity of targeting (e.g. with regard to politically salient demographic characteristics, and/or minimum sizes for targeted groups).
- Supporting advocacy and organizing for better platform policies and collaborative arrangements within the companies, and with advertisers, celebrities and other influencers.
- Requiring social media platforms to enable network portability, through regulations and/or revisions to the Communications Act.
- Supporting a third-party auditing/oversight system for social media companies around speech, disinformation, and research protocols.
- Articulating a clear set of agreed principles and practices for addressing the problem of disinformation.

Small Group Action Planning

After this initial presentation of ideas, participants were organized into breakout groups to pursue priority options and proposals on three topics: collaborative governance, platform pluralism, and disinformation. Each group was asked to build their discussions and presentations around an action planning framework that asked the following questions:
- ***What*** *is the idea for improving social media platform contributions to political speech/democratic discourse and practice?*
- ***Why*** *does this idea advance the positive value of political speech/democracy?*
- ***Who*** *are the actors who would need to support or engage with this idea to move it forward?*
- ***How*** *will you strategically and/or tactically mobilize the actors needed to move the idea?*

After some time in breakout groups, participants reconvened to present and refine their ideas with the larger group. Below, we summarize each breakout group's main ideas, and feedback on those ideas from the full group discussions that followed.

<u>Collaborative Governance</u>

This group understood collaborative governance as **a set of rules and roles that help diverse stakeholders come together to develop a shared purpose, generate and allocate resources to help them achieve their purpose, implement activities in coordinated ways, and hold each other accountable.** In the context of political speech on social media platforms, the collaborative governance group recognized a core challenge: despite many concerns and criticisms of their self-regulation, the leading platforms have strong financial and practical reasons for resisting collaboration and accountability, and continuing to operate largely independent of other actors.

The group discussed three specific arenas where collaborative governance could help address challenges to political speech. The three ideas are presented below, each with its own "why, who, and how."

1. **Collaboration to promote more constructive forums during US electoral campaigns, using a forum certification and amplification (upranking) system.** This idea, being developed by [The Pluribus Project](#), aims to define criteria for a high quality electoral forum (such as candidate representation, citizen participation, and substantive focus); to create a certification system for forums that meet the criteria; incentivize forum organizers to seek certification and campaigns to participate in certified campaigns; and to create strong incentives for certification and participation by gaining commitments from major social media platforms to uprank posts, pages and other content generated through certified forums.

   Why
   Currently candidate discourse in campaign cycles is driven by issues that mobilize partisan voters, and by the interests of large donors. By promoting forums that ensure representation of all viable candidates, meaningful citizen participation and formats that generate constructive discussion and debate, it may be possible to make discourse less polarized and more focused on the issues that most voters care about.

   Who
   This effort requires collaboration among one or more forum certifying organizations, national and local media and civic groups that host campaign forums, campaign teams, and major social media platforms that would need to commit to upranking certified forums.

   How
   Currently the effort is focused on getting a certification system designed and tested, and on outreach to media, candidates and platforms to encourage participation in experiments to test and prove the concept.

2. **Audit Framework:** The platforms and outside stakeholders could agree to a set of standards for moderating political speech, and then have a third-party body audit and report on the platforms' adherence to those standards. The auditor's access to platform data and its independence from the platforms could be ensured by a set of financial, legal, technical and other guidelines and procedures. It could be feasible and impactful to develop standards for user authenticity, limits on identity-based data targeting, and/or election integrity.

   Why
   As the full group discussed, the platforms face commercial pressures not to adopt rigorous standards for content moderation. On the other hand, shared standards that all the major platforms agreed to, with an independent auditor, could increase the platforms' confidence that they were not putting themselves at a competitive disadvantage. Such an audit framework could also address the concerns of many outside stakeholders to increase the transparency and

accountability of the platforms, while also assuring the platforms that their efforts would be judged fairly, using standards and technical approaches that they themselves agreed to. A globally agreed set of standards could also provide the platforms with additional insulation from authoritarian governments that sought to manipulate the platforms for political ends.

Who
The development of standards would involve the platforms and a global set of stakeholders. The group did not discuss the stakeholder group's composition in detail. However, it did recognize that the platforms would be concerned about the responses of governments and civil society groups that have been critical of their approach to content moderation, especially in regard to political issues.

How
 Given concerns about the platforms' willingness to participate, the group thought that it might make sense to begin with a positive and invitation to the major platforms to discuss their successes in content moderation, the standards that they have been using, and opportunities that they see to verify the effectiveness of their current standards and their implementation. This dialogue could evolve into the development of collaborative governance of an audit framework.

