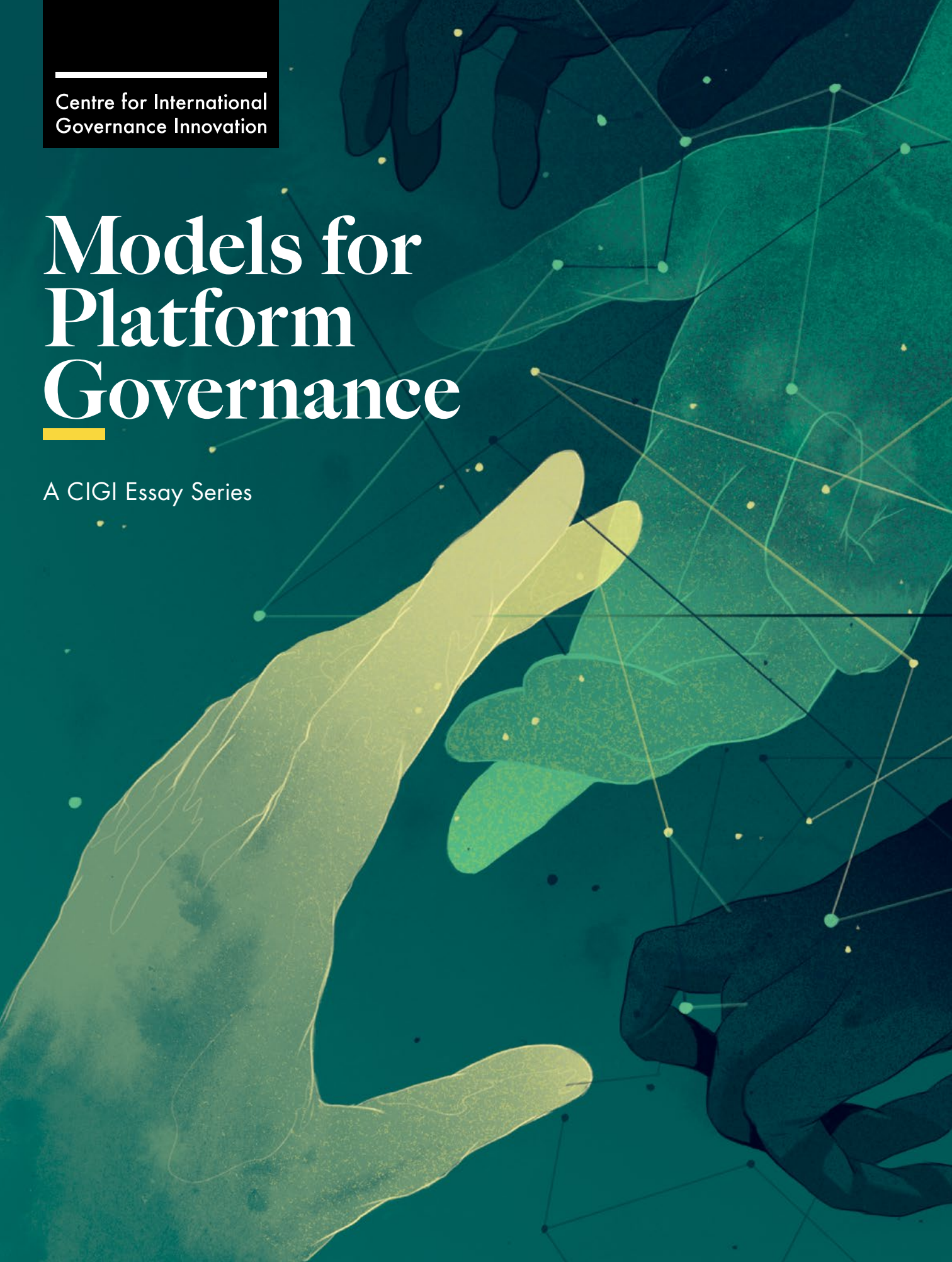
Centre for International Governance Innovation

Models for Platform Governance

A CIGI Essay Series



Contents

Credits

Director, Global Economy
ROBERT FAY

Senior Fellow
TAYLOR OWEN

Publisher
CAROL BONNETT

Managing Editor
ALLISON LEONARD

Publications Editor
LYNN SCHELLENBERG

Graphic Designer
SAMI CHOUHDARY

Illustrations by SIMÓN PRADES

Watch the series video at
cigionline.org/platforms

Centre for International Governance Innovation

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
www.cigionline.org

Taylor Owen

Introduction: Why Platform Governance?

Over the past three years, the debate about the role of digital technology in our society, our economy and our democracies has gone through a remarkable transformation. Following two decades of techno-optimism, whereby digital technology generally, and social media specifically, was viewed as broadly aligned with democratic good and so left to be governed in a laissez-faire environment, we are now in the midst of what could be called a "techlash."

The recent catalyst for this turn was the 2016 US presidential election — a moment that saw the election of Donald J. Trump with the aid of a Russian government adept at leveraging the digital infrastructure to manipulate the American electorate. But the root cause for this turn runs far deeper. It is the result of a notably laissez-faire policy approach that has allowed our public sphere to be privatized, embedding in our digital ecosystem the incentive structures of markets while allowing the social, economic and democratic costs of an unfettered system to be externalized, and therefore borne by the public.

Ultimately, the platform Web is made up of privately owned public spaces, largely governed by the commercial incentives of private actors, rather than the collective good of the broader society. Platforms are more like shopping malls than town squares — public to an extent, but ultimately managed according to private interests. Once nimble start-ups, Google, Facebook, Twitter and Amazon now serve billions of users around the world and increasingly perform core functions in our society. For many users, particularly those in emerging economies, these companies *are* the primary filtering point for information on the internet (Cuthbertson 2017). The private gains are clear to see — Google, Facebook and Amazon are among the most profitable companies in history. But in spite of myriad benefits offered by platforms, the costs are clear as well.

The social costs of the platform economy are manifesting themselves in the increasingly toxic nature of the digital public sphere, the amplification of misinformation and disinformation, the declining reliability of information, heightened polarization and the broad mental health repercussions of technologies designed around addictive models.

The economic costs are grounded in the market distortion created by increased monopolistic behaviour. The vast scale of the digital platform economy not only affords near-unassailable competitive advantages, but also invites abuses of monopoly power in ways that raise barriers to market entry (Wu 2018). Moreover, the ubiquity of the platform companies in the consumer marketplace creates special vulnerabilities because of the amount of control they wield over data, advertising and the curation of information.

The policy response is riddled with challenges.

The costs to our democracy are grounded not only in the decline of reliable information needed for citizens to be informed actors in the democratic process and the undermining of public democratic institutions, but in threats to the integrity of the electoral system itself.

As we collectively learn more about the nature of these problems, in all their complexity and nuance, this moment will demand a coordinated and comprehensive response from governments, civil society and the private sector. Yet, while there is growing recognition of the problem, there remains significant ambiguity and uncertainty about the nature and scale of the appropriate response.

The policy response is riddled with challenges. Since the digital economy touches so many aspects of our lives and our economies, the issues that fall under this policy rubric are necessarily broad. In countries around the world, data privacy, competition policy, hate speech enforcement, digital literacy, media policy and governance of artificial intelligence (AI) all sit in this space. What's more, they are often governed by different precedents, regulated by siloed departmental responsibility, and lack coordinated policy capacity. This confusion has contributed to a policy inertia and increased the likelihood that governments fall back on self-regulatory options.

And so democratic governments around the world have begun to search for a new strategy to govern the digital public sphere. Looking for an overarching framework, many are converging on what might be called a platform governance agenda.

The value of a platform governance approach is that first, it provides a framework through which to connect a wide range of social, economic and democratic harms; second, it brings together siloed public policy areas and issues into a comprehensive governance agenda; and third, it provides a framework for countries to learn from and coordinate with each other in order to exert sufficient market pressure.

But what might a platform governance agenda look like? There are three dimensions to consider in answering this question — policy coordination, scale of appropriate response and degree of regulatory risk.

First, there are no single-issue solutions to the challenges of technology and society. In order to address the breadth of policy areas in this space, we need a combination of content, data and competition policies that are implemented in coordination across government and between governments. This will demand a coordinated "whole-of-government" effort to bring together a wide range of policies. The challenges we confront are systemic, built into the architecture of digital media markets, therefore public policy response must be holistic and avoid reactions that solve for one aspect of the problem while ignoring the rest.

Second, within the platform governance agenda there is a need for multiple scales of responses for different policy issues: national implantation; international coordination; and international collaboration. As this essay series suggests, there is an urgent need for global platform governance, as no single state can shift the structure of the platform economy alone. Platforms are global organizations, which, in the absence of enforced national

rules, will default to their own terms of service and business practices. At the same time, because of the scale of the operation of these companies and the power they have accrued as a result, as well as the complexity of the new governance challenges they present, it is very difficult for any individual country to go it alone on regulation. However, this need for global governance is complicated by a parallel need for subsidiarity in policy responses. On some issues, such as speech regulation, policy must be *nationally implemented*. In these cases, countries can learn from and iterate off each other's policy experimentation. On others, such as ad-targeting laws, *international coordination* is necessary, so that countries can exert collective market power and align incentives. On others, such as AI standards, *international collaboration* is needed to ensure uniform application and enforcement and to overcome collective action problems.

Third, the issues that fall under the platform governance agenda are of varying levels of complexity and regulatory risk. Some policies have a high degree of consensus and limited risk in implementation. The online ad micro-targeting market could be made radically more transparent, and in some cases could be suspended entirely. Data privacy regimes could be updated to provide far greater rights to individuals and greater oversight and regulatory power to punish abuses. Tax policy could be modernized to better reflect the consumption of digital goods and to crack down on tax-base erosion and profit shifting. Modernized competition policy could be used to restrict and roll back acquisitions and to separate platform ownership from application or product development. Civic media could be supported as a public good. And large-scale and long-term civic literacy and critical-thinking efforts could be funded at scale by national governments. That few of these have been implemented is a problem of political will, not policy or technical complexity. Other issues, however, such as content moderation, liability and AI governance, are far more complex and are going to need substantive policy innovation.

The categorization of these three variables in Table 1 is not intended to be definitive. Many of these issues overlap categories, and the list of policies is certainly not exhaustive. But it may serve as a typology for how this broad agenda can be conceptualized.[1]

Table 1: Variables Affecting Platform Governance

Theme	Policy	Scale	Regulatory Risk
Content	Content moderation	Nationally led	High
	Ad transparency	International coordination	Low
	Bot and agent identification	International coordination	Moderate
	Civic journalism	Nationally led	Low
	Misinformation-focused cyber security	International collaboration	Moderate
	Research	International coordination	Low
	Digital literacy	Nationally led	Low
	Liability	International coordination	High
Data	Algorithmic accountability	International collaboration	High
	Data rights	International coordination	High
Competition	Modernized antitrust	International coordination	Moderate
	Mergers and acquisitions restrictions	Nationally led	Moderate
	Data portability and interoperability	International collaboration	Moderate
	Fair taxation	International collaboration	Low

Source: Author.

Just as we needed (and developed) new rules for the post-war industrial economy, we now need a new set of rules for the digital economy. Instead of rules to govern financial markets, monetary policy, capital flow, economic development and conflict prevention, we now need rules to regulate data, competition and content — the intangible assets on which most of the developed economy, and increasingly the health of our societies, now depend. This is the global governance gap of our time.

As this model evolves, there will be a need for other countries to not only collaborate on implementation, but also coordinate responses and iterate policy ideas. This work will invariably occur through state organizations such as the Group of Seven, the Group of Twenty, the Organisation for Economic Co-operation and Development, and the United Nations. But the situation will also demand new institutions to bring together the state and non-state actors needed to solve these challenging policy problems. One promising place for this policy coordination is the International Grand Committee on Big Data, Privacy and Democracy (IGC), which has evolved to use platform governance as its overarching frame. As a self-selected group of parliamentarians concerned with issues of platform governance, the IGC has the opportunity to be a catalyzing international institution for the design and coordination of a platform governance agenda.

And that is why CIGI has convened this essay series, and why we will bring a network of global scholars to Dublin in November 2019 to support this nascent international institution. Whatever the IGC's role ahead within this emerging realm of governance, we hope this conversation sparks a much-needed global governance process.

NOTE

1 Many of the policies discussed below are also articulated in Greenspon and Owen (2018) and in Owen (2019).

WORKS CITED

Cuthbertson, Anthony. 2017. "Who controls the Internet? Facebook and Google dominance could cause the 'Death of the Web.'" *Newsweek*, November 2. www.newsweek.com/facebook-google-internet-traffic-net-neutrality-monopoly-699286.

