



HARVARD Kennedy School

MOSSAVAR-RAHMANI CENTER
for Business and Government

Recommendations to the Biden Administration: On Regulating Disinformation and Other Harmful Content on Social Media

Caroline Atkinson

Peterson Institute for
International Economics

Paul Barrett

New York University

Lynda Clarizio

Nielsen U.S. Media

Dipayan Ghosh

Harvard Kennedy School

John Haigh

Harvard Kennedy School

Thomas Melia

PEN America

Michael Posner

New York University

Vivian Schiller

Aspen Digital

Clint Watts

Foreign Policy Research
Institute

2021

M-RCBG Faculty Working Paper Series | 2021-02

Mossavar-Rahmani Center for Business & Government
Weil Hall | Harvard Kennedy School | www.mrcbg.org

The views expressed in the M-RCBG Associate Working Paper Series are those of the author(s) and do not necessarily reflect those of the Mossavar-Rahmani Center for Business & Government or of Harvard University. The papers in this series have not undergone formal review and approval; they are presented to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). Papers may be downloaded for personal use only.



Recommendations to the Biden Administration *On Regulating Disinformation and Other Harmful Content on Social Media*

March 2021

EXECUTIVE SUMMARY

This white paper recommends a range of steps the Biden Administration should take to counter disinformation and other harmful content on major social media platforms. In recent years, the spread of disinformation online has eroded crucial democratic institutions and discourse, especially in connection with elections and with disproportionate impact on underrepresented communities. The Administration should move swiftly to address this threat in a variety of ways.

The recommendations fall into six categories:

I. *Industry standards and regulatory infrastructure*: The social media industry has not developed adequate standards and processes for curtailing disinformation and other harmful content. Moreover, no existing government body pays sustained attention to social media. In light of these gaps, the Administration should work with Congress to create such a regulatory body, possibly as a new Digital Bureau within the Federal Trade Commission. Authorizing legislation could require the industry to collaborate with the bureau to develop industry standards of conduct, which the bureau would then enforce.

II. *Platform liability and incentives for more vigorous content moderation*: Section 230 of the Communications Decency Act of 1996 needs to be updated. We recommend that the Administration collaborate with Congress to retain the law's liability shield for social media platforms but add important exceptions, or "carve-outs," for certain areas, such as civil rights infractions and cyber-stalking, where the shield would not apply. Limiting the shield in this manner would incentivize platforms to police those areas more vigorously. Modifications of Section 230 would need to be rationalized with the industry standards outlined in Section I.

III. *Executive branch actions*: In some areas, the Administration can act without Congress to improve collaboration between the Executive Branch and industry. For example, the Administration should encourage social media companies to participate more energetically in information-sharing programs, with a commitment to disseminate corporate intelligence on foreign and domestic disinformation activity. The industry also should provide this intelligence to federal law enforcement and intelligence agencies.

IV. *Financial incentives to encourage desirable company behavior*: The Administration should work with Congress to develop a system of financial incentives to encourage greater industry attention to the social costs, or “externalities,” imposed by social media platforms. A system of meaningful fines for violating industry standards of conduct regarding harmful content on the internet is one example. In addition, the Administration should promote greater transparency of the placement of digital advertising, the dominant source of social media revenue. This would create an incentive for social media companies to modify their algorithms and practices related to harmful content, which their advertisers generally seek to avoid.

V. *Transparent advertising*: We recommend that the Administration push for an enhanced version of the previously introduced Honest Ads Act. Rather than focus only on online political advertising, the act would apply new disclosure requirements to all advertising. This would obviate the need for endless debate about what constitutes a “political” ad.

VI. *Support for credible local news organizations*: The Administration should take steps to strengthen credible news organizations, especially at the local level, because of their importance to the functioning of our democracy. The reporting done by these outlets serves as a crucial counterweight to disinformation. But over the past 15 years, social media companies have siphoned off a huge portion of the advertising revenue that had sustained local journalism. The Administration should develop and support legislation that would help local news outlets survive.

