# PUBLIC SAFETY AND DISINFORMATION

Karen Kornbluh[*] & Eli Weiner[**]

CITE AS: 4 GEO. L. TECH. REV. 609 (2020)

TABLE OF CONTENTS

I. INTRODUCTION ..... 610
II. THE PROBLEM ..... 611
A. Misleading Outlets, Hyper-Partisan Clickbait, and Arbitraging The Trust Built Up by Independent Journalism. ..... 611
B. Personalized Political Propaganda Obscures The True Sponsors of Online Ads. ..... 612
C. Paid Influencers and Networks of Amplifiers Have Become Critical in 2020. ..... 614
D. Secret Groups, Encrypted Messaging, and Fringe Sites Linking to the Main Platforms. ..... 616
E. Loopholes Created by Different Platform Rules Are Applied Inconsistently and Without Transparency, Frustrating Accountability. 617
III. NEW DIGITAL DEMOCRACY ARCHITECTURE ..... 620
A. Update Offline Protections for an Online World ..... 621
B. Provide Choice by Funding a PBS of the Internet Through a "Superfund" Type Fee on Digital Advertising Revenue ..... 622
C. Create Accountability With a Code of Conduct Enforceable Through a Combination of the Following: ..... 622
IV. CONCLUSION ..... 623

[*] Senior Fellow and Director, Digital Innovation and Democracy Initiative, German Marshall Fund of the United States. This article is a precis of Karen Kornbluh and Ellen P. Goodman, *Safeguarding Digital Democracy Digital Innovation and Demoracy Roadmap*, GERMAN MARSHALL FUND OF THE UNITED STATES (Mar. 24 2020), http://www.gmfus.org/publications/safeguarding-democracy-against-disinformation [https://perma.cc/B4DN]. I am deeply thankful to the Democracy Fund, the Hewlett Foundation, and the Knight Foundation for supporting the Digital Innovation and Democracy Initiative at the German Marshall Fund of the U. S.; Rachel Tausenfreund who edited the original report; research assistant Ryan Whittington; and the *Georgetown Law Technology Review* for their helpful edits. Any mistakes herein are my own.

[**] Eli Weiner is a research assistant with the Digital Innovation and Democracy Initiative at the German Marshall Fund of the United States.

## I. INTRODUCTION

Too often, combatting public safety and disinformation on the Internet is presented as a false choice between continuing to allow platforms free reign to set rules of the road for our digital media ecosystem and giving the government more control over the content flowing across the networks. Currently, the government in India is moving to clamp down on Internet companies with sweeping new regulations that would force them to take down any content deemed "unlawful in any manner whatever."[1] Furthermore, in the United States Senate, legislative proposals would have government agencies condition immunity on whether Internet companies exhibit political bias or approve platform best practices.[2]

Other options would empower users instead of allowing either platforms or government to act as a censor. With an understanding of the digital information platforms themselves, new media gatekeepers could suggest options that would update the obligations of our old gatekeepers—to minimize the opportunity for user manipulation, boost public interest journalism, and promote democratic debate. A new media architecture would steer clear of vague rules and instead focus on updating offline protections, fostering choice, and public accountability. It would be technology-neutral and tailored with input from stakeholders. Moreover, in so doing, it would close the loopholes that allow bad actors to engage in online information warfare on the largest platforms without restricting free expression or stymieing welcome innovation.

At the German Marshall Fund of the United States, we have tackled these issues in a more detailed report.[3] The new architecture would have three major elements: (1) increasing offline rights and protections for consumers, elections, civil rights, and privacy; (2) promoting and sustaining local public news and media literacy; and (3) creating an accountability structure for these elements and content moderation.

---

[1] *See* INDIAN MINISTRY OF ELECTRONICS AND INFO. TECHNOLOGY, THE INFORMATION TECHNOLOGY [INTERMEDIARIES GUIDELINES (AMENDMENT) RULES] 2018 2 (Dec. 24, 2018), https://meity.gov.in/writereaddata/files/Draft_Intermediary_Amendment_24122018.pdf [https://perma.cc/5JWJ-XAJM].

[2] *See generally* Ending Support for Internet Censorship Act, S. 1914, 116th Cong. § 1 (2019), https://www.congress.gov/bill/116th-congress/senate-bill/1914/text [https://perma.cc/LRX6-JPVR] (modifying Section 230 of the Communications Decency Act to limit the immunity for platforms under that Section).

[3] *See generally* Karen Kornbluh & Ellen P. Goodman, *Safeguarding Digital Democracy Digital Innovation and Democracy Initiative Roadmap*, GERMAN MARSHALL FUND U.S. (Mar. 24, 2020), http://www.gmfus.org/publications/safeguarding-democracy-against-disinformation [https://perma.cc/B4DN-8CUE].