Greenspon, Ed and Taylor Owen. 2018. *Democracy Divided: Countering Disinformation and Hate in the Digital Public Sphere*. Ottawa, ON: Public Policy Forum. https://ppforum.ca/wp-content/uploads/2018/08/DemocracyDivided-PPF-AUG2018-EN.pdf.

Owen, Taylor. 2019. "Six Observations on Securing the Integrity of the Digital Public Sphere." Discussion paper for the International Meeting on Diversity of Content in the Digital Age.

Wu, Tim. 2018. *The Curse of Bigness: Antitrust in the New Gilded Age*. New York, NY: Columbia Global Reports.

ABOUT THE AUTHOR

Taylor Owen is a CIGI senior fellow and the editor of *Models for Platform Governance*. He is an expert on the governance of emerging technologies, journalism and media studies, and on the international relations of digital technology. At CIGI, Taylor is working on issues of international digital governance, as well as launching a podcast on the digital economy. Taylor holds the Beaverbrook Chair in Media, Ethics and Communications and is an associate professor in the Max Bell School of Public Policy at McGill University. Previously, he was an assistant professor of digital media and global affairs at the University of British Columbia and the research director of the Tow Center for Digital Journalism at Columbia University. His doctorate is from the University of Oxford, and he has been a Pierre Elliott Trudeau Foundation Scholar, a Banting Postdoctoral Fellow, an Action Canada and Public Policy Forum Fellow, the 2016 Public Policy Forum Emerging Leader and, until 2019, sat on the board of directors at CIGI. Taylor currently serves on the governing council of the Social Sciences and Humanities Research Council.

Centre for International
Governance Innovation

DATA GOVERNANCE IN THE DIGITAL AGE

A CIGI ESSAY SERIES

Find out why Canada needs
a national data strategy.

cigionline.org/data-governance

Pierre François Docquir

The Social Media Council: Bringing Human Rights Standards to Content Moderation on Social Media

Increasingly, we turn to social media platforms to access the information and ideas that structure the agenda and content of public debates (Newman et al. 2019). Giant social media companies have elevated themselves to a position of market dominance where they hold a considerable degree of control over what their users see or hear on a daily basis. We may know that content moderation and distribution — in other words, the composition of users' feeds and the accessibility and visibility of content on social media — happen through a combination of human and algorithmic decision-making processes but, overall, current practices offer very little in terms of transparency and virtually no remedy to individual users when their content is taken down or demoted (ARTICLE 19 2018, 15).

This situation has become a major issue for democratic societies. The responsibilities of the largest social media companies are currently being debated in legislative, policy and academic circles around the globe, but many of the numerous initiatives that have been put forward do not sufficiently account for the protection of freedom of expression and other fundamental rights. There is a strong consensus among international experts on freedom of expression that the mere regulation of speech by contract (that is, a company controlling its own platform on the basis of terms of service and community standards) fails to provide adequate transparency and protection for freedom of expression and other human rights (ibid.). The creation of content moderation duties in legislation — exemplified by Germany's Network Enforcement Act — tends to lead to the creation of systems where private actors are tasked with applying criminal law and other national legal provisions under short deadlines and the threat of very heavy fines (ARTICLE 19 2017). These systems increase the fragmentation of legal obligations for social media companies, creating a situation where individual users have little or no remedy to address hasty content removal and providing no guarantee for the protection of individual freedoms.

This situation has become a major issue for democratic societies.

Media landscapes and the diversity of roles fulfilled by tech companies have been evolving at a high pace and will continue to do so. Democracy now requires that we engage in a collective learning process to organize online content moderation in a manner compatible with the requirements of international standards on freedom of expression. From this perspective, the need for a mechanism capable of ensuring an effective public supervision of content moderation on social media platforms is increasingly recognized on all sides.

ARTICLE 19, a leading free speech global organization, has proposed the creation of the "Social Media Council" (SMC) — a model for a multi-stakeholder accountability mechanism that would provide an open, transparent, independent and accountable forum to address content moderation issues on social media platforms on the basis of international standards on human rights. The SMC model puts forward a voluntary approach to the oversight of content moderation: participants (social media platforms and all stakeholders) sign up to a mechanism that does not create legal obligations. Its strength and efficiency rely on voluntary compliance by platforms, whose commitment, when signing up, will be to respect and execute the SMC's decisions (or recommendations) in good faith. This proposal was endorsed by UN Special Rapporteur David Kaye, who recommended in April 2018 that "all segments of the ICT sector that moderate content or act as gatekeepers should make the development of industry-wide accountability mechanisms (such as a social media council) a top priority" (UN General Assembly 2018, para. 72).

ARTICLE 19 initially envisioned the SMC as having an ambitious scope: a network of national or regional SMCs entrusted with providing general guidance to social media platforms and deciding individual complaints brought by individual users, operating on the basis of international standards on human rights and coordinating through the mediation of an international SMC. Such multi-stakeholder, transparent, accountable and independent fora could weave freedom of expression within all aspects of online content moderation and distribution across all social media platforms, from integrating international standards in decisions to delete or demote content, to ensuring exposure to the broadest possible diversity of information and ideas through a form of human-rights-optimized algorithmic distribution.

ARTICLE 19, together with the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression and Stanford University's Global Digital Policy Incubator, submitted this proposal to a working meeting of academics, civil society organizations and social media companies, which generated intense discussions, as the resulting conference report records (Global Digital Policy Incubator, ARTICLE 19 and Kaye 2019). This conference and other subsequent meetings have helped shed light on the questions, big

and small, raised by the project of creating an SMC — see, for instance, the comments from the Electronic Frontier Foundation (McSherry 2019). There are different visions of what the exact roles and functions of this new mechanism should be, where it should be set up or how it would interact with other initiatives, such as the creation of an oversight board by Facebook (ARTICLE 19 2019).

A first point of discussion is the choice of rules that should preside over the oversight of content moderation. While there is a growing consensus that international standards on human rights provide the appropriate universal legal framework, there may be different ways to apply this body of rules. The SMC could simply refer to these rules directly, and the authoritative interpretation by international and regional courts and special mechanisms would provide all necessary guidance to inform the SMC's decisions. Another option would be to adopt a code of human rights principles for content moderation. The specific adaptation of international standards to online content moderation through the adoption of a code would ensure that the SMC operates under stricter guidance than the broad reference to international standards. In both cases, as is generally the situation with the application of international standards, a certain "margin of appreciation" — or margin of flexibility — would be part of the SMC mechanism. This flexibility would allow a differentiation in the application of international standards between different companies and their respective products (for example, Facebook is different from Twitter). It would also make room for companies to adopt their own views on the speech that is allowed on their platforms, although market dominance would result in a narrower margin of manoeuvre in this respect.

Another point of divergence is whether the SMC should have an adjudicatory or an advisory role. In an advisory capacity, it would provide general guidance to social media companies on the compatibility of terms of service or community standards with international standards on human rights. In this configuration, the SMC would be an open forum where stakeholders could elaborate recommendations or observations. The alternative would be to give the SMC the power to review individual decisions; the council would then have to decide whether, in the particular circumstances of a case, the decision made by the social media platform conformed to the requirements of international human rights standards. Such a mechanism should be accessible to all. There should also be clear and precise rules of procedure on questions such as admissibility conditions, time limits, admissibility of evidence, elements covered by confidentiality, exchange of arguments and views, elements of publicity, and the adoption and publication of decisions.

The SMC model puts forward a voluntary approach to the oversight of content moderation.

ARTICLE 19 discussed the various possible orientations of an SMC, as well as some more technical issues such as the rules of procedure or the funding mechanism, in a background paper supporting a current online consultation.[1] ARTICLE 19 considers that the different visions for the SMCs are not mutually exclusive: they could be designed to be complementary. In that perspective, the question is not so much whether SMCs should be set at the global level or the national level — there are strong arguments for each — but how they could all work together. The local SMC, anchored in the local context, with its members very familiar with the complexities of the linguistic, social, cultural, economic and political circumstances of the country, would bring an increased credibility to the whole system by producing a nuanced understanding that a distant, international forum cannot reach, and it could develop solutions adapted to the local context. And the global SMC would bring a sense of universality to the system: it would elaborate a universal code of human-rights-based principles on content moderation, and it would provide a framework for national SMCs to resolve divergences. It is possible that a local SMC could bring valuable local expertise to the oversight board that Facebook is building, should that

particular experiment prove compatible with international standards on human rights; a memorandum of understanding between the oversight board and a local SMC could provide a framework within which the board could seek specific insights from the local body.

At the moment, there are various legislative initiatives that would rely on self-regulatory mechanisms within a legal framework of co-regulation, under the guise of bringing a swift end to the dissemination of often vaguely defined harmful content. The SMC offers a model that can deliver a form of co-regulation that fully ensures the protection of the fundamental right to freedom of expression. Moreover, the SMC model offers a stopgap between the state body charged with overseeing the self-regulatory mechanism and the social media companies, without which companies are likely to apply mechanisms and execute decisions that do not comply with international human rights standards.

The SMC is not the only idea that seeks to deal with the issue of content moderation as a matter of urgent democratic importance — see, for instance, the proposal for a moderation standards council (McKelvey, Tworek and Tenove 2019) and the model from Global Partners Digital (Bradley and Wingfield 2018). Not only is this question *dans l'air du temps*, it is also emerging at the exact point of convergence between the goals and interests of human rights groups and those of social media platforms: avoiding the pitfalls of harsh legislative approaches that often come with disproportionate sanctions; contributing to restoring trust from users through transparency and accountability; providing an effective yet adaptable form of regulation that can easily accommodate the constant evolution of tech platforms; and ensuring that moderation of speech is done on the universal grounds of international law. ARTICLE 19 is urging interested members of the public to be involved by exploring its presentation on the SMC and to share thoughts in a public survey.[2] Now is the time to help us shape the future of social media regulation.

AUTHOR'S NOTE

Views in this article do not necessarily reflect the positions of ARTICLE 19.

NOTES

1 Readers are invited to view the consultation paper and complete the consultation survey at www.article19.org/resources/social-media-councils-consultation/; the survey closes November 30, 2019.

2 Please visit www.article19.org/resources/social-media-councils-consultation/ for more information and to complete the survey before November 30, 2019.

WORKS CITED

ARTICLE 19. 2017. "Germany: The Act to Improve Enforcement of the Law in Social Networks." Legal analysis. London, UK: ARTICLE 19. www.article19.org/wp-content/uploads/2017/09/170901-Legal-Analysis-German-NetzDG-Act.pdf.

———. 2018. "Side-stepping rights: Regulating speech by contract." Policy brief. London, UK: ARTICLE 19. www.article19.org/wp-content/uploads/2018/06/Regulating-speech-by-contract-WEB-v2.pdf.

———. 2019. "Facebook oversight board: Recommendations for human rights-focused oversight." March 27. www.article19.org/resources/facebook-oversight-board-recommendations-for-human-rights-focused-oversight/.

Bradley, Charles and Richard Wingfield. 2018. "A Rights-Respecting Model of Online Content Regulation by Platforms." May. London, UK: Global Partners Digital. www.gp-digital.org/content-regulation-laws-threaten-our-freedom-of-expression-we-need-a-new-approach/.

Global Digital Policy Incubator, ARTICLE 19 and David Kaye. 2019. *Social Media Councils: From Concept to Reality*. Conference report, February. Stanford, CA: Global Digital Policy Incubator. https://cyber.fsi.stanford.edu/gdpi/content/social-media-councils-concept-reality-conference-report.

McKelvey, Fenwick, Heidi Tworek and Chris Tenove. 2019. "How a standards council could help curb harmful online content." *Policy Options*, February 11. https://policyoptions.irpp.org/magazines/february-2019/standards-council-help-curb-harmful-online-content/.

McSherry, Corynne. 2019. "Social Media Councils: A Better Way Forward, Window Dressing, or Global Speech Police?" Electronic Frontier Foundation, May 10. www.eff.org/fr/deeplinks/2019/05/social-media-councils-better-way-firward-lipstick-pig-or-global-speech-police.

Newman, Nic, Richard Fletcher, Antonis Kalogeropoulos and Rasmus Kleis Nielsen. 2019. *Reuters Institute Digital News Report 2019*. Oxford, UK: Reuters Institute for the Study of Journalism. www.digitalnewsreport.org/.

UN General Assembly. 2018. *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*. A/HRC/38/35, April 6. https://undocs.org/pdf?symbol=en/A/HRC/38/35.