3. **Independent research protocol:** The platforms and outside researchers could agree to a set of principles and guidelines for independent research. The starting point would be agreement between the researchers and one or more platforms on a set of research questions, hypotheses and methods. The protocol could then include the following elements:

   - Independence of the researchers, including access to all information needed to answer the agreed research question(s), and freedom to publish in a peer-reviewed academic journal without prior review by the platform, as long as the publication does not reveal technical trade secrets
   - User privacy, ensured by use of anonymized, aggregated information, with deletion protocols
   - Transparency and accountability, ensured by having research plans reviewed and approved in advance of the research by university research oversight committees, and or by a qualified, neutral third party body (such as the called the Open Science Framework.
   - Conformity with ethics standards for research, either through review by a university ethics body or by another qualified, independent ethics body.

Why
The platforms have restricted researcher access to data, and have sometimes released studies by their own staff that independent researchers have been unable to verify. Recent efforts to organize larger scale collaboration between platforms and researchers have not fully addressed these limitations. In both situations, the value and credibility of the research products have

been open to question. A research protocol that addresses concerns about independence, user privacy, transparency, accountability and ethics as well as platform concerns about confidentiality of their technical secrets could be a powerful tool for overcoming these challenges.

Who and how
One of the participants has already developed a research protocol with the characteristics named above, and a major platform has agreed to use the protocol for a specific research project. If the protocol proves effective and is well publicized, it could be voluntarily adopted by other researchers and platforms. There are some staff at the platforms who have research backgrounds and might support such a protocol as a good way to respond to current concerns about the way research is being conducted. Over time, as it comes into wider use, such a protocol might become an informal "best practice" standard for research on social media platform impacts. It could also become institutionalized if one or more platforms adopt it as "their" protocol for working with independent researchers, individually or in larger collaborations.

Participants offered several comments and suggestions during the full group discussion.
- With regard to the audit framework, other initiatives, such as the Global Network Initiative, have tried to develop shared standards for audits of the platforms, with very limited success. It may be more effective to undertake fully independent audits using publicly available data than to seek the platforms' agreement on audit standards or methods. Even government regulators such as the Federal Trade Commission, which required Facebook to hire an independent auditor to monitor compliance with its privacy commitments, have not effectively enforced this audit requirement.
- It might, however, be possible to get the platforms to agree to create a very large panel (a set of anonymized users whose platform activity could be followed over time) that could be used for multiple research and audit purposes. Such a panel would require the technical cooperation of the platforms, but would not necessarily require them to agree on specific standards or questions to be answered. With a panel approach and an agreed research protocol for access to it, a great deal could be learned.
- Ultimately, the point of a collaborative governance approach should be to create effective third-party oversight over the platforms. It needs to be collaborative because the platforms are powerful, but it must also be credible and effective in achieving goals that the platforms have not been advancing.

Platform Pluralism

What
The platform pluralism group defined its goal as follows: ***Successfully foster the development of new social media platforms that are plural in purpose and rule set, public in spirit, and participatory in governance.*** Potential purposes for the new platforms could include political or community organizing, communicating across difference, fostering community ties, collaborative problem-solving, or any number of other purposes.

In addition to the development of the new social media platforms themselves, group members noted the following two additional, necessary sub-goals for their project to succeed:

- A tool to merge content from multiple social media platforms using APIs and a shared name space. This could be a kind of social media "browser" or "aggregator."
- A new policy framework and citizen movement that would support platform interoperability, aggregation rights, and stronger anti-trust regulation/enforcement.

Why

Platform pluralism would increase user choice, platform innovation, and competition. Civic discourse could be improved by the existence of successful platforms whose basic mission is to support public values, like PBS or NPR do for television and radio. Platform pluralism would facilitate innovation among users and social entrepreneurs interested in fostering pro-democratic speech and action. The non-profit Cortico is an example of an organization working to build the kinds of new civic space for public conversations that the group wants to foster.[1] At the same time, a side-effect of pluralism would be more platforms that attract users who engage in hate speech, like gab.