## II. THE PROBLEM

"On the Internet, no one knows you're a dog," reads the caption from the famous New Yorker cartoon of a dog at a computer.[4] While personal anonymity has always been a feature of the Internet—until Facebook's policy of verifying accounts and Twitter's blue checks—a host of newer design features have developed which allow parasitic campaigns to manipulate users on the platforms, weaponizing tribal fears and corrupting the information ecosystem with disinformation. Partisan and clickbait outlets have the same design features as independent journalistic outlets, though they adhere to none of the standards of journalism (e.g., the masthead, corrections, separation of news from opinion). Online ads leverage users' data usually without their understanding to test and target content to the most susceptible; even so-called "organic content" (activity by other users) can be the hidden result of bots or coordinated activity. There are five major categories of online activity through which disinformation campaigns can deceive users: (A) misleading outlets; (B) personalized political propaganda; (C) paid influencers and networks of amplifiers; (D) secret groups, encrypted messaging, and fringe sites linking to main platforms; and (E) inconsistently-applied loopholes created by different platform rules.

### A. Misleading Outlets, Hyper-Partisan Clickbait, and Arbitraging The Trust Built Up by Independent Journalism

The user interfaces used by platforms ensures the appearance of their stories is the same as those from traditional journalistic organizations, while separating the stories from the outlet that produces it. This interface denies the reader access to information developed by independent journalism to offer readers transparency about the news sources (including bylines, mastheads, separated news from opinion, corrections, and codes and standards)—or this interface can even obscure whether the news outlet provides this information at all. Readers who see some stories with fact-checks may assume that stories without those have been checked and deemed factual, when in fact, only a portion are fact-checked, and stories bear no indication if they are satire or opinion. Not only do these practices boost trust in disinformation, they also, over time, undermine trust in all news.

---

[4] Michael Cavna, *'Nobody Knows You're a Dog': As Iconic Internet cartoon Turns 20, Creator Peter Steiner Knows the Joke Rings as Relevant as Ever*, WASH. POST (July 31, 2013), https://www.washingtonpost.com/blogs/comic-riffs/post/nobody-knows-youre-a-dog-as-iconic-internet-cartoon-turns-20-creator-peter-steiner-knows-the-joke-rings-as-relevant-as-ever/2013/07/31/73372600-f98d-11e2-8e84-c56731a202fb_blog.html [https://perma.cc/DMW4-5VDD].

Meanwhile, legacy newspapers struggle as platforms have continued to cannibalize the revenue of local independent journalism. Google and Facebook now capture 58 percent of the advertising market.[5] One of the nation's leading newspaper companies with 30 newspapers around the country, McClatchy, recently filed for bankruptcy.[6] Since 2004, more than 1,800 local print outlets have closed, and at least 200 counties have no newspaper at all.[7]

These statistics imply a civic emergency. Areas with limited local news have less politically aware populations.[8] One study found that the city of Denver experienced a decrease in civic engagement after the closure of *The Rocky Mountain News* and the shrinking of the *Denver Post*.[9] Layoffs and closings have hamstrung the ability to hold public and corporate officials accountable.[10]

### B. Personalized Political Propaganda Obscures The True Sponsors of Online Ads

Disinformation campaigns advertise to small audiences of users enticing them to share memes, take quizzes, donate, follow "news" sites and fictitious accounts, and to join groups. The ads are targeted to audiences based on data gathered about them and people like them. For example, in one effort, women over the age of twenty-five who had expressed interest in pregnancy were served a targeted ad featuring anti-vaccination conspiracies.[11]

---

[5] Todd Spangler, *Amazon on Track to Be No. 3 In U.S. Digital Ad Revenue but Still Way Behind Google, Facebook*, VARIETY (September 9, 2018), https://variety.com/2018/digital/news/amazon-us-digital-ad-revenue-google-facebook-1202947923/ [https://perma.cc/U8RC-UQP5].

[6] Katy Robertson & Marc Tracy, *McClatchy, a Major U.S. Newspaper Chain, Files for Bankruptcy*, N. Y. TIMES (Feb. 13, 2020), https://www.nytimes.com/2020/02/13/business/media/mcclatchy-bankruptcy.html [https://perma.cc/A2MJ-M65H].

[7] PEN AM., LOSING THE NEWS: THE DECIMATION OF LOCAL JOURNALISM AND THE SEARCH FOR SOLUTIONS 27 (2019), https://pen.org/wp-content/uploads/2019/12/Losing-the-News-The-Decimation-of-Local-Journalism-and-the-Search-for-Solutions-Report.pdf [https://perma.cc/TN2C-ZGMV].

[8] *Id.* at 14.

[9] *Id.* at 14.

[10] Julie Bosman, *How the Collapse of Local News Is Causing a 'National Crisis*,*'* N. Y. TIMES (Nov. 20, 2019), https://www.nytimes.com/2019/11/20/us/local-news-disappear-pen-america.html [https://perma.cc/BG5P-JXZ2].

[11] Meira Gebel, *Anti-Vaccination Ads on Facebook Are Targeting Pregnant Women, While a Measles Outbreak Spreads Across The Country*, BUS. INSIDER (Feb. 14, 2019), https://www.businessinsider.com/anti-vaccine-facebook-ads-target-pregnant-women-as-measles-spreads-2019-2 [https://perma.cc/BKV6-NGRS].