ABOUT THE AUTHOR

Pierre François Docquir is a researcher and expert in the fields of human rights law and internet and media law and regulation. He is the head of the Media Freedom Programme at ARTICLE 19 after joining the organization in 2015 as the senior legal officer. He previously served as vice-president of the Conseil supérieur de l'audiovisuel in Belgium and was a researcher at the Université libre de Bruxelles, where he obtained his Ph.D. in law in 2009.

Subscribe at www.bigtechpodcast.com

big·tech

A podcast about technology's impact on our democracy, economy and society.

Join Co-hosts Taylor Owen and David Skok as they sit down with leading scholars, policy makers and entrepreneurs to discuss how emerging technologies are reshaping our democracy, economy and society.

PRESENTED BY

Centre for International Governance Innovation / The Logic

Joan Donovan

Navigating the Tech Stack: When, Where and How Should We Moderate Content?

On Saturday, August 3, 2019, a gunman opened fire in a Walmart in El Paso, Texas, killing 22 people and wounding 27 before he was taken into custody by police. As news of the attack spread, so did a white supremacist manifesto, allegedly written by the shooter and uploaded hours before the shooting to an anonymous forum, 8chan (Harwell 2019b). This document was archived and reproduced, amplified on other message boards and social media, and eventually reported in the press. This was the third mass shooting linked to the extremist haven 8chan in six months, and followed the same pattern as the synagogue shooting in Poway, California, in April and the Christchurch, New Zealand, mosque shootings in March: post a racist screed to 8chan; attack a targeted population; and influence national debates about race and nation.

What will it take to break this circuit, where white supremacists see that violence is rewarded with amplification and infamy? While the answer is not straightforward, there are technical and ethical actions available.

After the white supremacist car attack in Charlottesville, Virginia, in 2017, platform companies such as Facebook, Twitter and YouTube began to awaken to the fact that platforms are more than just a reservoir of content. Platforms are part of the battleground over hearts and minds, and they must coordinate to stop the amplification of white supremacy across their various services. Advocacy groups, such as Change the Terms,[1] are guiding platform companies to standardize and enforce content moderation policies about hate speech.

But, what happens to content not originating on major platforms? How should websites with extremists and white supremacist content be held to account, at the same time that social media platforms are weaponized to amplify hateful content?

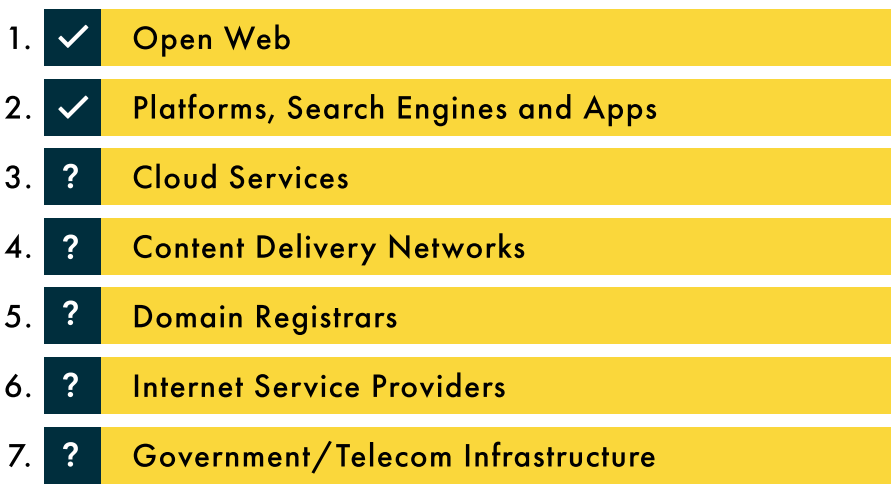
Following the El Paso attack, the original founder of 8chan has repeatedly stated that he believed the site should be shut down (quoted in Harwell 2019a), but he is no longer the owner. The long-term failure to moderate 8chan led to a culture of encouragement of mass violence, harassment and other depraved behaviour. Coupled with a deep commitment to anonymity, the current owner of 8chan resists moderation on principle, and back tracing content to original posters is nearly impossible. The more heinous the content, the more it circulates. 8chan, among other extremist websites, also contributed to the organization of the Unite the Right Rally in Charlottesville, where Heather Heyer was murdered in the 2017 car attack and where many others were injured.

In the wake of Charlottesville, corporations grappled with the role they played in supporting white supremacists organizing online (Robertson 2017). After the attack in Charlottesville and another later in Pittsburgh in October 2018, in which a gunman opened fire on the Tree of Life synagogue, there was a wave of deplatforming and corporate denial of service (Koebler 2018; Lorenz 2018), spanning cloud service companies (Liptak 2018), domain registrars (Romano 2017), app stores (O'Connor 2017) and payment servicers (Terdiman 2017). While some debate the cause and consequences of deplatforming specific far-right individuals on social media platforms, we need to know more about how to remove and limit the spread of extremist and white supremacist websites (Nouri, Lorenzo-Dus and Watkin 2019).

Researchers also want to understand the responsibility of technology corporations that act as the infrastructure allowing extremists to connect to one another and to incite violence. Corporate decision making is now serving as large-scale content moderation in times of crisis, but is corporate denial of service a sustainable way to mitigate white supremacists organizing online?

On August 5, 2019, one day after two mass shootings rocked the nation, Cloudflare, a content delivery network, announced a termination of service for 8chan via a blog post written by CEO Matthew Prince (2019). The decision came after years of pressure from activists. "Cloudflare is not a government," writes Prince, stating that his company's success in the space "does not give us the political legitimacy to make determinations on what content is good and bad" (ibid.). Yet, due to insufficient research and policy about moderating the unmoderatable and the spreading of extremist ideology, we are left with open questions about where content moderation should occur online.

Figure 1: Content Moderation in the Tech Stack



Source: Author.

When discussions of content moderation take a turn for the technical, we tend to hear a lot of jargon about "the tech stack" (Figure 1). It is important to understand how the design of technology also shows us where the power lies.

Most debates about content moderation revolve around individual websites' policies for appropriate participation (level 1) and about major platforms' terms of service (level 2). For example, on level 1, a message board dedicated to hobbies or the user's favourite TV show may have a policy against spamming ads or bringing up political topics. If users don't follow the rules, they might get a warning or have their account banned.

On level 2, there is a lot of debate about how major corporations shape the availability and discoverability of information. While platforms, search engines and apps have policies against harassment, hate and incitement to violence, it is difficult to enforce these policies given the enormous scale of user-generated content. In Sarah Roberts's new book *Behind the Screen: Content Moderation in the Shadows of Social Media* (2019), she documents how this new labour force is tasked with removing horrendous violence and pornography daily, while being undervalued despite their key roles working behind the scenes at the major technology corporations. Because of the commercial content moderation carried out by these workers, 8chan and other extremist sites cannot depend on social media to distribute their wares.

For cloud service providers on level 3, content moderation occurs in cases where sites are hosting stolen or illegal content. Websites with fraught content, such as 8chan, will often mask or hide the location of their servers to avoid losing hosts. Nevertheless, actions by cloud service companies post-Charlottesville did destabilize the ability for the so-called alt-right to regroup quickly (Donovan, Lewis and Friedberg 2019).

On level 4 of the tech stack, content delivery networks (CDNs) help match user requests with local servers to reduce network strain and speed up websites. CDNs additionally provide protection from malicious access attempts, such as distributed denial-of-service attacks that overwhelm a server with fake traffic. Without the protection of CDNs such as Cloudflare or Microsoft's Azure, websites are vulnerable to political or profit-driven attacks, such as a 2018 attempt to overwhelm Github (Kottler 2018) or a 2016 incident against several US banks (Volz and Finkle 2016).

It is difficult to enforce these policies given the enormous scale of user-generated content.

Cloudflare, the CDN supporting 8chan, responded by refusing to continue service to 8chan. Despite attempts to come back online, 8chan has not been able to find a new CDN at this time (Coldewey 2019).

In the aftermath of Charlottesville, Google froze the domain of a neo-Nazi site that organized the event and GoDaddy also refused services (Belvedere 2017). In response to the El Paso attack, another company is taking action. Tucows, the domain registrar of 8chan, has severed ties with the website (Togoh 2019). It is rare to see content decisions on level 5 of the tech stack, except in the cases of trademark infringement, blacklisting by a malware firm or government order.

Generally speaking, cloud services, CDNs and domain registrars are considered the backbone of the internet, and sites on the open Web rely on their stability, both as infrastructure and as politically neutral services.

Level 6 is a different story. Internet service providers (ISPs) allow access to the open Web and platforms, but these companies are in constant litigious relations with consumers and

the state. ISPs have been seen to selectively control access and throttle bandwidth to content profitable for them, as seen in the ongoing net neutrality fight. While the divide between corporations that provide the infrastructure for our communication systems and the Federal Communications Commission is overwhelmed by lobbying (West 2017), the US federal and local governments remain unequipped to handle white supremacist violence (Andone and Johnston 2017), democratic threats from abroad (Nadler, Crain and Donovan 2018), the regulation of tech giants (Lohr, Isaac and Popper 2019) or the spread of ransomware attacks in cities around the country (Newman 2019). However, while most ISPs do block piracy websites, at this stage we have not seen US ISPs take down or block access to extremist or white supremacist content. Other countries, for example, Germany, are a different case entirely as they do not allow hate speech or the sale of white supremacist paraphernalia (Frosch, Elinson and Gurman 2019).

The US federal and local governments remain unequipped to handle white supremacist violence.

Lastly, on level 7, some governments have blacklisted websites and ordered domain registrars to remove them (Liptak 2008). Institutions and businesses can block access to websites based on content. For example, a library will block all manner of websites for reasons of safety and security. In the case of 8chan, while US President Trump has called for law enforcement to work with companies to "red flag" posts and accounts, predictive policing has major drawbacks (Munn 2018).

At every level of the tech stack, corporations are placed in positions to make value judgments regarding the legitimacy of content, including who should have access, and when and how. In the case of 8chan and the rash of premeditated violence, it is not enough to wait for a service provider, such as Cloudflare, to determine when a line is crossed. Unfortunately, in this moment, a corporate denial of service is the only option for dismantling extremist and white supremacist communication infrastructure.

The wave of violence has shown technology companies that communication and coordination flow in tandem. Now that technology corporations are implicated in acts of massive violence by providing and protecting forums for hate speech, CEOs are called to stand on their ethical principles, not just their terms of service. For those concerned about the abusability of their products, now is the time for definitive action (Soltani 2019). As Malkia Cyril (2017) of Media Justice argues, "The open internet is democracy's antidote to authoritarianism." It's not simply that corporations can turn their back on the communities caught in the crosshairs of their technology. Beyond reacting to white supremacist violence, corporations need to incorporate the concerns of targeted communities and design technology that produces the Web we want.

Regulation to curb hateful content online cannot begin and end with platform governance. Platforms are part of a larger online media ecosystem, in which the biggest platforms not only contribute to the spread of hateful content, but are themselves an important vector of attack, increasingly so as white supremacists weaponize platforms to distribute racist manifestos. It is imperative that corporate policies be consistent with regulation on hate speech across many countries. Otherwise, corporate governance will continue to be not merely haphazard but potentially endangering for those who are advocating for the removal of hateful content online. In effect, defaulting to the regulation of the country with the most stringent laws on hate speech, such as Germany, is the best pathway forward for content moderation, until such time that a global governance strategy is in place.

ACKNOWLEDGEMENTS

Brian Friedberg, Ariel Herbert-Voss, Nicole Leaver and Vanessa Rhinesmith contributed to the background research for this piece.

NOTE

1 See www.changetheterms.org/.

WORKS CITED

Andone, Dakin and Chuck Johnston. 2017. "Report on Charlottesville rally faults police over planning, failure to protect public." CNN, December 2. www.cnn.com/2017/12/01/us/charlottesville-riots-failures-review/index.html.

Belvedere, Matthew J. 2017. "GoDaddy CEO: We booted the neo-Nazi Daily Stormer website for inciting violence." CNBC, August 15. www.cnbc.com/2017/08/15/godaddy-ceo-we-booted-the-neo-nazi-daily-stormer-website-for-inciting-violence.html.

Coldewey, Devin. 2019. "8chan's new internet host was kicked off its own host just hours later." *Techcrunch*, August 5. https://techcrunch.com/2019/08/05/8chans-new-internet-host-was-kicked-off-its-own-host-just-hours-later/.