•••

The following people developed the recommendations in this paper:

Caroline Atkinson

Member of the Executive Committee of the Board of the Peterson Institute for International Economics

Paul Barrett

Deputy Director of the Center for Business and Human Rights at the Stern School of Business at New York University

Lynda Clarizio

Former President of Nielsen U.S. Media

Dipayan Ghosh

Co-director of the Digital Platforms and Democracy Project at Harvard Kennedy School

John Haigh

Co-director of the Mossavar-Rahmani Center for Business and Government at Harvard Kennedy School

Thomas Melia

Washington Director, PEN America

Michael Posner

Director of the Center for Business and Human Rights at the Stern School of Business at New York University

Vivian Schiller

Executive Director of Aspen Digital, a program of the Aspen Institute

Clint Watts

Distinguished Research Fellow at the Foreign Policy Research Institute and
Non-resident Fellow at the Alliance for Securing Democracy, German Marshall Fund

•••

The participants bring different backgrounds and knowledge to the analysis of these important challenges. All agree on the need for the Administration to address disinformation and other forms of harmful content — and on the general direction of the proposals that follow. But not all members of the working group agree with every aspect of the recommendations. Moreover, the views expressed here are those of the individual participants and do not necessarily reflect the views of their organizations.

PREFACE

The authors of this white paper share a commitment to addressing disinformation, misinformation, and other harmful content on the most prominent social media sites, including Facebook, YouTube, and Twitter. By disinformation, we mean content that is intentionally false and designed to influence public opinion in a harmful way. Purveyors of misinformation, by contrast, may not know that they are spreading false content. Other forms of harmful content include hate speech and incitement to violence. All of these are problems that have assumed much greater urgency in recent years, as a variety of actors have spread rank falsehoods online, distorted the truth, disproportionately harmed communities of color, and contributed to an increasingly polarized society. The internet did not create these deep divisions, but it is dangerously amplifying them.

While these problems are global in nature and require a coordinated international response, our recommendations are directed primarily towards the U.S. government. Regulation of social media platforms by the government will have an impact both domestically and abroad; therefore, policy makers in the United States need to recognize these global consequences as they develop new laws and regulations.

Some members of our group have worked for social media and information technology companies or in the news media. Others have served in the U.S. government. We share a commitment to an open internet, one that promotes free expression and contributes to society by enhancing communication and the sharing of knowledge and information across borders. Social media platforms play an important role in advancing education, promoting commerce, giving voice to the disenfranchised and oppressed, and allowing for political engagement by people with divergent views. This sharing of information and ideas is the lifeblood of a democratic society.

The problem we seek to address is the escalation of disinformation; hate speech, including racism and misogyny; and other harmful content, which now have an outsized influence on social media platforms. The social media companies need to be at the center of any effort to mitigate this problem, in part because they alone have real-time access to material that appears on their sites and the capacity to quickly identify and then downgrade or remove harmful content. But because the leading companies have not taken appropriate responsibility for correcting the problem, we recommend significant changes be enacted by the U.S. government to create stronger incentives for companies to act — measures the government is uniquely suited to undertake.

The main incentives currently prompting social media companies to address political disinformation and other harmful content include: (1.) a sense of social obligation, which too often has been overridden by their business models and drive for growth, and (2.) the threat of legislation and other regulation, which to date has not been sufficiently plausible to induce the companies to make meaningful changes.

These issues will not be addressed effectively by the companies making promises through their corporate social responsibility programs. The changes that are needed must be part of an enhanced regulatory and governance framework.

We make these recommendations while mindful of limits on federal action that would restrict free speech. Outside of a few narrowly defined exceptions, the First Amendment to the U.S. Constitution forbids the government from regulating the content of speech. The government's role also is limited in a practical sense in that it does not have access to problematic content in a timely manner and lacks the technical means to take corrective actions — for example, by making adjustments to algorithms.