Although Facebook now "prohibits ads that include claims debunked by third-party fact checkers or, in certain circumstances, claims debunked by organizations with particular expertise," it decided to exempt ads from political candidates from fact-checking requirements on the grounds that it was important to allow the ads to be subject to public scrutiny.[12] However, as hundreds of Facebook employees warned in an open letter voicing their objection to the policy, "it's hard for people in the electorate to participate in the 'public scrutiny' that we're saying comes along with political speech," because these ads are only shown to small groups.[13] Google's ad policies prohibit misleading content, and the company announced that it will only restrict misinformation in political ads that "could significantly undermine participation or trust in an electoral or democratic process," suggesting that "misleading content" will be defined narrowly to exclude misinformation about specific candidates or policies.[14] The discrepancy among platform rules creates loopholes that cross-platform disinformation campaigns exploit.[15]

Even if ads do not contain falsehoods, the lack of a shared information space undermines public debate. Facebook employees warned that "[t]hese ads are often so micro-targeted that the conversations on our platforms are much more siloed than on other platforms."[16] Information regulators in the United Kingdom and Spain, as well as members of the U.S. Congress, have similarly urged that platforms pause in the distribution of campaign ads. Twitter CEO Jack Dorsey announced that the company will no longer sell political ads that reference elections, candidates, parties, legislation and regulations, elected or appointed officials, or judicial decisions.[17] Google has restricted micro-targeting in political ads.[18] A recent survey by the Knight

---

[12] *See Advertising Policies: Misinformation*, FACEBOOK, https://facebook.com/policies/ads/prohibited_content/misinformation [https://perma.cc/7XEP-WY58].

[13] *Read The Letter Facebook Employees Sent to Mark Zuckerberg About Political Ads*, N.Y. TIMES (Oct. 28, 2019), https://www.nytimes.com/2019/10/28/technology/facebook-mark-zuckerberg-letter.html [https://perma.cc/3XQJ-TRKH].

[14] Scott Spencer, *An Update on Our Political Ads Policy*, GOOGLE (Nov. 20, 2019), https://www.blog.google/technology/ads/update-our-political-ads-policy/ [https://perma.cc/E69L-9W6Y].

[15] *See* Emily Glazer & Patience Haggin, *Political Ads are Flourishing Online. Few Agree How to Regulate Them*, WALL ST. J. (Nov. 15, 2019), https://www.wsj.com/articles/as-political-ad-spending-balloons-online-consensus-on-regulation-is-elusive-11573813803 [https://perma.cc/Y2Q4-JANX].

[16] *Read The Letter Facebook Employees Sent to Mark Zuckerberg About Political Ads*, *supra* note 13.

[17] *Political Content Policy*, TWITTER (2020), https://business.twitter.com/en/help/ads-policies/prohibited-content-policies/political-content.html [https://perma.cc/W65W-9CCP].

[18] *See* Spencer, *supra* note 14.

Foundation and Gallup found that more than 70 percent of Americans oppose the use of personal data for microtargeting purposes by political campaigns.[19]

In addition to issues with fact-checking and micro-targeting, current real-time ad labelling and after-action public libraries provide varying and inadequate information to potential voters, depriving them of the ability to know who is sponsoring ads. In the absence of industry-wide standards, platform practices differ from each other in terms of what kinds of ads they deem political. In addition, the data in the after-action databases is not robust, scoping is incorrect, identification is insufficient, metrics are fuzzy, advertisement data is unverifiable, and targeting information is lacking.[20] Journalism professor Jonathan Albright found it easy to use a false identity when buying advertisements, and Mozilla researchers found bugs and technical issues in the Facebook ad library.[21] Google's database functions better, but it does not include ads about topics, only candidates.[22] Moreover, even when the information in the databases is updated and correct, it can still fail to reveal the parties funding the ads.

### C. Paid Influencers and Networks of Amplifiers Have Become Critical in 2020

These networks of so-called "organic" (not ad-based) activity can flood the information zone with disinformation, manipulating algorithmic recommendations to fill trending lists and search engines. Autocratic governments have long flooded the information ecosystem to distract from inconvenient news and to deceive the public about critical or independent

---

[19] Dannagal G. Young & Shannon C. McGregor, *Mass Propaganda Used to be Difficult, but Facebook Made it Easy*, WASH. POST (Feb. 14, 2020), https://www.washingtonpost.com/outlook/2020/02/14/mass-propaganda-used-be-difficult-facebook-made-it-easy/ [https://perma.cc/9RGF-AG5L].

[20] *See generally* Paddy Leerssen et al., *Platform Ad Archives: Promises and Pitfalls*, 8 INTERNET POL. REV. 1 (2019).