Cyril, Malkia A. 2017. "The Antidote to Authoritarianism: Without the open internet, Americans lose an essential tool in the fight against discriminatory mass surveillance." *The Atlantic*, May 8. www.theatlantic.com/technology/archive/2017/05/the-antidote-to-authoritarianism/525438/.

Donovan, Joan, Becca Lewis and Brian Friedberg. 2019. "Parallel Ports: Sociotechnical Change from the Alt-Right to Alt-Tech." https://dx.doi.org/10.14361/9783839446706-004.

Frosch, Dan, Zusha Elinson and Sadie Gurman. 2019. "White Nationalists Pose Challenge to Investigators." *The Wall Street Journal*, August 5. www.wsj.com/articles/shootings-highlight-law-enforcement-challenges-to-combating-domestic-terror-11564947769.

Harwell, Drew. 2019a. "Three mass shootings this year began with a hateful screed on 8chan. Its founder calls it a terrorist refuge in plain sight." *The Washington Post*, August 4. www.washingtonpost.com/technology/2019/08/04/three-mass-shootings-this-year-began-with-hateful-screed-chan-its-founder-calls-it-terrorist-refuge-plain-sight/.

———. 2019b. "8chan vowed to fight on, saying its 'heartbeat is strong.' Then a tech firm knocked it offline." *The Washington Post*, August 5. https://beta.washingtonpost.com/technology/2019/08/05/defiant-chan-vowed-fight-saying-its-heartbeat-is-strong-then-tech-firm-knocked-it-offline/.

Koebler, Jason. 2018. "Deplatforming Works." *Vice*, August 10. www.vice.com/en_us/article/bjbp9d/do-social-media-bans-work.

Kottler, Sam. 2018. "February 28th DDoS Incident Report." *GitHub* (blog), March 1. https://github.blog/2018-03-01-ddos-incident-report/.

Liptak, Andrew. 2008. "A Wave of the Watch List, and Speech Disappears." *The New York Times*, March 4. www.nytimes.com/2008/03/04/us/04bar.html.

———. 2018. "Two more platforms have suspended Gab in the wake of Pittsburgh shooting." *The Verge*, October 28. www.theverge.com/2018/10/28/18034126/gab-social-network-stripe-joyent-deplatforming-hate-speech-pittsburgh-shooting.

Lohr, Steve, Mike Isaac and Nathaniel Popper. 2019. "Tech Hearings: Congress Unites to Take Aim at Amazon, Apple, Facebook and Google." *The New York Times*, July 16. www.nytimes.com/2019/07/16/technology/big-tech-antitrust-hearing.html.

Lorenz, Taylor. 2018. "The Pittsburgh Suspect Lived in the Web's Darkest Corners." *The Atlantic*, October 27. www.theatlantic.com/technology/archive/2018/10/what-gab/574186/.

Munn, Nathan. 2018. "This Predictive Policing Company Compares Its Software to 'Broken Windows' Policing." *Vice*, June 11. www.vice.com/en_us/article/d3k5pv/predpol-predictive-policing-broken-windows-theory-chicago-lucy-parsons.

Nadler, Anthony, Matthew Crain and Joan Donovan. 2018. *Weaponizing the Digital Influence Machine: The Political Perils of Online Ad Tech*. October 17. New York, NY: Data & Society Research Institute. https://datasociety.net/wp-content/uploads/2018/10/DS_Digital_Influence_Machine.pdf.

Newman, Lily Hay. 2019. "Ransomware Hits Georgia Courts as Municipal Attacks Spread." *Wired*, July 1. www.wired.com/story/ransomware-hits-georgia-courts-municipal-attacks-spread/.

Nouri, Lella, Nuria Lorenzo-Dus and Amy-Louise Watkin. 2019. "Following the Whack-a-Mole: Britain First's Visual Strategy from Facebook to Gab." Global Research Network on Terrorism and Technology Paper No. 4. London, UK: Royal United Services Institute, July 4. https://rusi.org/publication/other-publications/following-whack-mole-britain-firsts-visual-strategy-facebook-gab.

O'Connor, Clare. 2017. "Google Drops Alt-Right Favorite Gab From Play Store As Internet Hate Speech Purge Continues." *Forbes*, August 18. www.forbes.com/sites/clareoconnor/2017/08/18/google-drops-alt-right-favorite-gab-from-play-store-as-internet-hate-speech-purge-continues/#3cf6d11c61c5.

Prince, Matthew. 2019. "Terminating Service for 8Chan." *Cloudflare* (blog), August 5. https://blog.cloudflare.com/terminating-service-for-8chan/.

Roberts, Sarah. 2019. *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven, CT: Yale University Press.

Robertson, Adi. 2017. "Two months ago, the internet tried to banish Nazis. No one knows if it worked." *The Verge*, October 9. www.theverge.com/2017/10/9/16446920/internet-ban-nazis-white-supremacist-hosting-providers-charlottesville.

Romano, Aja. 2017. "Neo-Nazi site Daily Stormer resurfaces with Russian domain following Google and GoDaddy bans." *Vox*, August 16. www.vox.com/culture/2017/8/16/16156210/daily-stormer-russian-domain-godaddy-google-ban.

Soltani, Ashkan. 2019. "Abusability Testing: Considering the Ways Your Technology Might Be Used for Harm." USENIX.org, January 28. www.usenix.org/node/226468.

Terdiman, Daniel. 2017. "After Charlottesville, PayPal says it won't do business with hate groups." *Fast Company*, August 15. www.fastcompany.com/40454274/after-charlottesville-paypal-says-it-wont-do-business-with-hate-groups.

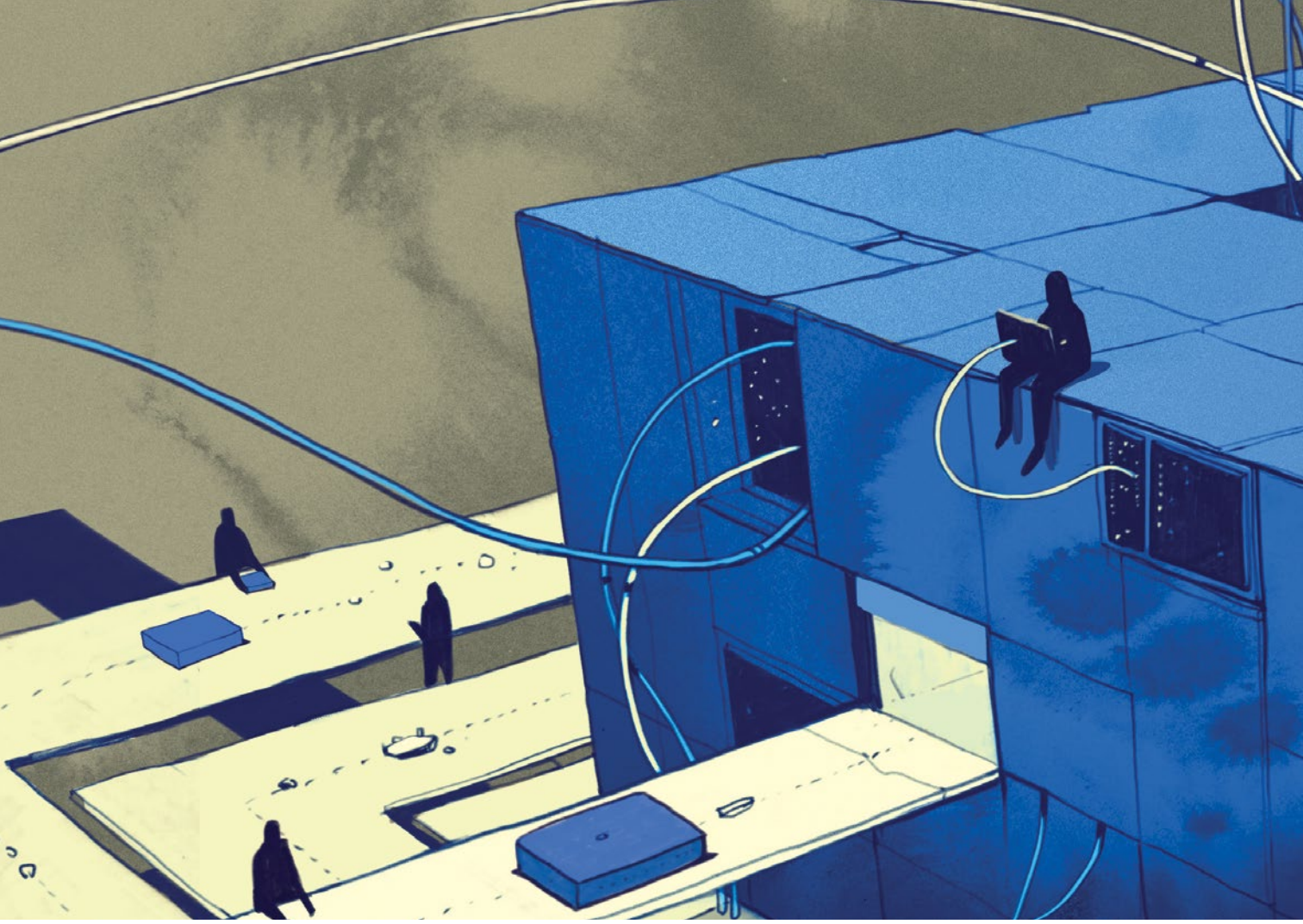
Togoh, Isabel. 2019. "Tucows Drops 8chan Domain Registration After El Paso Shooting." *Forbes*, August 5. www.forbes.com/sites/isabeltogoh/2019/08/05/tucows-drops-8chan-domain-registration-after-el-paso-shooting/#54ec871affbf.

Volz, Dustin and Jim Finkle. 2016. "U.S. indicts Iranians for hacking dozens of banks, New York dam." Reuters, March 24. www.reuters.com/article/us-usa-iran-cyber/u-s-indicts-iranians-for-hacking-dozens-of-banks-new-york-dam-idUSKCN0WQ1JF.

West, Geoff. 2017. "Money flows into net neutrality debate ahead of FCC vote." Center for Responsive Politics, December 14. www.opensecrets.org/news/2017/12/money-flows-into-net-neutrality-debate-ahead-of-fcc-vote/.

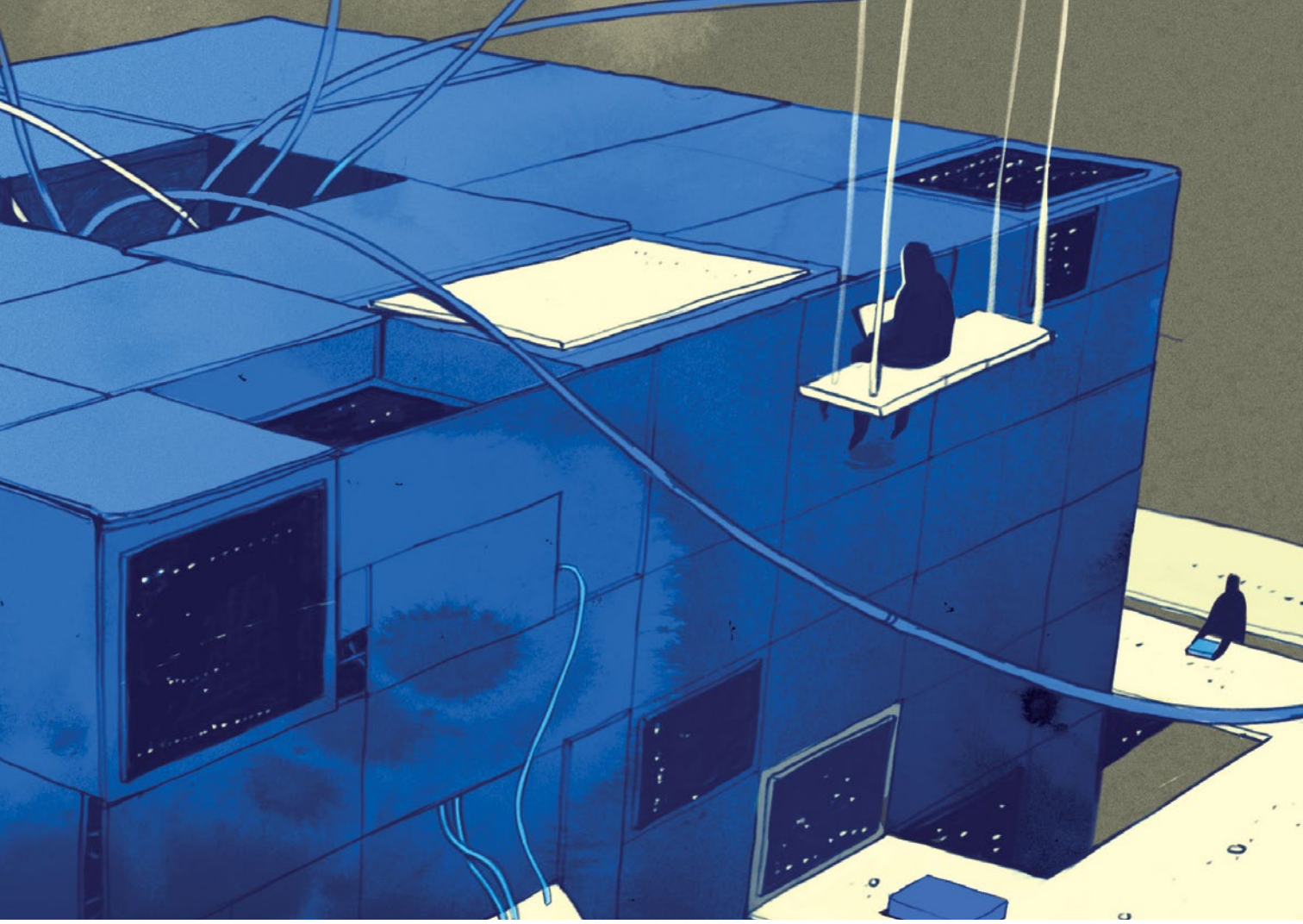
ABOUT THE AUTHOR

Joan Donovan is the director and lead researcher of the Technology and Social Change Research Project at the Shorenstein Center on Media, Politics and Public Policy at Harvard Kennedy School. Joan's research and teaching interests are focused on media manipulation, effects of disinformation campaigns, and adversarial media movements.