Our recommendations focus on six areas: (1.) creation of industry standards of conduct and a new regulatory infrastructure for oversight of the social media industry and, more broadly, the commercial internet; (2.) amendment of Section 230 of the Communications Decency Act to incentivize more vigorous content moderation; (3.) measures the Biden Administration can take using its existing independent executive authority; (4.) enactment of new financial incentives, such as fines, to encourage desirable reforms; (5.) adoption of an expanded version of the previously introduced Honest Ads Act; and (6.) support for credible local news as a counterweight to disinformation.

The paper does not deal with two other vitally important topics—privacy and competition policy— each of which deserves a detailed examination of its own. Some of the authors [have written](#) separately about privacy, competition, and disinformation.

I. INDUSTRY STANDARDS AND REGULATORY INFRASTRUCTURE

The extent of political disinformation, hate speech, and other harmful content illustrates that the social media industry has not done enough to police itself. Specifically, the leading social media companies have not developed standards and processes for addressing harmful content that recognize the broader social harms caused by their activities. As a result, this industry requires greater governmental oversight.

For comparison, the Federal Communications Commission oversees telecommunications, radio, and television. The Securities and Exchange Commission maintains efficient, transparent equity markets. But no existing government body pays sustained attention to social media. This needs to change. The objective should be to prioritize the public interest over corporate interests and address the societal harms occurring as a result of the activities of social media platforms.

Options include creating a new regulatory agency or strengthening and better coordinating the activities of an existing agency. Although the creation of a stand-alone agency would be ideal, political obstacles to such an initiative would be considerable and unlikely to be overcome in the short term. Instead, we recommend enhancing the authority of the Federal Trade Commission to oversee the commercial internet, including social media companies.

The Administration should work with Congress to pass legislation that requires the industry to work with a new Digital Bureau within the FTC to draft industry-wide standards of corporate conduct, which the government would have ultimate authority to approve and enforce. The standards would define the level of reasonable care expected of social media companies in addressing demonstrable harms. They could include transparency standards for how social media algorithms rank and recommend content, limits on the prevalence of various forms of harmful content, and minimum protections of user privacy.

The legislation could empower the FTC's Digital Bureau to directly enforce the industry standards. Alternatively, the legislation could require individual companies to incorporate the standards into their terms-of-service agreements with users. Then, if the companies fail to observe the standards, the bureau would have the authority to initiate an enforcement action under Section 5(a) of the FTC Act, which prohibits "unfair or deceptive acts or practices in or affecting commerce."

Beyond enforcement authority, the FTC's Digital Bureau would provide a locus for robust, ongoing inquiry and analysis of industry trends. This capacity would allow the bureau to see over the horizon in a sector that is continually changing. For example, the bureau would work with other agencies, such as the Federal Bureau of Investigation, to identify new forms of disinformation, including incitement to violence related to political extremism. At the same time, the FTC's Digital Bureau would work with the FTC's Bureau of Competition and the Antitrust Division of the Department of Justice to identify nascent threats to competition.

Over time, policy makers should consider transforming the FTC's Digital Bureau into a fully independent Digital Platform Agency. Doing so would underscore the importance of overseeing powerful and influential social media companies — much as the creation of the FCC and SEC in the 1930s signaled the urgent need for oversight of broadcast media and equity markets.

II. PLATFORM LIABILITY AND INCENTIVES FOR CONTENT MODERATION

Having explained in the previous section our main recommendation for affirmative regulation of the core business practices of the social media industry, we now turn to Section 230 of the Communications Decency Act. We believe Section 230 should be preserved but amended so that, in addition to providing a liability shield against certain civil claims, it creates incentives for the companies to police their sites more vigorously. In this sense, Section 230 would become an adjunct to the broader regulatory infrastructure, standards of conduct, and enforcement activity we recommended in Section I.

[Section 230 \(c\) \(1\)](#) states: “No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.” This means that platforms are not held liable in court for harm related to content posted by users. The statute has exceptions, or “carve-outs,” to this legal immunity. Platforms are potentially liable for content that violates federal criminal law, sex trafficking laws, and intellectual property law.