How

The group noted that a proof of concept for this idea exists in the online tool Gobo, which allows users to choose the algorithms/filters that control their social media feeds. They discussed the following possible ways forward:

- A policy campaign around interoperability, anti-trust enforcement, privacy reforms, and the need for social media diversity. The campaign could draw on traditionally American themes, and revolve around empowering people to leave a platform "and take their friends with them." Alternatively, the campaign could begin in a place like Scandinavia or France, where the environment might be more susceptible to immediate policy change.
- Forming a National Commission on Communications Policy in the U.S.
- Pro-competition actions to prevent dominant incumbents from using their market power to absorb or block new platforms.
- Using blockchain technology for identify verification.
- A lawsuit to articulate/establish consumer rights.
- Seeking to de-escalate the fight with existing incumbents through building something analogous to the broadcaster-affiliate model, where the parties' financial interests are aligned.
- A proposal to Michael Bloomberg and his associated philanthropies for public purpose social media platforms linked to cities. This could include a participatory budgeting component, with crowdsourcing and other forms of online collaboration for monitoring and improving urban spaces.

---

[1] A non-profit in cooperation with the MIT Media Lab, Cortico builds systems that bring under-heard community voices, perspectives and stories to the center of a healthier public dialogue.

The group discussed the idea of a National Policy Commission on Communications Policy in more depth. They suggested that such a group could take advantage of lessons learned from the [Knight Commission on Trust, Media and Democracy](#) to identify clear problem statements and a plan of action to address them. Such a commission should be independent, but would ideally be backed by Congress with a goal to identify immediate, short-term policy ideas. Areas of focus could include the First Amendment, Section 230 of the Communications Decency Act, and the Fairness Doctrine. It would also be important to discuss a format/standard for data export to support the goal of platform pluralism.

Who
To move these ideas forward, the group suggested it will be important to identify more partners to work on the technical side of aggregation problem, including an export protocol and format. It will also be key to identify partners for policy and advocacy. This could include a policy lead who can help develop a platform, and initiate a campaign for the formation of plural, participatory, public networks.

For the National Policy Commission idea, specifically, the group suggested it would be key to identify champions in Congress. These could include Senators Warner, Markey, Schatz, and Blumenthal. Other key partners could include Sir Tim Berners-Lee, the Center for Democracy and Technology, and the Hewlett Foundation.

Full Group Feedback
During the full group discussion, participants offered the following comments and suggestions on this idea, organized by theme.

Feasibility:
- There is a long history of new platforms being subsumed by the dominant players (for example, Tweetdeck was a social network aggregator that was acquired by Twitter). It will be key to strengthen competition policy to prevent this tactic from succeeding.
- It is important to consider if there is a demand for platforms that serve these purposes. Past efforts that have tried to serve the public interest, like [Nextdoor](#), have struggled.
- Politically, it may be very difficult to bring these policies forward if the existing platforms identify them as threats to their business model, or if the reforms seek to establish a right to display the existing platforms' content.
- Sequencing is important. If the right regulations are not in place, it will not be effective to try to fund new social networks to compete with Facebook and Twitter.

Strategy:
- It may be helpful to consider organizing multiple smaller platforms to provide a counterweight to the dominant players. A platform like LinkedIn might support these kinds of changes and help with advocacy.
- The idea of interoperability suggests there is a kind of individual, digital right at stake. There could be an effort to extend the idea of network rights to include one's "social

graph," which might help shift the balance of power in any public debate on these issues.

Impact:
- It is not clear that having more social media platforms would help with security. In systems with distributed structures, the weakest link can bring the whole system down.
- On the other hand, platform diversity would make it harder to target the whole U.S. population through just one platform.
- To mitigate potential negative impacts, efforts to increase inter-operability and platform diversity would need to go hand in hand with increased privacy protections.

Stopping Disinformation

What
In order to address disinformation on existing platforms, this group first defined "disinformation" as "campaigns designed to intentionally spread demonstrably false information." The group named its goal as ***developing ways to identify real sources of both foreign and domestic disinformation***.

Why
The promulgation of disinformation via social media platforms is destructive to democracy. Creating incentives, developing mechanisms, and building capacity to identify and eliminate sources of disinformation would mitigate the significant damage that social media has inflicted on global democratic rights, processes, and institutions. Addressing this issue directly could set new precedents for norms of "good behavior" for users and strong governance mechanisms for existing and future platforms.

Who
The disinformation group identified the following categories of actors needed to address disinformation on existing platforms:
- Government(s)
- Platforms
- Academia
- Advertising industry
- Media

How
For each category of key actors, the group discussed necessary actions for addressing the disinformation problem.