Susan Etlinger

What's So Difficult about Social Media Platform Governance?



During the past few years, the global conversation about responsible technology has intensified. Increasingly, we are acknowledging that technology is not and can never be neutral, that it holds significant implications for people and society, and that intelligent technologies have consequences that can disenfranchise or target vulnerable populations.

None of this is news to historians of technology, sociologists, political scientists and others who study human behaviour and history. Yet, it seems nothing was further from the minds of the social media platforms' founders when they conceived of their services. These tech giants' products now act as the de facto communication channels for a significant portion — up to half — of the 4.3 billion internet users worldwide.

Even a brief glance at these companies' websites[1] offers a window into their conceptions of themselves. Facebook's mission is "to give people the power to build community and bring the world closer together." Twitter's is "to give everyone the power to create and share ideas and information instantly without barriers." YouTube aims "to give everyone a voice and show them the world."

These companies have been woefully unprepared for the ways their platforms could be weaponized. Facebook, Twitter and YouTube were used to distribute and amplify misinformation in both the Brexit vote in the United Kingdom and the 2016 presidential election in the United States.[2]

They were used to livestream the terror attacks on two mosques in Christchurch, New Zealand.[3] They've been used as a tool for ethnic cleansing in Myanmar (Mozur 2018). White supremacists and other similar groups regularly use them to spread hate speech.

Clearly, these effects stand at odds not only with democratic values but with the platforms' stated intentions and ideals. In response, the

companies have instituted or augmented existing fairness, trust and safety teams to address these issues organizationally and systemically. They've built principles and checklists and thick, detailed volumes of community guidelines. They've contracted with companies that employ thousands of moderators around the world to review content.

Still, the incidents continue. Given the threats to democracy and society, as well as their considerable resources and technical capability, what exactly is preventing social media companies from more effectively mitigating disinformation, hate speech and aggression on their platforms?

Addressing these issues is not as straightforward as it seems. In addition to the legal, social and cultural dynamics at play, there are other factors we must consider: the scale of social media platforms; the technologies on which they are built; and the economic environments in which they operate. Any one of these factors alone would present significant challenges — the scale of operations, the sophisticated yet brittle nature of intelligent technologies and the requirements of running publicly traded companies — but in concert they prove far more complex.

The Scale of Social Media

The scale of social media today is unprecedented. For example, in contrast to the early days — when Facebook was created as a site to rate women's attractiveness, before it morphed into a kind of digital yearbook — the company today has 2.41 billion monthly active users (as of June 30, 2019), and "more than 2.1 billion people use Facebook, Instagram, WhatsApp, or Messenger every day on average."[4] YouTube has over one billion users — comprising approximately one-third of internet users — and operates in 91 countries in 80 different languages. According to YouTube, people around the world watch over a billion hours of video every single day. Twitter has 139 million daily active users, who send 500 million tweets per day. All these numbers add up to a level of impact and complexity in communications that has never before been seen in human history.

Understanding the Technology Issues

Social media platforms today are far more technically sophisticated than when they were introduced in the mid 2000s. One of the key changes is the use of what are broadly referred to as artificial intelligence (AI) technologies, which enable the platforms to recognize and interpret language and images and use data to develop models that draw inferences, classify information and predict interests, among other things.

These capabilities enable the platforms to recommend videos, products or political posts to us, translate posts from languages different from our own, identify people and objects, and provide accessibility options to people with visual impairments or other needs. They can also be used to identify potentially problematic content or account behaviour.

What's fundamentally different about these technologies is that, unlike computer programs of the past, AI learns from data and behaves autonomously in some circumstances. But, as sophisticated as they are becoming, intelligent technologies pose serious trade-offs. There are several core issues at play.

Safety versus Freedom of Speech

When it comes to content moderation, AI programs are not adept at understanding context and nuance, so they make mistakes that can result in "false positives" (flagging an innocuous video, statement or photo) or "false negatives" (missing a violent or otherwise undesirable post). In the world of social media, false positives prompt protests over censorship, for example, when a platform removes a post by an organization that is sharing it to raise awareness of a human rights violation, while false negatives expose the company to legal liability, if, say, it fails to recognize and remove prohibited content within a stipulated time period.

As a result, social media companies use human content moderators to review ambiguous posts, a task some have dubbed "the worst job in history" (Weber and Seetharaman 2017). Content moderators, often employees of large contracting firms, watch and classify hundreds of thousands of posts per day. Some of the tasks, such as labelling places and

classifying animals, are generally harmless, if rather tedious, but the people who deal with the most extreme content can develop symptoms of post-traumatic stress triggered by ongoing exposure to disturbing imagery. Some content moderators, after repeated exposure to certain material, begin to "embrace the fringe viewpoints of the videos and memes that they are supposed to moderate" (Newton 2019).

The essentially subjective, probabilistic and often culturally specific nature of image and language classification creates additional complexity. Bias, implicit or explicit, tends to become amplified (Zhao et al. 2017). One person's protest is another person's riot. One person's treasured photo of a family member breastfeeding her child is another person's pornography.

Facebook's 2016 removal of "Napalm Girl," Nick Ut's historic and Pulitzer Prize-winning photograph of a naked nine-year-old girl, Kim Phúc, screaming in pain as she fled a napalm attack during the Vietnam War, highlights both the contextual nature of media and the constraints of algorithmic decision making. Facebook's decision to remove the photo prompted an open letter to the company in *Aftenposten*, Norway's largest newspaper, charging the company with censorship and abuse of power (Hansen 2016). Facebook countered, "While we recognize that this photo is iconic, it's difficult to create a distinction between allowing a photograph of a nude child in one instance and not others" (quoted in Levin, Wong and Harding 2016). Ultimately, the company reinstated "Napalm Girl."

But Facebook's argument cannot be viewed as an isolated editorial choice; it's a systems issue as well. The algorithms that the company uses to moderate content learn from decisions, so social media platforms must also consider the precedents they are setting in their systems, what unintended and undesirable consequences such decisions might unleash, and what possible technical interventions could prevent those unwanted consequences from occurring in the future, all across dozens of languages, multiple varieties of data (text, photo, meme, video) and millions of posts per day.

Limitations of Language Technologies

While language technology continues to improve rapidly, it remains highly dependent on high volumes of labelled and clean data to achieve an acceptable level of accuracy. "To determine in real-time whether someone is saying something that suggests a reasonable likelihood of harm," says Carole Piovesan, partner and co-founder of INQ Data Law, "requires a team of local interpreters to understand the conversation in context and in the local dialect."[5] From a practical standpoint, it is much easier for algorithms and content moderators to interpret a fast-moving event in a language spoken by hundreds of millions of people than it is to interpret content in a language or dialect spoken by a small population.

One person's protest is another person's riot.

This type of scenario also requires a process to ensure that there is consensus on the meaning of and action to be taken on a questionable post — a process that is critical but time-consuming in situations where even milliseconds can determine how far a message will spread. This makes it difficult, if not impossible, for social media companies to comply with laws that require removal of prohibited posts within the hour. Furthermore, as Piovesan points out, there are secondary effects we must consider: "If you were to offer your service only in those regions where you have enough competent interpreters to properly understand the conversation, you risk increasing the divide between those who have access to social platforms and those who do not."

Lack of Digital Norms

Social networks now find themselves having to make decisions, often imperfect ones, that affect people across the world. It would be far simpler for them to establish a common culture and set of norms across their platforms. But, says Chris Riley, director of public policy at Mozilla, "that would be very hard to do without meeting the lowest common denominator of every country in which they

operate. It's not possible to engineer a social network that meets the social expectations of every country. Furthermore, it's not advisable to try, because to do so would promote cultural hegemony."[6] This does not mean that it is fruitless for the global community to attempt to agree upon governance structures. Rather, these structures must be flexible enough to account for language, culture, access, economic and other variables without introducing undesirable unintended consequences.

The Economics of Social Media Companies

Social media companies' mission statements focus on sharing, community and empowerment. But their business models are built on, and their stock prices rise and fall every quarter on the strength of, their ability to grow, as measured in attention and engagement metrics: active users, time spent, content shared.

But this isn't simply a question of profit. The large social media companies — Facebook (including Instagram, Messenger and WhatsApp), Twitter and YouTube (owned by Google) — are publicly traded. They must meet their responsibility to society while also upholding their fiduciary responsibility to shareholders, all while navigating a complex web of jurisdictions, languages and cultural norms. By no means does this excuse inaction or lack of accountability. Rather, it illustrates the competing access needs of a wide range of stakeholders and the dynamics that must inform governance structures in the digital age.

Considerations for Global Platform Governance

The realities of scale, technology and economics will continue to challenge efforts to implement practical and effective governance frameworks. The following are a few ideas that warrant further discussion from a governance perspective.

Secure Mechanisms for Sharing Threat-related Data

We will continue to see bad actors weaponize the open nature of the internet to destabilize democracy and governments, sow or stifle dissent and/or influence elections, politicians, and people.[7] Technologies such as deepfakes (O'Sullivan 2019) and cheapfakes (Harwell 2019) will only make it harder to distinguish between legitimate content and misinformation. We'll also see human (troll) and bot armies continue to mobilize to plant and amplify disinformation, coordinate actions across multiple accounts and utilize "data voids" (Golebiewski and boyd 2018) and other methods to game trending and ranking algorithms and force their messages into the public conversation.

As we have seen, this behaviour typically occurs across multiple platforms simultaneously. A sudden increase on Twitter in the number and activity of similarly named accounts in a particular region may be a precursor to a campaign, a phenomenon that could also be occurring simultaneously on Facebook or another social network.

But, as in the cyber security world, social media platforms are limited in the extent to which they are able to share this type of information with each other. This is partially a function of the lack of mechanisms for trust and sharing among competitors, but, says Chris Riley, "it's also hard to know what to share to empower another platform to respond to a potential threat without accidentally violating privacy or antitrust laws." In effect, this limitation creates an asymmetric advantage for bad actors, one that could potentially be mitigated if social media platforms had a secure and trustworthy way to work together.

Third-party Content Validation

One approach that has been discussed in policy circles is the idea of chartering third-party institutions to validate content, rather than having the social media platforms continue to manage this process. Validation could potentially become the responsibility of a non-partisan industry association or government body, leaving platforms to focus their resources on account behaviour, which is, generally speaking, less subjective and more tractable than content.

Viewing Takedowns in Context

The outcry over Facebook's 2016 decision to remove the "Napalm Girl" image from its

platform illustrates why it's so important to look at content takedowns not in isolation but in context. Looking at them in context would require understanding the reason for removing the content: whether it was a purely automated or human-influenced decision, the policies according to which the platform took that decision, the speed at which it was accomplished and the available appeals process. As fraught and newsworthy as content takedowns may sometimes be, it's neither productive nor sustainable to view them in isolation. Rather, governance structures should examine the organization's commitment to responsibility over time as evidenced by process, policy, resource allocation and transparency.