Enacted in 1996, Section 230 responded to the needs of fledgling companies operating via a then-new commercial internet. Section 230 (c) (1) protected these companies from potentially crippling litigation over what their users did and said online. The law’s other main provision, [Section 230 \(c\) \(2\)](#), protected the same companies from liability related to their decisions to remove objectionable content. Together, the two parts of Section 230 provided a shield against claims related to decisions to leave content up or take it down. Over time, the courts interpreted the law expansively, prompting some observers to argue that social media companies were using Section 230 to deflect legitimate criticism of their content moderation practices. Another complaint about the law was that, while it once protected fragile new entrants to the digital marketplace, in more recent years, it has served the interests of some of the most powerful and profitable enterprises in the U.S. economy.

Most members of our group believe that Section 230 (c) (2) ought to remain as an incentive for platforms to remove problematic content. But we recommend amending (c) (1) to limit the reach of liability protection when platforms leave up allegedly harmful content posted by users. Specifically, we suggest following the existing structure of Section 230 (c) (1), which includes carve-outs to legal immunity. As noted, the existing carve-outs make platforms legally responsible for civil claims related to federal crimes, sex trafficking, and intellectual property violations — and therefore, presumably more vigilant about removing offending content in those areas. We recommend expanding the kinds of deleterious subject matter for which platforms would be potentially liable. This, in turn, would incentivize the platforms to be more energetic and proactive in removing harmful content in those areas.

We propose extending the list of carve-outs to incorporate those found in the recently introduced [SAFE TECH Act](#), including civil rights violations, targeted harassment, cyber stalking, wrongful death, and paid advertising. [Others](#), including certain members of this group, have urged the inclusion of incitement to violence, hate speech, and disinformation. This approach would modify Section 230’s protection against legal liability without removing it completely. It would provide incentives for social media companies to take needed self-regulatory steps without opening the flood-gates to litigation. Over time, the effects of the modifications of Section 230 would need to be tracked and possibly rationalized with the standards of conduct developed by regulatory agencies and companies as outlined in Section I.

III. EXECUTIVE BRANCH ACTIONS

While it works with Congress to pursue the recommendations in Sections I and II, the Administration can take other steps unilaterally or in collaboration with industry. These include:

Organizational clarity: The Administration should clarify which agencies have responsibility for addressing disinformation and other harmful content. Once these responsibilities are mapped, the Administration should identify what regulatory authority already exists and can be exercised immediately, without additional legislation. Mechanisms to coordinate across agencies need to be established. Finally, the Administration should identify key objectives for each of the responsible agencies, as well as processes for them to report on their progress.

Disclosure of operator attribution: The Administration and industry should determine how social media platforms ought to identify and disclose the nature and scope of disinformation operations. These discussions should explore the appropriate threshold needed, in terms of level of confidence in the existence and prevalence of these operations, and to whom the disclosures should be made — law enforcement, the intelligence community, the public.

Security and defensive response standards: The Administration should work with social media companies to develop industry-wide standards and procedures for how the industry ought to respond to known or suspected disinformation operations, whether domestic or foreign. These standards and procedures could be incorporated into the broader code of industry conduct approved and enforced by the new government regulatory authority discussed in Section I.

Information sharing between industry and government: Companies should be encouraged to participate more vigorously in information-sharing programs, with a commitment to disseminate corporate intelligence on foreign and domestic disinformation activity. The industry also should provide this intelligence to federal law enforcement and intelligence agencies. Together, the industry and government should create a national repository of publicly available data on malign content removed by the companies for use by researchers. Stepping up activity of this sort will require careful attention to individual privacy and civil liberties.

Transparency of corporate policy: The companies should be publicly accountable for any corporate counter-disinformation policies and practices to which they have committed, such as Facebook's initiative to take down content promulgated by the Russian Internet Research Agency. Aspects of these practices that cannot be made public should still be disclosed confidentially to the government.