Government actors should act on, amend, or develop laws and regulations that incentivize the removal of disinformation (e.g. Titles 10 and 50, and Section 230 of the Communications Decency Act). Other key actions include improving norms for information sharing on

disinformation threats among government and the platforms (private sector); increasing investment into tools for detecting disinformation; and mandating the use of tools that assess source credibility (e.g. NewsGuard, the Pluribus Project, etc.).

Platforms, in tandem with the other key actors, should consciously increase transparency surrounding their privacy and security settings, ad sources, blocking capabilities, various forms of microtargeting, and the uses of the data this generates. Platforms should also stop fighting legislation like the Honest Ads Act (with suggested regulation through FTC or FCC) and focus on threat prioritization to increase their abilities to be proactive, rather than reactive in combatting disinformation campaigns. They should consider developing a position like Chief Advertising, Behavior, or Content Officer, whose primary role would be to identify actors spreading disinformation and their campaigns, seek to stop them, and mitigate their impacts.

The advertising industry should demand more transparency of the platforms and facilitate the empowerment of users to block unwanted ads on platforms. The advertising industry should view the platforms as "information fiduciaries" and treat them accordingly. Industry associations such as the Interactive Advertising Bureau should help with the development and implementation of oversight mechanisms.

Academics should use research to address disinformation in a proactive way. Researchers should conduct a pre-emptive (vs. reactive) analysis of 30-50 countries identified as "high-risk" to investigate the role of disinformation.

In addition to the above actions, it is also important to generate external pressure through traditional media (via op-eds, published reports, etc.) and a lobbying effort.

Full Group Feedback
- For the platforms, addressing disinformation in a more comprehensive way would be costly. It would mean hiring a lot more people, as well as increased public relations and liability risks. It is not clear where the leverage to incentivize such a change would come from.
- Policymakers often seek ideas from external sources, like members of this group, that they can champion. Members of this group should be ready to advocate around ideas like a Code of Ethics.
- Government officials and even some platform representatives repeatedly misinterpret Section 230 as *limiting* the platforms' ability to moderate speech, when in fact Section 230 *protects* and *authorizes* content moderation by the private sector. This message needs to be conveyed to officials and others more clearly.
- Within the platforms, there is a disconnect between the various teams whose work touches on disinformation. To effectively take down accounts used to spread disinformation, there needs to be a separate, responsive unit within each of the platforms. Within Google, Jigsaw could be an example of such a unit.

Future of the Initiative and Final Reflections

Ideas to take forward

Following small group report outs, CBI asked participants to comment on opportunities to take an idea forward collaboratively or otherwise. Participants expressed interest in exploring the Code of Ethics model, the idea of platform pluralism, and a longer-term project that could address communications policy in the 21st century (e.g. First Amendment and/or Section 230). Participants emphasized the need to expand the discussion to include key policy makers (and potentially representatives of agencies with national security and intelligence mandates), social media "influencers" or celebrities, former employees of the major platforms, and other relevant industry members from companies such as Wikipedia, Microsoft, IBM, and Salesforce. These voices could help structure and guide the direction of the initiative, and increase the legitimacy of future recommendations or ideas.

Final reflections

CBI facilitators David Fairman and Toby Berkman concluded the meeting by asking participants for their final reflections on the discussion and the meeting as a whole. Participants expressed gratitude for the opportunity to meet, excitement at the potential to continue working on these issues with new colleagues, and renewed determination, despite the challenges, to continue both collaborative and individual work on improving the governance of political speech on social media platforms.

Appendix A: Participant List

Name	Title	Affiliation
Jonathan Albright	Director of the Digital Forensics Initiative	Tow Center for Digital Journalism, Columbia University
Susan Benesch	Executive Director	Dangerous Speech Project
Toby Berkman	Senior Associate	Consensus Building Institute
Chipo Dendere	Professor	Amherst College
Eric Effron	Editorial Director	NewsGuard
David Fairman	Managing Director	Consensus Building Institute
Dipayan Ghosh	Pozen Fellow & Co-director of the Platform Accountability Initiative	Harvard Kennedy School
Mike Godwin	Distinguished Senior Fellow	R Street Institute
Gordon Goldstein	Adjunct Senior Fellow	Council on Foreign Relations
Stephen Heintz	President	Rockefeller Brothers Fund
Niloofar Razi Howe	Security Investor & Entrepreneur	
Young Mie Kim	Professor	University of Wisconsin-Madison
David Kirkpatrick	Founder	Techonomy Media
Michael Posner	Professor	NYU Stern Center for Business and Human Rights
Deb Roy	Associate Professor, Media Arts & Sciences; Co-Founder, Chairman of the Board	MIT Media Lab; Cortico
Richard Stengel	Contributor, MSNBC; Senior advisor, Snapchat	See previous
Lucas Welch	Executive Director	Pluribus Project
Ethan Zuckerman	Director, Center for Civic Media	MIT