Conclusion

As we consider governance frameworks to address the issues presented by social media, we must also consider their implications for broader business and social ecosystems and, to the extent possible, plan for the emergence of technologies that will present novel challenges and threats in the future. These plans will require a measure of flexibility, but that flexibility is critical to ensure that we become and remain equal to the task of combatting disinformation, protecting human rights and securing democracy. Whatever path we choose, one thing is clear: the global digital governance precedents we set for social media today will affect us, individually and collectively, far into the future.

NOTES

1 See mission statements for Facebook, Twitter and YouTube, respectively, at https://investor.fb.com/resources/default.aspx, https://investor.twitterinc.com/contact/faq/default.aspx and www.youtube.com/about/.

2 See www.theguardian.com/news/series/cambridge-analytica-files.

3 See www.christchurchcall.com/.

4 See https://newsroom.fb.com/company-info/.

5 All quotes by C. Piovesan from phone interview conducted by the author, September 10, 2019.

6 All quotes by C. Riley from a meeting with the author on August 6, 2019, San Francisco, CA.

7 See https://datasociety.net/research/media-manipulation/.

WORKS CITED

Golebiewski, Michael and danah boyd. 2018. "Data Voids: Where Missing Data Can Easily Be Exploited." *Data & Society*, May 11. https://datasociety.net/output/data-voids-where-missing-data-can-easily-be-exploited/.

Hansen, Espen Egil. 2016. "Dear Mark. I am writing this to inform you that I shall not comply with your requirement to remove this picture." *Aftenposten*, September 8. www.aftenposten.no/meninger/kommentar/i/G892Q/Dear-Mark-I-am-writing-this-to-inform-you-that-I-shall-not-comply-with-your-requirement-to-remove-this-picture.

Harwell, Drew. 2019. "Faked Pelosi videos, slowed to make her appear drunk, spread across social media." *The Washington Post*, May 24. www.washingtonpost.com/technology/2019/05/23/faked-pelosi-videos-slowed-make-her-appear-drunk-spread-across-social-media/?noredirect=on.

Levin, Sam, Julia Carrie Wong and Luke Harding. 2016. "Facebook backs down from 'napalm girl' censorship and reinstates photo." *The Guardian*, September 9. www.theguardian.com/technology/2016/sep/09/facebook-reinstates-napalm-girl-photo.

Mozur, Paul. 2018. "A Genocide Incited on Facebook, With Posts From Myanmar's Military." *The New York Times*, October 15. www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html.

Newton, Casey. 2019. "The Trauma Floor: The secret lives of Facebook moderators in America." *The Verge*, February 25. www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona.

O'Sullivan, Donie. 2019. "When seeing is no longer believing. Inside the Pentagon's race against deepfake videos." CNN Business. www.cnn.com/interactive/2019/01/business/pentagons-race-against-deepfakes/.

Weber, Lauren and Deepa Seetharaman. 2017. "The Worst Job in Technology: Staring at Human Depravity to Keep It Off Facebook." *The Wall Street Journal*, December 27. www.wsj.com/articles/the-worst-job-in-technology-staring-at-human-depravity-to-keep-it-off-facebook-1514398398.

Zhao, Jieyu, Tianlu Wang, Mark Yatskar, Vicente Ordonez and Kai-Wei Chang. 2017. "Men Also Like Shopping: Reducing Gender Bias Amplification using Corpus-level Constraints." arXiv.org, July 29. https://arxiv.org/pdf/1707.09457.pdf.

ABOUT THE AUTHOR

Susan Etlinger is a senior fellow at CIGI, where her work focuses on AI and big data. She is also an industry analyst at Altimeter Group, a division of Prophet Brand Strategy, and a globally recognized expert in digital strategy. An author of a series of reports and frameworks on topics including AI, data, analytics and digital ethics, Susan is also a frequent keynote speaker at academic and industry conferences around the world. Her TED talk, "What Do We Do with All This Big Data?" has been translated into 25 languages and viewed more than 1.3 million times. Susan's research is regularly cited in university curricula, and she has been quoted in numerous media outlets including *The Wall Street Journal*, *Fast Company*, *The New York Times* and NPR.

Robert Fay

Digital Platforms Require a Global Governance Framework

Platforms are at the core of the digital economy. They form its backbone and are its conduits. They are used for search, social engagement and knowledge sharing, and as labour exchanges and marketplaces for goods and services. Activities on platforms are expanding at a tremendous pace that is likely to continue, especially with 5G implementation looming. Platforms such as Google, Facebook, Twitter and Amazon span the globe, serve billions of users and provide core functions of our society, analogous to the role served by public utilities. But the governance around their functions is not well developed as it is for public utilities. In fact, there is a governance chasm.

Indeed, while platforms are pervasive in everyday life, the governance across the scale of their activities is ad hoc, incomplete and insufficient. They are a ready and often the primary source of information for many people and firms, which can improve consumer choice and market functioning. Yet, this information may be inaccurate, by design or not, and used to influence the actions of individuals — and, recently, the outcomes of elections. Their operations are global in scope, but regulation, the little that exists, is domestic in nature. They help to facilitate our private lives, but can also be used to track and intrude into our private lives. The use of private data is opaque, and the algorithms that power the platforms are essentially black boxes. This situation is unacceptable.

To be sure, there are many governance initiatives under way. Some countries are developing national strategies for artificial intelligence (AI) and big data. Some are examining and developing policy responses to the issue and implications of fake news. Several are developing national cyber strategies. Many are revisiting and revising legislation around privacy. The Group of Seven and the Group of Twenty (G20) have begun some initiatives, as have the Organisation for Economic Co-operation and Development (OECD) and the United Nations. Even the platforms themselves have called for some form of regulation.

But as yet there is no comprehensive global discussion or action. Governance innovation is required to create an *integrated* framework at the national and international levels. This framework needs a broad combination of policies, principles, regulations and standards, and developing it will involve experimentation, iteration and international coordination, as well as the engagement of a wide variety of stakeholders.

The Current Situation Has Been Seen Before

In some senses, the current situation is reminiscent of the rapid development of financial services globally in the 1990s and 2000s. Fuelled by light-touch regulation, and in no small measure by hubris, banks grew in size and power, leading to some exceptionally large global banks. In many instances, this expansion was encouraged by the prevailing notion that it was a global good, creating new financial services for new customers with greater efficiency, via financial wizardry that turned out to be opaque in terms of network effects, risks and consequences. The view then was that self-interest and reputation would constrain bad behaviour.

Sound familiar? We witnessed the significant social consequences that resulted, and the plummet in the public's trust in institutions.

The current regulatory framework around platforms epitomizes light-touch regulation. Like the few global banks that had dominated financial services before the Great Recession, there are a few global tech giants that dominate platforms, but how they operate is opaque. And they are exhibiting bad behaviour more and more frequently through a tangled web of connections so complicated that it would take a machine learning algorithm to figure them out. More insidious is the surveillance capitalism operating via the advertising-driven business model of many of the platforms. Further, there are concerns shared around the globe, regardless of individual societies' different values, about how information is being used, from issues of privacy to monetization.

The current regulatory framework around platforms epitomizes light-touch regulation.

Indeed, the potential negative impact of the misuse of information collected by the platforms would make the negative impact of the global financial crisis pale in comparison, given how technologies permeate every aspect of our lives and will continue to do so, both at an increasing pace and in ways we cannot even envisage right now. Indeed, the Internet of Things, 5G and digital identities embody systemic risk, through their interconnectedness. And the risks are profound: from cyber warfare to state surveillance and privacy invasion, to data breaches and large economic and personal income losses and, ultimately, a loss in trust. And there is an East-West geopolitical divide as the United States and China compete head-to-head for supremacy in the data and AI realm with others caught in the middle.

Yet, the potential for these technologies to improve the everyday lives of people is substantial and derives in part through the interconnectedness that is based on trust. In fact, the greatest benefits arise when all can participate in one global internet economy rather than in a digital economy splintered into different realms.

The time to act — globally — is now: to develop a governance framework, globally, and to ensure that these technologies are used for the greater good, globally.

A Model and a Way Forward

One way forward is to draw from the lessons of the financial crisis and how policy makers dealt with the large economic and financial issues that resulted. In particular, in the heart of the crisis, the Financial Stability Board (FSB) was created (from the existing Financial Stability Forum) and given a mandate by the G20 to promote the reform of international financial regulation and supervision, with a role in standard setting and promoting members' implementation of international standards.

Some background about the functioning of the FSB will help to set the stage. The main decision-making body is the plenary, which consists of representatives of all members: 59 representatives from 25 jurisdictions; six representatives from four international financial institutions; and eight representatives from six international standard-setting, regulatory, supervisory and central bank bodies. In carrying out its work the FSB promotes global financial stability by "coordinating the development of regulatory, supervisory and other financial sector policies and conducts

outreach to non-member countries. It achieves cooperation and consistency through a three-stage process."[1] The three-stage process consists of a vulnerabilities assessment, policy development and implementation monitoring. Each area has several working groups that comprise not only individuals from member countries and international standard-setting bodies, but also from non-member countries and organizations that may also be affected.

Reforming international financial regulation and supervision was a daunting task — indeed, it's work that continues to this day — given that mandates and cultures vary tremendously across the institutions involved in regulating and delivering financial services — central banks, private banks, capital markets, securities and insurance regulators, standard setters, policy makers and so on. And reform efforts met with tremendous resistance, including complaints about rising regulatory burdens and costs. Ultimately, however, it was clear that regulation had been too lax and urgently needed to be addressed.

The FSB's innovative multi-stakeholder processes have been essential to carrying out its responsibilities. Despite the daunting task that the FSB faced, the significant progress in financial sector reform is proof that these processes can achieve real and substantive reforms. These processes provide a model that might be useful for platform governance, but how we do go about adapting them?

Create a Digital Stability Board

A way forward is to create a new institution — let's call it the Digital Stability Board (DSB) — and give it a mandate by global leaders. A plenary body would set objectives and oversee work of the DSB and consist of officials from countries who initially join the organization. In addition, the DSB would work with standard-setting bodies, governments and policy makers, regulators, civil society and the platforms themselves via a set of working groups with clear mandates that would report back to the plenary. For example, the broad objectives for the DSB could be to:

- **Coordinate** the development of standards, regulations and policies across the many realms that platforms touch. The areas would include — but not be limited to — governance along the data and AI value chain (including areas such as privacy, ethics, data quality and portability, algorithmic accountability, etc.); social media content; competition policy; and electoral integrity. The objective of coordination would be to develop a set of principles and standards that could be applied globally while allowing for domestic variation to reflect national values and customs.

- **Monitor** development, advise on best practices, and consider regulatory and policy actions needed to address vulnerabilities in a timely manner.

- **Assess** vulnerabilities arising from these technologies, including their impact on civil society and the regulatory and policy actions needed to address them on a timely basis.

- **Ensure** that this work feeds into other organizations such as the World Trade Organization, which needs to modernize trade rules to reflect big data and AI, but also to develop a framework with which to assess the implications for trade and trade rule compliance.

The goal is not to reinvent work already in progress. There are many notable and substantive initiatives across the globe that could be drawn into the DSB, but they are generally not coordinated and, in many cases, have narrow mandates that may not be representative of wider interests. The goal is to coordinate these efforts and fill in gaps as required. The following are some initiatives already begun.

Standard Setting

The Institute of Electrical and Electronics Engineers (IEEE) has launched the Global Initiative on Ethics of Autonomous and Intelligent Systems (IEEE Global Initiative). The International Telecommunication Union has its Global Symposium for Regulators. Domestic equivalents would also be drawn in as required. The International Organization for Standardization and the International Electrotechnical Commission have begun standards development activities under the aegis of the Joint Technical Committee

(JTC 1). The FSB itself is examining the implications of fintech and how it may require an update of regulatory rules and standards.

Big Data and AI Governance

The OECD has recently released its AI Principles;[2] the European Union has enacted the General Data Protection Regulations related to privacy and has several other initiatives in motion; the United Nations, through its High-level Panel on Digital Cooperation, has just released its report calling for the UN Secretary-General to facilitate an agile and open consultation process to develop updated mechanisms for global digital cooperation (UN 2019). Canada, along with eight other countries, participates in the Digital 9 group, in which participants share world-class digital practices, collaborate to solve common problems, identify improvements to digital services, and support and champion growing digital economies.[3]

Policy

The UK government has outlined many initiatives, including its *Online Harms Paper* (HM Government 2019), as well as Unlocking digital competition: Report of the Digital Competition Expert Panel (HM Treasury 2019). Canada has released its national intellectual property strategy[4] and cyber strategy (Canada 2018), and recently announced its Digital Charter.[5]

Democracy

Many significant efforts exist in this area, including reports from the European Commission High Level Expert Group on Fake News and Online Disinformation (2018), the UK Parliament's Digital, Culture, Media and Sport Committee on disinformation and fake news (UK House of Commons 2019), the Knight Commission on Trust, Media and Democracy (2019), the LSE Truth, Trust & Technology Commission (London School of Economics and Political Science 2019) and the French Government's *Créer un cadre français de responsabilisation des réseaux sociaux: agir en France avec une ambition européenne* (Potier and Arbiteboul 2019).