Algorithmic and content commitments: Companies should be encouraged to continue making voluntary commitments that bear on disinformation and other harmful content. Such commitments include Twitter's ceasing the sale of political advertising, given its inability to adequately address disinformation in advertising, and Google's limitation of targeted political ads.

Presidential commission: Even as the Administration maps existing responsibilities within the Executive Branch, it ought to create a six-month presidential commission to refine the government's agenda and make further recommendations.

IV. FINANCIAL INCENTIVES TO ENCOURAGE DESIRABLE COMPANY BEHAVIOR

Social media platforms have positive effects on society. They facilitate communication, education, and healthy political organization. But they also impose social costs, or “externalities,” currently not accounted for in their corporate or financial decisions. These costs include disinformation, hate speech, and other harmful content. We propose altering the financial incentives for platforms that fail to address these harms. Reducing the financial returns social media companies receive if they spread harmful content will encourage them to address the harms.

If corporate activities impose costs on society, a simple way to address this “market failure” is to charge the corporations for the external costs they impose. This is often the least invasive way to address such social costs. Moreover, this response incentivizes the desired behavior from the companies without extensive government intervention into specific decisions by firms and consumers.

Establish fines for violations of newly articulated industry-wide standards: Once industry standards of conduct are created, as discussed in Section I, fines would be established for violations. This approach requires clarity on standards, with appropriate measures to determine the nature and magnitude of harms and the associated fines. It would include developing an accounting system to track firm performance against the standards. For example, for financial accounting, we have GAAP. For issues related to environment, social, and governance (ESG), relatively new standards are evolving. A similar effort to develop metrics for measuring the proliferation of harmful online content could be developed. This could incorporate as a metric the notion of “prevalence” of offending content, meaning the number of times users would encounter the content in question per, say, 10,000 views. Social media companies would need to apply the accounting system and provide transparency of activities and measures within that system, much as they do with GAAP reporting.

Another form of incentive — digital services taxation — has generated discussion and enactment of legislation, both abroad and in certain U.S. states. The issue of digital taxation is beyond the scope of this report. Further research is needed to understand the effects such taxes have on social media companies and whether they create incentives for the companies to address harmful content.

Foster greater transparency on ad placement: Another salutary incentive could be created in the digital advertising market. Social media companies derive the vast bulk of their revenue from advertisers. The companies that buy advertising do not like their ads placed alongside controversial content, including disinformation and hate speech. The Administration should foster greater transparency related to how social media companies place ads on their platforms. This would encourage social media platforms to modify their algorithms and practices to diminish the amount of harmful content presented adjacent to advertising — as well as the amount shown to users.

V. TRANSPARENCY IN ADVERTISING

The Administration should urge Congress to adopt a strengthened version of the Honest Ads Act. As introduced in past sessions of Congress, the Honest Ads Act [seeks to impose](#) on social media companies the political advertising disclosure requirements that already apply to the broadcast industry. We recommend that the Administration work with lawmakers to pass an enhanced version of the Act, which would require that all advertising online be covered by this disclosure provision.

Under this enhanced approach, lawmakers, regulators and social media companies would not have to address the difficult challenge of defining what constitutes a “political” ad. The more robust legislation we recommend would include a requirement that the true and original source of funding for each ad be disclosed. Failure to make these disclosures should be punishable by a substantial fine.

The FTC should have regulatory jurisdiction to oversee compliance with this law, as it has greater capacity to do so than the Federal Election Commission. The FTC already has started to take steps to crack down on social media influencers who do not disclose paid sponsorships of products they promote, so overseeing advertising compliance would be a logical extension of that work. The new Digital Bureau of the FTC we proposed in Section I would take on oversight of advertising.

An enhanced Honest Ads Act would thwart efforts by the social media companies and others to limit the scope of advertising disclosure and define the term “political” as narrowly as possible. By having this disclosure requirement apply to all ads, such attempts to limit transparency would be moot. Broader legislation also would provide a vehicle for other improvements in advertising regulation, such as requiring that forwarded ads retain their on-screen designation as advertisements.