Appendix B: Meeting Agenda

Friday, February 1

Lunch will be available in the Dining Hall starting 11:30am

12:00pm   Connecting Question and Introductions (over lunch)
                      Welcome from Stephen Heintz, President, Rockefeller Brothers Fund

12:30   Strengths & Limitations of Social Media Platforms as Forums for Political Speech
                      Introductory remarks
                      Discussion:
- *What are the strengths and limitations of the current platforms?*
- *What are the underlying drivers responsible for these strengths and limitations?*
- *What would characterize a good forum for political speech on social media?*

1:45   Existing Governance Mechanisms for Political Speech on Social Media
                      Introductory remarks
                      Discussion:
- *How is political speech governed now on the platforms? (Who makes the rules? Who interprets and enforces them? What redress exists for those who disagree with the rules and/or their enforcement?)*
- *What are the strengths and limitations of these existing approaches, considering key contextual factors (e.g. state-society relations, platform penetration, political polarization)?*
- *How would we know if we had good governance of political speech on social media platforms?*

3:00   Break

3:15   Options & Proposals for Improving Governance
                      Introductory remarks
                      Discussion:
- *What ideas/innovations have been proposed to address the limitations of existing approaches?*
- *What are the most promising opportunities for improving governance, in different contexts?*

5:00   Break

5:45   Reception

6:30   Dinner
- Connecting questions at tables

8:00   After-dinner conversation in the Hayloft

Saturday, February 2 (Revised)

9:15   Participants present propositions for improving platforms as contributors to democratic discourse and practice; discussion to clarify propositions

10:15   Cluster propositions, organize breakout groups

10:45   Break

11:00   Breakout groups meet

12:45   Working lunch with reports from breakout groups and feedback

1:45   Continued feedback to breakout groups

3:15   Future of the initiative and this group

3:40   Final reflections

4:00   End

Appendix C: Revised Characteristics of a Good Forum for Political Speech

1. **Open and accessible to anyone who wishes to participate and who accepts the platform's standards for speech, assuming shared governance**
   - Based on an assumption of platform pluralism and user choice: there is value in having many platforms (with different standards), among which users can choose; to have portability of social networks across platforms; and to enable users to personalize their experience of engagement in political discourse
2. **Transparent, while protecting vulnerable speakers.**
   - [Transparency of identities,] with enforced prohibitions on false identities and bots
   - Transparency of paid advertising/posting and platform use of user data for advertising
   - Transparency of policies and strategies[algorithms] that moderate political speech (including financial metrics and rules)
3. **Designed and moderated to promote constructive experiences and discourse**
   - Including mechanisms for identifying and removing false content
   - Encouraging and facilitating constructive engagement across diverse and differing views
   - Including a transparent record of moderation decisions
4. **Based on shared governance.**
   - Governance should be nested, to support localism and user-determined moderation within groups
   - Governance should promote public interests in informed discourse on core democratic processes (e.g. elections, legislative and exec. processes), while preserving free speech
   - Recognizing and supporting the appropriate role of government for specific categories of "bad actions/actors" (but recognizing risk of governments as bad actors)
   - Ensuring accountability of platforms and users for enforcement of agreed standards, including cross-platform mechanisms for accountability

Appendix D: Potential Principles of Good Governance

1. The platform's standards for political speech, the interpretation and implementation of the standards, and the processes for complaint and appeal, should be <u>jointly developed and overseen by platform managers and user representatives who are diverse in their political allegiances, social identities, and economic positions</u> (e.g. including advertisers, individual users, news and opinion media, organized groups with a range of views on political speech, and others).

2. The participants in governance, and the processes of standard setting, interpretation, implementation and appeal <u>should be transparent, and there should be periodic opportunities for online consultation and input</u> on governance issues.

3. Governance <u>should be multi-level, in the sense that regional and/or national stakeholders should be engaged in an ongoing process of translating and applying global standards</u> so that they are appropriate to the specific societal contexts and risks in which the platform operates.

4. <u>Governmental bodies with regulatory authority should be consulted and national laws respected, as long as their requirements do not conflict with internationally agreed human rights and rights to freedom of opinion and expression</u>; in cases of conflict, ongoing dialogue should be used to address governmental concerns.