The platforms have also announced initiatives: Google (2019) has recently released its proposal "Giving users more transparency, choice and control over how their data is used in digital advertising," Apple and Microsoft have called for stronger privacy laws, and Facebook is in the process of developing its content review board (Bloomberg 2018; Microsoft 2019; Harris 2019).

Why a New Institution?

To give this important governance framework initiative any hope of succeeding requires the formation of a new institution — the DSB. The current Bretton Woods institutions have their hands full and do not have the expertise in all of these areas. Allowing them to formulate the reforms would likely leave the process piecemeal — its current state. As with the undertaking of financial regulatory reform, undertaking reform in the digital sphere requires creating this new institution and would also both signal and acknowledge the importance of setting global standards and policies for big data, AI and the platforms. The stage is already set to move forward with an organization like the DSB. Recognizing the key role that data now plays in the global economy, as well as the importance of trust to underpin the uses of data, the current Japanese G20 presidency has "data free flows with trust" as one of its key themes.[6]

The International Grand Committee on Big Data, Privacy and Democracy — which comprises a diverse set of 11 countries and more than 400 million citizens — could serve as a natural springboard to launch the DSB. The Committee's focus has been on the behaviour of platforms, including their role in disseminating fake news. The diversity of the membership — small and large countries, with differing cultures, values and institutions — makes it ideal for launching the DSB. Funding would come from its member countries alongside voluntary donations and in-kind contributions via participation in the DSB working groups.

This year is the seventy-fifth anniversary of Bretton Woods. Announcing the formation of this institution would recognize the important role that the Bretton Woods institutions have played in promoting a rules-based system — which has led to vast improvements in living standards — while also recognizing the need to update these arrangements to reflect the

profound implications arising from digital platforms.

One of the new institution's first steps would be to document all of the current activities — looking for commonalities and key areas of divergence, gaps and institutions involved. Eventually, it would develop a universal declaration on AI ethics, patterned on the universal declaration on human rights. These will require substantive work and are worthy goals for a new institution to undertake.

It Is a Beginning, Not an End

The DSB is a starting point. It is likely that a diverse set of countries and stakeholders would want to take part, and in fact would be necessary for the organization to have legitimacy.

It is also likely that there will be resistance to new forms of regulation and ways of doing business in the technology platform space, just as the creation of the FSB drew opposition. But as with financial sector reform, these efforts are essential to achieve the full benefits from the platforms. They will build and solidify trust, and trust is ultimately what will attract users to a platform. In keeping with the open nature of the World Wide Web, the process for reform should be open to all countries and organizations that wish to join, either at the outset or as the reform process matures.

And the time to start is now.

NOTES

1 See www.fsb.org/work-of-the-fsb/#coordination.

2 See https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

3 See www.canada.ca/en/government/system/digital-government/improving-digital-services/digital-9.html#Charter.

4 See www.ic.gc.ca/eic/site/108.nsf/eng/h_00000.html.

5 See www.ic.gc.ca/eic/site/062.nsf/eng/h_00108.html.

6 See https://g20.org/en/overview.

WORKS CITED

Bloomberg. 2018. "Apple's Tim Cook Calls for More Regulations on Data Privacy." Bloomberg News, March 24. www.bloomberg.com/news/articles/2018-03-24/apple-s-tim-cook-calls-for-more-regulations-on-data-privacy.

Canada. 2018. *National Cyber Security Strategy: Canada's Vision for Security and Prosperity in the Digital Age*. PS4-239/2018E. www.publicsafety.gc.ca/cnt/rsrcs/pblctns/ntnl-cbr-scrt-strtg/index-en.aspx.

European Commission High Level Expert Group on Fake News and Online Disinformation. 2018. *Final report*. March 12. https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation.

Google. 2019. "Giving users more transparency, choice and control over how their data is used in digital advertising." Proposl for feedback and discussion, version 1.0, August. http://services.google.com/fh/files/misc/industry_request_for_comment_v1.0.pdf.

Harris, Brent. 2019. "Establishing Structure and Governance for an Independent Oversight Board." Facebook Newsroom, September 17. https://newsroom.fb.com/news/2019/09/oversight-board-structure/.

HM Government. 2019. *Online Harms White Paper*. Presented to Parliament by the Secretary of State for Digital, Culture, Media & Sport and the Secretary of State for the Home Department by Command of Her Majesty, April. www.gov.uk/government/consultations/online-harms-white-paper.

HM Treasury. 2019. *Unlocking digital competition: Report of the Digital Competition Expert Panel*. March. www.gov.uk/government/publications/unlocking-digital-competition-report-of-the-digital-competition-expert-panel.

Knight Commission on Trust, Media and Democracy. 2019. *Crisis in Democracy: Renewing Trust in America*. Washington, DC: Aspen Institute. http://csreports.aspeninstitute.org/documents/Knight2019.pdf.

London School of Economics and Political Science. 2019. *Tackling the Information Crisis: A Policy Framework for Media System Resilience*. London, UK: LSE Truth, Trust & Technology Commission. www.lse.ac.uk/media-and-communications/truth-trust-and-technology-commission/The-report.

Microsoft. 2019. "Microsoft calls for greater collaboration between tech companies and government to strengthen consumer trust in digital services." Microsoft News Center India, June 27. news.microsoft.com/en-in/microsoft-greater-collaboration-tech-companies-government-strengthen-consumer-trust-digital-services/.

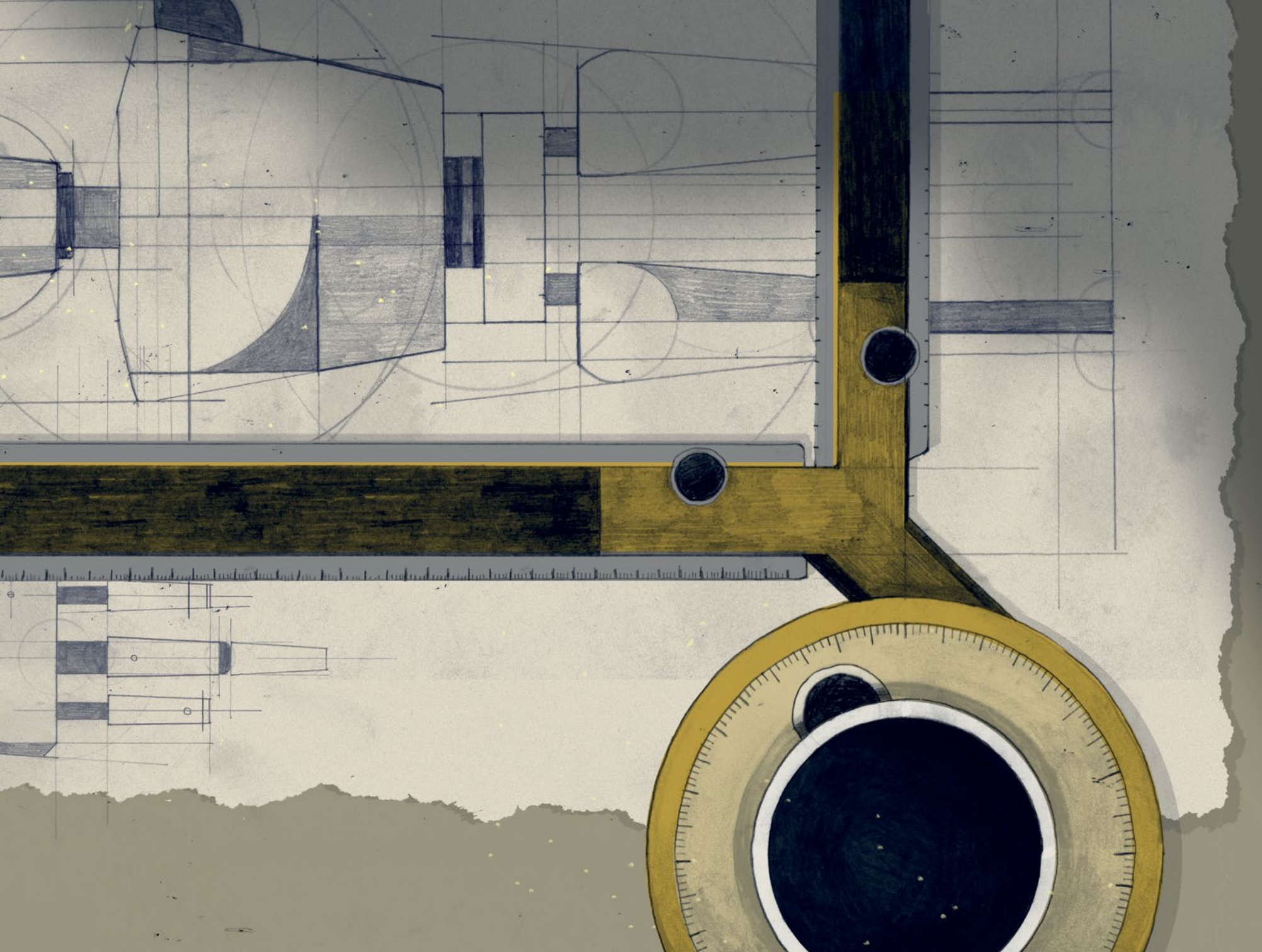
Potier, Frédéric and Serge Abiteboul. 2019. *Créer un cadre français de responsabilisation des réseaux sociaux: agir en France avec une ambition européenne*. May. France: Secrétariat d'Etat au numérique. www.ladocumentationfrancaise.fr/var/storage/rapports-publics/194000427.pdf.

UK House of Commons. 2019. *Disinformation and "Fake News": Final Report*. February 18. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf.

UN. 2019. *The Age of Digital Interdependence*. Report of the UN Secretary-General's High-level Panel on Digital Cooperation, June. https://digitalcooperation.org/report/.

ABOUT THE AUTHOR

Robert (Bob) Fay is director of global economy at CIGI and is responsible for research direction and related activities. He has extensive experience in macro- and micro-economic research and policy analysis. Prior to joining CIGI, Bob held several senior roles at the Bank of Canada (BoC), most recently as senior director overseeing work to assess developments and implications arising from the digitization of the Canadian economy. As deputy director of the International Department at the BoC, he assessed global economic developments and their implications for Canada and investigated a wide variety of issues, including those related to the international monetary system and global financial architecture.



Michel Girard

Global Standards for Digital Cooperation

Robust global standards and third-party certification programs are essential to anchor big tech platform governance. First, credible and enforceable global standards are needed for consumers and civil society to regain trust in big tech platforms. Second, they represent the only available pathway to avoid an unwieldy patchwork of national regulations. As well, global standards are required to create a level playing field, where smaller firms can compete against big tech platforms. Finally, without global standards, there cannot be an inclusive digital economy and society, as called for by the UN Secretary-General's High-level Panel on Digital Cooperation. The creation of a Digital Stability Board — as Robert Fay proposes in this series — can spur the development of credible global standards and third-party certification programs to properly frame big data analytics. Given the stakes, and the large number of standards and specification bodies involved, there is an urgent need for international cooperation and coordination in this space.

What Standards Are and Why They Matter

Although they're not visible to the average consumer, standards and conformity assessment activities keep the economy running. They cover everything from setting the size of the simplest screw thread to managing the most complex information technology network. Standards provide a level playing field for industry and help build trust between participants in supply chains. Standards serve as a "handshake" between various components of systems and allow for interoperability. Standards also play a pivotal role in protecting the health and safety of consumers in a range of sectors, including food and consumer products, infrastructure and the workplace.

Standards set out requirements, specifications, guidelines or characteristics that can be consistently applied to ensure that products, materials, processes and services perform as intended — qualitatively, safely and efficiently. They are drafted in a way that allows another party to test and certify that a product, process or system meets the requirements of a specific standard. Put simply, they make things work, help innovations spread and facilitate efficient trade among provinces, countries, economic regions and the international community of nations.

Standards are generally developed through a formalized rule-making process involving engineers and other technical experts, regulators and consumer interests. The process aims at balancing competing interests in order to offer a technical solution that is broadly accepted and shares the benefits of technological compatibility as widely as possible.