[21] *See* Jonathan Albright, *Facebook and the 2018 Midterms*, MEDIUM (Nov. 4, 2018), https://medium.com/s/the-micro-propaganda-machine/the-2018-facebook-midterms-part-i-recursive-ad-ccountability-ac090d276097 [https://perma.cc/8RUL-UUUA]; *Facebook's Ad Archive API is Inadequate*, MOZILLA BLOCK (Apr. 29, 2019), https://blog.mozilla.org/blog/2019/04/29/facebooks-ad-archive-api-is-inadequate/ [https://perma.cc/CXH6-U94F].

[22] *See* Matthew Rosenberg, *Ad Tool Facebook Built to Fight Disinformation Doesn't Work as Advertised*, N.Y. TIMES (July 25, 2019), https://www.nytimes.com/2019/07/25/technology/facebook-ad-library.html [https://perma.cc/MAN5-QE4G]; Taylor Hatmaker, *Google Releases a Searchable Database of U.S. Political Ads*, TECHCRUNCH (Aug. 15, 2018), https://techcrunch.com/2018/08/15/google-political-ad-library/ [https://perma.cc/B4CK-FN3K].

views. But this tactic formerly required considerable resources, and often ran into the roadblock of skeptical newspapers and broadcast news editors. Now, the disinformation campaign toolkit is available off-the-shelf from commercial vendors, widespread and affordable for enterprising political outfits or nation-states looking to maximize the budget lines earmarked for information warfare.

The NATO Strategic Communications Center of Excellence confirmed that it remains shockingly easy—and shockingly cheap—to purchase comments, likes, views, and followers from third parties operating on the major platforms.[23] Private "black PR" firms increasingly offer their services to run online influence operations by using paid trolls operating fake accounts. Nathaniel Gleicher, Facebook's head of cybersecurity policy, labeling "the professionalization of deception" as a growing threat.[24]

Search engines are also manipulated by these actors using a variety of tactics identified by researchers, including "keyword stuffing," or adding popular keywords to unrelated websites to promote content in search-engine rankings; "link bombs," or increasing the number of other sites that link to the page; "mutual admiration societies," or groups of websites with links designed to appear as legitimate citations that instead point to each other;[25] and "data voids" that create news around an unused search term (e.g., "crisis actor" or "caravan") and then post content with disinformation that is found by users searching for the new term.[26]

Bots, trolls, and networks of true believers can work in coordinated fashion to increase the number of times an individual sees disinformation from different sources, crafting a sealed information environment. With such an enormous capacity to magnify a given message, repetition becomes reality. Indeed, "[t]he volume and recency of disinformation matter," according to a Hewlett Foundation review, and "people are more likely to be affected by

---

[23] Davey Alba, *Fake 'Likes' Remain Just a Few Dollars Away, Researchers Say*, N.Y. TIMES (Dec. 6, 2019), https://www.nytimes.com/2019/12/06/technology/fake-social-media-manipulation.html [https://perma.cc/L5J6-4882].

[24] Craig Silverman, Jane Lytvynenko & William Kung, *Disinformation For Hire: How A New Breed Of PR Firms Is Selling Lies Online,* BUZZFEED NEWS (Jan. 6, 2020), https://www.buzzfeednews.com/article/craigsilverman/disinformation-for-hire-black-pr-firms [https://perma.cc/7JKH-DV85].

[25] JOSHUA A. TUCKER ET AL., HEWLETT FOUND., SOCIAL MEDIA, POLITICAL POLARIZATION, AND POLITICAL DISINFORMATION: A SCIENTIFIC STUDY 30 (2018).

[26] *See generally* MICHAEL GOLEBIEWSKI & DANAH BOYD, DATA & SOC'Y, DATA VOIDS: WHERE MISSING DATA CAN EASILY BE EXPLOITED (2018), https://datasociety.net/wp-content/uploads/2018/05/Data_Society_Data_Voids_Final_3.pdf [https://perma.cc/N3JU-TYNN].

inaccurate information if they see more and more recent messages reporting facts, irrespective of whether they are true."[27]

Amplifiers and their networks cause algorithms to sense engagement and further amplify the content they push, to the point where it emerges as a newsworthy or trending topic. These algorithms prioritize content for newsfeeds and recommendations. Search results are optimized for user "engagement" (measured by the number of comments, shares, likes, etc.) to attract and keep users' attention so that they will stay online to be served more ads. These networks work across platforms. According to the Senate Intelligence Committee, "achieving the 'viral' spread of YouTube videos generally entails capitalizing on the reach and magnitude of Facebook and Twitter networks to spread links to the video hosted on YouTube."[28]

### D. Secret Groups, Encrypted Messaging, and Fringe Sites Linking to the Main Platforms

In addition to promoting misleading news, micro-targeting users with personalized persuasion, and flooding the news zone, disinformation campaigns manipulate users by creating and infiltrating accounts, pages, and groups, pretending to represent collections of Americans with a common interest. For example, the Internet Research Agency (IRA) in Russia created a fake "Blacktivist" page that garnered 11.2 million engagements over the course of the IRA's campaign. In general, during the 2016 elections, more than 62,000 users committed to attend 129 events organized by Russian trolls,[29] including through Russian-created Facebook pages such as Heart of Texas and United Muslims of America, which had over 300,000 followers.[30]