VI. SUPPORT FOR CREDIBLE LOCAL NEWS OUTLETS

Reporting by credible local news organizations helps to inform the public on a wide range of issues, strengthening our democratic society. And [studies](#) indicate that the public perceives local news as more trustworthy and unbiased. Local reporting appears regularly on social media sites and serves as an important counterweight to disinformation. But the social media platforms have siphoned off a huge share of the advertising revenue that historically has funded news gathering. Largely as a result of these circumstances, local news organizations are contracting and dying off, creating a vacuum in which disinformation thrives. Over the [past two years alone](#), some 300 U.S. local news organizations have closed and 6,000 journalists have lost their jobs.

We recommend several approaches to supporting local news gathering:

Creation of a fund from fines and other financial policies: Section IV discussed the option of imposing fines and other financial policies on social media companies. The revenue from these policies could be applied to the creation of a new independent fund to support local news gathering. The fund would distribute grants to both nonprofit and for-profit organizations, allowing the recipients to retain journalists, sustain their operations, and maintain editorial independence. It would be critical to create an independent structure to administer this fund, insulated from political partisanship and interference. The fund should support both existing local outlets and startups. It also should help local news organizations develop sustainable models, both commercial and nonprofit, that result in more local reporters serving communities across the country.

Negotiating power for news outlets: The Administration and Congress should facilitate the ability of news outlets to negotiate with social media companies over the use of news content. One model would create a mechanism for multiple news outlets to bargain collectively with social media companies over payment amounts and arrangements. France has established this type of system, leaving control over pricing in the hands of social media companies and the news organizations. Several large French media companies recently reached an agreement with Google that will provide this type of compensation. The proposed [Journalism Competition and Preservation Act](#), sponsored by Representative David Cicilline (D., R.I.) and Senator Amy Klobuchar (D., Minn.), would give U.S. media companies greater leeway under anti-trust law to bargain collectively with social media companies. Another model exists in Australia, where a new law requires payments by social media companies to news organizations and imposes mandatory arbitration by the government if a funding agreement cannot be reached. If this type of model is pursued, the Administration would need to ensure that it does not result in the platforms taking down news content in order to avoid paying compensation.

Tax credits for subscribers and local advertisers: Recognizing the need to support local news gathering, the Administration should support an enhanced version of the [Local Journalism Sustainability Act](#), which would provide individual tax credits to people who subscribe to one or more local newspapers or online news sources for their personal use. Though not tied directly to reform of social media, this initiative would encourage taxpayers to provide financial support to for-profit and nonprofit local news sources. As currently drafted, the bill provides an initial tax credit of \$250, which decreases in subsequent years. We recommend a larger, sustained refundable tax credit, which would create a more powerful incentive to buy subscriptions. The tax-credit approach decentralizes decision making about which news outlets receive financial support, empowering subscribers themselves to make that choice. The tax credit should be refundable and available to those who don't itemize their taxes so that everyone can benefit.

As a number of organizations have [proposed](#), the federal government also should create a tax credit of \$2,500 to \$5,000 for small businesses that buy local advertising from commercial or nonprofit news organizations. This would help local news gathering and aid local businesses that have suffered financially during the coronavirus pandemic.

CONCLUSION

The media ecosystem has changed. Advances in computing efficiency, data storage capacity and connectivity have enabled the rise of new internet business models that have brought societal benefits but also significant harms. In the absence of adequate self-regulation, the social media industry now requires sustained government oversight. This paper offers targeted policy recommendations for the Administration's consideration as it seeks to reduce disinformation, hate speech, incitement of violence, and other forms of harmful online content that threaten democratic institutions. Policymakers can pursue this important goal while fully respecting the First Amendment and seeking to promote innovation that drives economic growth. Company vows to improve are no longer sufficient. It is time for smart, precise government intervention.