But as disinformation moves to Facebook's groups (the private version of pages) and encrypted messaging—which receives limited moderation and is not accessible to the public—even more users are susceptible to what researcher Jonathan Albright calls "shadow organizing" (when bad actors seed disinformation) without detection.[31] Shadow organizing can happen across

---

[27] TUCKER ET AL., *supra* note 25, at 40.
[28] S. REP. NO 116-XX, pt. 2, 58 (2019), https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf [https://perma.cc/4HLU-H59F].
[29] Dustin Volz & David Ingram, *Facebook: Russian Agents Created 129 U.S. Election Events*, REUTERS (Jan. 25, 2018), https://www.reuters.com/article/us-usa-trump-russia-facebook/facebook-russian-agents-created-129-u-s-election-events-idUSKBN1FE37M [https://perma.cc/S5JS-P8LG].
[30] S. REP. NO. 116-XX, pt. 2, *supra* note 28 at 47.
[31] *See* Jonathan Albright, *The Shadow Organization of Facebook Groups*, MEDIUM (Nov. 4, 2018), https://medium.com/s/the-micro-propaganda-machine/the-2018-facebook-midterms-part-ii-shadow-organization-c97de1c54c65 [https://perma.cc/29UE-KRV2].

multiple platforms—starting in fringe sites with more lenient rules such as Gab or 4Chan, and spreading to private Facebook groups and then beyond. Nina Jankowicz warns that private groups, along with the fringe sites that link to the mainstream platforms, are "where unsavory narratives ferment and are spread, often with directions about how to achieve maximum impact."[32]

The danger of these secluded, online propagandizing and recruitment grounds are growing increasingly apparent. For example, homegrown militia movements that traffic in conspiracy theories and refuse to recognize the authority of the federal government are organizing among members of police departments through private Facebook groups. Facebook groups for militia organizations like the Three Percenters and the Oath Keepers (which believe that the federal government plans to take away Americans' guns, install martial law and set up concentration camps to kill dissenters), along with Neo-Confederate, Islamophobic, and white supremacist groups, count hundreds of active and former police officers among their ranks.[33]

### E. Loopholes Created by Different Platform Rules Are Applied Inconsistently and Without Transparency, Frustrating Accountability

When platforms say they do not want to police speech, they are disregarding a core part of their business. The Lawyers' Committee for Civil Rights Under Law wrote Mark Zuckerberg that "Facebook constantly regulates speech on its platform with curation algorithms that decide which content gets amplified and which gets buried. You have decided it is acceptable to regulate speech to increase user engagement."[34]

Disinformation disproportionately weaponizes animosity against immigrants, Muslims, Jews, women, and African Americans. Around the world, coordinated online hate speech against racial and ethnic minorities has led to violence. Rumors about Muslims circulating on WhatsApp have resulted in hangings in India.[35] In March 2018, the chairman of the U.N.

---

[32] Joe Uchill, *Privacy Plan Could Worsen Facebook's Echo Chamber Problem*, AXIOS (Mar. 7, 2019), https://www.axios.com/facebook-privacy-plan-echo-chamber-misinformation-b87173a4-5aab-4e6c-a1b6-23aeabafeed5.html [https://perma.cc/9TCZ-8MR3].

[33] *See* Will Carless & Michael Corey, *The American Militia Movement, a Breeding Ground for Hate, is Pulling in Cops on Facebook*, REVEAL NEWS (June 24, 2019), https://www.revealnews.org/article/the-american-militia-movement-a-breeding-ground-for-hate-is-pulling-in-cops-on-facebook/ [https://perma.cc/ZVE7-LKXV] .

[34] Letter from Kristen Clarke, Pres. & Exec. Dir., Law. Committee for C.R. Under Law, to Mark Zuckerberg, Chief Exec. Officer, Facebook (Nov. 4, 2019), https://lawyerscommittee.org/letter-to-facebook-regarding-failure-to-address-misinformation-online/ [https://perma.cc/N4WH-XMGL].

[35] Timothy Mclaughlin, *How WhatsApp Fuels Fake News and Violence in India*, WIRED (Dec.

Independent International Fact-Finding Mission on Myanmar said social media companies had played a "determining role" in the violence in the country, having "substantively contributed to the level of acrimony and dissension and conflict."[36] These comments were echoed a year later by the U.N. Special Rapporteur, who warned that "[p]ublic institutions linked to [Myanmar's] military, its supporters, extremist religious groups and members of the government continue to proliferate hate speech and misinformation on Facebook."[37]

According to the Philanthropy for Active Civic Engagement, when bad actors use various harassment techniques to "distort or drown out disfavored speech,"[38] they disproportionately target, "journalists, women, and ethnic or racial minorities."[39] Since August 2019, at least three mass shooters announced their plans on a fringe website and then spread their ideology on the larger platforms.[40]

Platforms have adopted new rules and hired tens of thousands of staff and contractors to limit hateful content, but the application and enforcement of these rules appear to be inconsistent. Leading U.S. civil rights and human rights organizations have accused Facebook of "reckless disregard for civil rights."[41] The Anti-Defamation League also points out inconsistency and lack

---

12, 2018), https://www.wired.com/story/how-whatsapp-fuels-fake-news-and-violence-in-india/ [https://perma.cc/95EG-WH6Y]; Vindu Goel et al., *How WhatsApp Leads Mobs to Murder in India*, N.Y. TIMES (July 18, 2018), https://www.nytimes.com/interactive/2018/11/23/technology/whatsapp-india-killings-ES.html [https://perma.cc/H5AS-82N2].

[36] Mehdi Hasan, *Dear Mark Zuckerberg: Facebook is an Engine of Anti-Muslim Hate the World Over. Don't You Care?*, INTERCEPT (Dec. 7, 2019), https://theintercept.com/2019/12/07/facebook-mark-zuckerberg-muslims-islamophobia/ [https://perma.cc/4NP7-VSCZ].

[37] Tom Miles, *U.N. Urges Social Media, Investors to Promote Human Rights in Myanmar*, REUTERS (Mar. 5, 2019), https://www.reuters.com/article/us-myanmar-rights-un/u-n-urges-social-media-investors-to-promote-human-rights-in-myanmar-idUSKCN1QM1MP [https://perma.cc/CS49-29PJ].

[38] Tim Wu, *Is the First Amendment Obsolete?*, KNIGHT FIRST AMEND. INST. (2017), https://knightcolumbia.org/content/tim-wu-first-amendment-obsolete#/_ftn2 [https://perma.cc/DQQ7-89GX].

[39] Kelly Born, *Social Media: Driving or Diminishing Civic Engagement?*, MEDIUM (June 21, 2018), https://medium.com/infogagement/https-medium-com-infogagement-social-media-driving-or-diminishing-civic-engagement-9850954910ed [https://perma.cc/9CH7-NR3X].

[40] April Glaser, *8chan Is a Normal Part of Mass Shootings Now*, SLATE (Aug. 4, 2019), https://slate.com/technology/2019/08/el-paso-8chan-4chan-mass-shootings-manifesto.html [https://perma.cc/9D75-5EE8].

[41] Letter from the Leadership Conf. on Civ. & Hum. Rts. to Mark Zuckerberg, Chief Exec. Officer, Facebook (Oct. 21, 2019), http://civilrightsdocs.info/pdf/policy/letters/2019/Zuckerberg_ltr_10-21-19_final.pdf [https://perma.cc/89R6-TNGS].

of transparency in enforcement as major problems.[42] Freedom House warns that social media have "provided an extremely useful and inexpensive platform for malign influence operations by foreign and domestic actors alike."[43] Indeed, the anti-immigration page VDare and the white supremacy newsletter American Free Press are still available.[44] Richard Spencer remains on Twitter.[45] Alt-right influencers and content are still widely available on YouTube, including white nationalist activist Martin Sellner, who had documented contact with the perpetrator of the mass shooting in New Zealand in March 2019.[46] The enforcement mechanisms and moderation schemes are clearly working at a deficit, if they work at all.

The major platforms also have rules against what Facebook calls "coordinated inauthentic behavior."[47] However, they appear to enforce these rules more consistently against foreign state operations than domestic individuals or groups. When, for example, BuzzFeed and independent researchers identified two networks of pro-Trump Facebook pages that disseminated false or misleading information in a coordinated manner, the company responded that such networks did "not violate its policy against coordinated inauthentic behavior."[48] And the partisan and unreliable Daily

---

[42] *Confronting the Rise in Anti-Semitic Domestic Terrorism*: *Hearing Before the Subcomm. on Intelligence & Counterterrorism of the H. Comm. on Homeland Sec.*, 116th Cong. 13, 16 (2020) (statement of Jonathan Greenblatt, CEO and National Director, Anti-Defamation League), https://homeland.house.gov/imo/media/doc/Testimony-Granblatt.pdf [https://perma.cc/A2FT-KCHU].

[43] ADRIAN SHAHBAZ & ALLIE FUNK, FREEDOM HOUSE, THE CRISIS OF SOCIAL MEDIA 1 (2019), https://freedomhouse.org/sites/default/files/2019-11/11042019_Report_FH_FOTN_2019_final_Public_Download.pdf [https://perma.cc/Z8LV-VPJ2].

[44] *See* Julia Carrie Wong, *White Nationalists are Openly Operating on Facebook. The Company Won't Ac*t, GUARDIAN (Nov. 21, 2019), https://www.theguardian.com/technology/2019/nov/21/facebook-white-nationalists-ban-vdare-red-ice [https://perma.cc/MBS5-7X7G].

[45] *See* @RichardBSpencer, TWITTER https://twitter.com/RichardBSpencer?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor [https://perma.cc/C935-ZCPY] .

[46] Mark Di Stefano, *YouTube Reinstated A Prominent European White Nationalist After He Appealed His Removal,* BUZZFEED NEWS, (Aug. 29, 2019), https://www.buzzfeed.com/markdistefano/youtube-martin-sellner-ban-reinstated [https://perma.cc/FND8-CMUP].

[47] Alexandra S. Levine, Nancy Scola, Steven Overly, & Cristiano Lima, *Why the Fight Against Disinformation, Sham Accounts and Trolls Won't be Any Easier in 2020*, POLITICO (Dec. 1, 2019), https://www.politico.com/news/2019/12/01/fight-against-disinformation-2020-election-074422 [https://perma.cc/K9ST-N88G].

[48] Craig Silverman & Jane Lytvynenko, *Facebook Says Anonymous Pages Posting Coordinated Pro-Trump Content Do Not Break Its Rules*, BUZZFEED NEWS (Nov. 20, 2019), https://www.buzzfeednews.com/article/craigsilverman/coordinated-pro-trump-facebook-pages [https://perma.cc/7PNV-LWDH].

Wire, which garners more engagement for its content than any other significant publisher on Facebook, was found to utilize a coordinated promotion operation, yet it remains online.[49]

Fundamentally, the problem remains one of transparency; it is difficult to hold platforms accountable for the application of their rules, since their enforcement actions are not saved or made auditable, nor is platform traffic. For example, determining details about the Russian influence campaign in the 2016 election required the concerted efforts of both the Senate Intelligence Committee[50] and Special Counsel Robert Mueller.[51] In the case of airline crashes, government officials on the National Transportation Safety Board are able to collect the "black box" flight data recorder to find out what happened and help recommend updated safety regulations to the Federal Aviation Agency,[52] but such post-fact analysis cannot be conducted on the platforms.

## III. NEW DIGITAL DEMOCRACY ARCHITECTURE

Understanding how central manipulation is to the spread of disinformation allows us to craft solutions that focus on updating tried and true protections that empower users in the offline world rather than relying on top-down government control of speech or the passive hope that Silicon Valley will self-innovate its way to a satisfactory resolution.

A new architecture would ensure that companies' policies are consistent and enforced in a manner that is clear and responsive to the public. Additionally, imposing similar obligations on similar companies would protect them from accusations of taking political sides when they take action. But this new architecture should be flexible and *content- and technology-neutral* without sacrificing regulatory protection or realistic enforcement options.

The new architecture would update offline laws that safeguard consumers and elections, as well as civil rights and privacy, for the online information ecosystem. It would create a fund for independent journalism,

---

[49] Judd Legum, *Keeping it "Real"*, POPULAR INFO. (Oct. 30, 2019), https://popular.info/p/keeping-it-real [https://perma.cc/2LJE-KNYL].

[50] *See* David McCabe, *G.O.P.-Led Senate Panel Affirms Russia Attacked Election, and Urges Action*, N.Y. TIMES (Oct. 8, 2019), https://www.nytimes.com/2019/10/08/business/senate-report-russia-election.htm [https://perma.cc/H5G8-9YGZ].

[51] *See* Eric Geller, *Collusion Aside, Mueller Found Abundant Evidence of Russian Election Plot*, POLITICO (Apr. 18, 2019), https://www.politico.com/story/2019/04/18/mueller-report-russian-election-plot-1365568 [https://perma.cc/YX9X-P78V].

[52] *See* FAA Procedures for Handling National Transportation Safety Board Recommendations, FAA Order No. 1220.2G (May 13, 2011), https://www.faa.gov/documentLibrary/media/Order/1220.2G.pdf [https://perma.cc/TMR9-KNBP].

creating the equivalent of the Public Broadcasting Service (PBS) for the Internet. It would also strengthen the old "self-regulatory" approach to Internet regulation with an industry-civil society code of conduct—focused on practices, not content—backed up by monitoring and enabled by data sharing, with a regulatory and civil enforcement backstop.

### A. Update Offline Protections for an Online World

- *Use "light patterns" to empower users.* Designing user interfaces can help inform and empower users with better labeling of news, ads, altered video and audio, accounts, coordination, and even algorithmic recommendations. Platforms should provide users with the ability to customize algorithmic recommendations and track content complaints easily.
- *Restore the campaign finance bargain.* The Honest Ads Act (that would impose broadcast disclosure rules on platform ads)[53] has bipartisan support and should set the floor for disclosure. In addition, platforms should verify who is actually funding the ads (rather than listing front groups) and platform fact-checking policies should be consistent and applied to politicians. Platforms should also, as recommended by a Federal Election Commission commissioner,[54] limit micro-targeting of political ads.
- *Update civil and human rights law*. Discrimination, harassment, and privacy laws—such as public accommodation laws—should be updated for the digital age. Platforms should create and enforce rules that are consistent, transparent, and appealable for content removal and algorithmic prioritization.
- *Strengthen privacy rights*. The U.S. needs a uniform privacy law to provide users with the ability to protect their privacy and ensure that platforms are not allowing their data to be used to manipulate them. California has enacted a new privacy law[55] inspired by the European Union's General Data Protection Regulation.[56] Meanwhile, federal privacy legislation is gaining momentum that might go beyond the "notice and consent" framework to take certain practices off the table (such as collection and sale of biometric, location, or health

[53] Honest Ads Act, S. 1356, 116th Cong. (2019), https://www.congress.gov/bill/116th-congress/senate-bill/1356/text [https://perma.cc/D5E5-ZBE6].
[54] Ellen L. Weintraub, *Don't Abolish Political Ads on Social Media. Stop Microtargeting.*, WASH. POST (Nov. 1, 2019), https://www.washingtonpost.com/opinions/2019/11/01/dont-abolish-political-ads-social-media-stop-microtargeting/ [https://perma.cc/M6ZU-FP97].
[55] CAL. CIV. CODE §§ 1798.100–1798.199 (2020).
[56] 2016 O.J. (L 119) 679.

information; information collected from microphones or cameras; or cross-device tracking), and to create new governance procedures for companies collecting personal information.[57]

- *Stipulate national security information sharing*. Platforms should share information with each other and with government agencies on violent extremism as well as with the public on foreign election interference.

B. Provide Choice by Funding a PBS of the Internet Through a "Superfund" Type Fee on Digital Advertising Revenue

- *Platforms have syphoned away the ad revenue that once supported public interest local journalism*. A fund to support noncommercial public interest journalism, fact checkers, and media literacy could be created by taxing platform ad revenue—raising the cost of a business relying on data collection and the viral spread of disinformation while also supporting more "signal" in the system. A commitment from platforms to highlight and boost this content would also loosen algorithmic control over information flows. Just as support for public media in the past extended to communications infrastructure, so in the digital space, there should be noncommercial infrastructure as an alternative to provisioning by dominant "Big tech" companies.

C. Create Accountability With a Code of Conduct Enforceable Through a Combination of the Following:

- *Increase competition*. Lack of competition can undermine the health of the public square by limiting or skewing speech options. Policymakers understood this when they subjected broadcasters to ownership limits and prohibited them from cross-owning stations and print newspapers. Antitrust suits will move slowly and be tough to win under current law, whereas regulatory oversight can introduce more competition. Data portability would provide tools for users to export their network to competing platforms with the appropriate privacy safeguards in place. Interoperability would facilitate competition by enabling communication across networks. Some have suggested

[57] Cameron F. Kerry, *Breaking Down Proposals for Privacy Legislation: How Do They Regulate?*, BROOKINGS INST. (Mar. 8, 2018), https://www.brookings.edu/research/breaking-down-proposals-for-privacy-legislation-how-do-they-regulate/ [https://perma.cc/P7YB-JX5Q].

implementation by requiring platforms to maintain APIs for third-party access under terms that are fair, reasonable, and non-discriminatory.[58]

- *Open the Data Black Box by mandating platforms to share data*. Platforms should provide (1) intellectual property- and privacy-protected after-action disclosure of how content is algorithmically curated; (2) what targeting policies are used; (3) moderation decision logs; and (4) access to traffic data so civil society watchdogs, researchers, and governments may help assess information flows.
- *Refocus code of conduct*. Platforms and civil society should develop a technology-neutral code of conduct focused on practices, not content.
- *Establish oversight and enforcement*. Independent data sharing would make it possible to verify that platforms are complying with the code. Monitoring can be done by independent third parties, a new Digital Democracy Board, or an existing agency with oversight authority. If necessary, Section 230 immunity[59] could be conditioned on complying with the code.

## IV. CONCLUSION

Today, citizens themselves have few tools to evaluate a product's security, privacy, transparency, or algorithms. The United States has abdicated its traditional leadership role on Internet policy while Europe is stepping into the void, and the Russian and Chinese governments are leveraging the lack of international consensus to use the Internet for political repression and control, both in their own countries and abroad. Meanwhile, smaller countries, left with few options, are forced to operate in a geopolitical arena with little international agreement or guidance. It is time to take active steps to ensure that the Internet is a tool to strengthen, not undermine, democratic values. In order to do so, we must agree on a common framework for understanding these challenges and embrace practical solutions that protect privacy and free expression while strengthening the information ecosystem.

[58] OFFICE OF SENATOR MARK WARNER, POTENTIAL POLICY PROPOSALS FOR REGULATION OF SOCIAL MEDIA AND TECHNOLOGY FIRMS (DRAFT) 22 (2018), https://www.warner.senate.gov/public/_cache/files/d/3/d32c2f17-cc76-4e11-8aa9-897eb3c90d16/65A7C5D983F899DAAE5AA21F57BAD944.social-media-regulation-proposals.pdf [https://perma.cc/8G59-F8AR].

[59] Communications Decency Act of 1996, 47 U.S.C. § 230(c)(2) (2018).