Many standards bodies were created at the beginning of the twentieth century to support industrialization. After World War II, new international organizations such as the International Organization for Standardization (ISO) were established as trade liberalization discussions gained traction. Today, thousands of standards development organizations (SDOs) are managing more than one million national standards and more than 330,000 international standards.

Principles for Standards Development and Maintenance

Standards to support the industrial economy were generally developed according to formalized rules stipulating the processes to be followed and involved engineers and other technical experts, regulators and consumer interest groups. International standards development bodies conform to the six principles of the World Trade Organization (WTO) for standards development and maintenance. These principles shed light on the philosophy behind technical standards development activities, and should be used by any organization entrusted with the development of global standards covering big data analytics, so as to bring credibility to the process and the outcomes.

Transparency

All essential information regarding current work programs, as well as about proposals for standards, guides and recommendations under consideration, and the final results of programs and recommendations, should be made easily accessible to at least all interested parties in the territories of at least all WTO members. Procedures should be established so that adequate time and opportunities are provided for written comments.

Openness

Membership in an international standards body should be made open on a non-discriminatory basis to, at the least, the relevant bodies of all WTO members. Participation at the policy development level and at every stage of standards development should also be open to members, without discrimination. Developing country members, in particular those with an interest in a specific standardization activity, should be provided with meaningful opportunities to participate at all stages of standard development.

Impartiality and Consensus

All relevant bodies of WTO members should be provided with meaningful opportunities to contribute to the elaboration of an international standard so that the standard

development process will not give privilege to, or favour the interests of, one or more particular supplier, country or region. Consensus procedures should be established that seek to take into account the views of all parties concerned, and to reconcile any conflicting arguments.

Effectiveness and Relevance

In order to prevent unnecessary trade barriers, international standards need to be relevant and to effectively respond to regulatory and market needs, as well as to scientific and technological developments in various countries. They should not distort the global market, have adverse effects on fair competition, or stifle innovation and technological development. In addition, they should not give preference to the characteristics or requirements of specific countries or regions when different needs or interests exist in other countries or regions. Whenever possible, international standards should be based on performance rather than design or descriptive characteristics.

Coherence

To avoid the development of conflicting international standards, international standardizing bodies must avoid duplicating, or overlapping with, the work of other international standardizing bodies. Cooperation and coordination with other relevant international bodies is essential.

Development Dimension

Constraints to effective participation in standards development, in particular the constraints on developing countries, should be taken into consideration in the standards development process. Tangible ways of facilitating developing countries' involvement in international standards development should be sought. The impartiality and openness of any international standardization process require that developing countries are not excluded de facto from the process.

Standardization in the ICT Sector

When it comes to the information and communications technology (ICT) sector, standard-setting activities can only be described as extraordinarily complex, opaque, evolutionary, bottom-up and unpredictable.

A number of factors led to the development of new models for setting standards and specifications (such as consortia and open-source software collaboratives), in parallel to traditional SDOs. In its infancy, the ICT sector followed the same path as other industries and relied on the traditional standards development model. However, with digitization in the 1970s, new approaches were needed to quickly set a bewildering number of new standards and specifications in order to achieve interoperability (Updegrove 2007). Starting in the 1980s, standards consortia organizations began to appear. Approximately 60 percent of all standards and specifications covering the ICT sector were created by consortia, including well-recognized interoperability standards such as USB drives, DVDs, the Blu-ray optical disc format, HTML, UHD, XML, MIDI and PCI Express (Biddle et al. 2012).

The entire edifice of digitization is based on software development and coding. As digitization emerged, so did new approaches to draft, test and ensure new ICT products' interoperability, from software to code language and apps. Although traditional SDOs are still used to generate rules for broad applications such as cyber security management systems or cloud computing, by and large, software developers shunned

Standards serve as a "handshake" between various components of systems and allow for interoperability.

traditional SDOs and standards/specifications consortia in favour of open-source software platforms. Microsoft, for example, which relied heavily on traditional SDOs to ensure interoperability, testing and certification of products such as cloud computing in the early 2000s, now uses development platforms such as GitHub to host and review code and build software with a community of 24 million developers.

However, unlike traditional SDOs, consortia and open-source development platforms are simply not designed to solicit broad public participation for making choices between various approaches or to integrate social, ethical, cultural or other considerations as a new product is being designed. When a project is assigned to an open-source software development platform, fundamental questions surrounding the "whether," the "what" and the "why" and the possible alternatives to an approach have already been answered. Instead of focusing on these fundamentals, participants are invited to work together to fix bugs and to help on the "how," including product design, outreach and marketing, to ensure new projects actually work as intended when launched. This process raises serious accountability and responsibility issues whenever software may have an impact on the health, safety and security of users.

The Case for Global Standards to Frame Big Data Analytics

There are five main arguments that call for voluntary, global standards to frame big data analytics, which includes big data platforms.

- Innovation is outpacing legal and regulatory frameworks and regulators' ability to respond to new issues.
- Governments are responding by developing approaches to frame new issues on their own, but fundamental principles are not harmonized around the world, leaving both regulators and big tech platforms unsure of how to enforce or comply. Inconsistencies in approaches are adding costs for framework implementation and contributing to lack of compliance because of conflicting requirements.
- Big data analytics is not the exclusive domain of big tech platforms, but becoming embedded in all industries, including traditional market players. While in the past each sector built a standardization framework in silos, market participants now employ legions of ICT software engineers and data scientists to work on big data analytics. Foundational documents can underpin new innovations in all market segments and allow for interoperability, not only with big tech platforms, but also between other players.
- The geopolitical dynamics of increased nationalism are weakening a number of international organizations aimed at supporting globalization through treaties and binding agreements. The international standards development community is one of the few stable institutions providing an international trust mechanism able to balance essential sovereignty concerns with global trade, because it is in the business of developing voluntary normative documents.
- If we do not pre-emptively establish normative standards to help society manage the risks accompanying big data, the consequences will almost certainly be unintended and unanticipated harm. The difference between these digital innovations and historical innovations in the tangible goods economy is that the unprecedented rate of progress and innovative possibilities today can outpace sober second thoughts.

The Need for Enhanced Coordination and Cooperation

The Undergrove survey (referenced above) identified more than 200 organizations involved in standardization in the digital industries sector, from long-established international bodies to technology-specific consortia to open-source specification platforms. Traditional standardization bodies such as the ISO, the International Electrotechnical Commission (IEC) and the Institute of Electrical and Electronics Engineers (IEEE) have been active in developing standards to frame big data analytics. The ISO and the IEC, for example,

created new committees and working groups under the Joint Technical Committee — JTC 1 — to develop foundational standards covering big data and artificial intelligence. The IEEE is spearheading a global initiative on the ethics of autonomous and intelligent systems, a new series of standards under its 7000 series and a new ethics certification program for autonomous and intelligent systems. In Canada, the Chief Information Officers Strategy Council has been accredited by the Standards Council of Canada to develop big data standards.

In addition, a large number of collaborative development platforms (open-source development and informal group projects) have become the preferred method for software-based interoperability development. However, while addressing certain industry needs, consortia and open-source platforms generally do not satisfy regulators' need to adhere to more formal international requirements regarding government use of global standards developed in the private sector.

Standardization work in this space would greatly benefit from enhanced cooperation and coordination. There is a need to establish a more robust dialogue between big tech software engineers, data scientists, regulators and civil society, and to agree on an international standards road map to properly frame international digital cooperation. However, no organization has been mandated by governments to coordinate global standards development activities and to ensure that all standards-setting bodies in the digital space adhere to the WTO principles, in order to ensure the credibility of both the process and the outcomes. A Digital Stability Board could be entrusted with these tasks.

Without global standards framing data value chains, international digital cooperation will remain a pipe dream. The world needs credible data governance standards covering issues such as privacy, the respect of fundamental rights, cyber security and data residency. We need big tech platforms to adopt those standards, and we need credible third-party certification programs to ensure that the standards are met.

Although it may sound counterintuitive, when it comes to global standards setting, leaders of big tech platforms at Google, Apple, Facebook, Amazon and the like are standing in the way of progress. Big tech leaders need to rethink their strategy regarding global voluntary standards setting. Stakeholders and shareholders don't want (and won't stand for) a patchwork of unenforceable, company-specific data governance policies, such as Google's proposed industry "standard" on data collection and digital advertising (Google 2019). Now is the time for big tech platform leaders to stop obfuscating and to join others in the global standards development sandbox. We need all players to work toward "one standard, one test" in order to reap the benefits of big data analytics.

WORKS CITED

Biddle, Brad, Frank X. Curci, Timothy F. Haslach, Gary E. Marchant, Andrew Askland and Lyn Gaudet. 2012. "The Expanding Role and Importance of Standards in the Information and Communications Technology Industry." *Jurimetrics* 52 (2): 177–208. www.jstor.org/stable/23239825?seq=1#page_scan_tab_contents.

Google. 2019. "Giving users more transparency, choice and control over how their data is used in digital advertising." Proposal for feedback and discussion, Version 1.0, August. https://services.google.com/fh/files/misc/industry_request_for_comment_v1.0.pdf.

Updegrove, Andrew. 2007. "ICT Standards Setting Today: A System under Stress." *First Monday* 12 (6).

ABOUT THE AUTHOR

Michel Girard is a senior fellow at CIGI, where he strives to drive a dialogue on standards for big data and artificial intelligence (AI), what they are and why they matter in these emerging sectors of the economy. His work highlights issues that should be examined in the design of new technical standards governing big data and AI in order to spur innovation while also respecting privacy, security and ethical considerations, and offers policy recommendations to facilitate the use of these standards and their incorporation into regulatory and procurement frameworks. In addition to his work at CIGI, Michel provides standardization advice to help innovative companies in their efforts to access international markets. He contributes to the CIO Strategy Council's standardization activities and was recently appointed to the International Electro Technical Commission's Market Strategy Board, which develops and maintains more than 10,000 international standards for electrical and electronic technologies.

Robert Gorwa

Regulating Them Softly

If you consume a decent amount of news coverage and popular commentary on technology policy issues, you doubtless frequently come across the statement that the companies that run popular platforms for user-generated content such as Facebook, Instagram, Twitter and YouTube are "unregulated."

Although this argument hints at the sentiment amassing an odd form of bipartisan support in the United States and a number of countries in Western Europe — that platform companies currently have insufficient responsibilities to their users and to the public, and their operation is creating a host of negative social and political externalities — it is off the mark in two respects. The first, frequently noted by frustrated academics in opinion pieces and Twitter threads, is that companies serving as intermediaries for user-generated content *are regulated*; it's just that these "intermediary liability" provisions, as enacted in legislation such as the European Union's E-Commerce Directive and the United States' Communications Decency Act, are intentionally laissez-faire, crafted carefully to protect free expression and allow for innovation.

The second, less discussed problem, is that regulation comes in many flavours. The regulator's tool box includes various forms of "soft law" alongside the more traditional "hard" legislation. In the past decade, a network of "informal governance" initiatives — backroom deals, informal codes of conduct and other voluntary mechanisms — have been a major channel through which certain stakeholders, such as security-focused actors within the European Union, shaped the policies of platforms long before the current "techlash."

To ignore these other forms of governance is to miss a hugely important dimension of today's global politics of online content.

To ignore these other forms of governance is to miss a hugely important dimension of today's global politics of online content: the contentious battles between firms, governments and civil society actors that have shaped the global terms of service affecting the billions of people using the American superplatforms each day.

Questions of Corporate Governance

The global governance of corporations has always been a fraught, difficult affair. In the search for profit, environmental standards and labour rights have been ignored, books have been cooked, tax authorities evaded, and worse.

In the past few decades, in the absence of a world government (or meaningful international coordination) for policing and punishing bad actors, a growing number of private organizations and initiatives have been created in an effort to shape corporate behaviour through voluntary standards and transnational rules. Some of the earliest instances of this trend involved codes of conduct, often initiated under the umbrella of large international organizations such as the World Health Organization, which notably struck a deal with Nestlé in 1984 after a multi-year consumer boycott and international activist campaign that followed the company's infant formula scandal (Sikkink 1986).

Since then, initiatives have been increasingly developed by groups of non-governmental and industry organizations, with dozens of efforts that have sought to create standards and outline best practices around sustainability (for example, ISO14001 and the Forest Stewardship Council), labour rights (such as the Fair Labor Association and the Worker Rights Consortium) and many other areas (Fransen and Kolk 2007).

This scaffolding of various voluntary arrangements, public-private partnerships, industry-specific measures and other informal regulatory instruments, often called "transnational governance," is today an essential feature of the global regulatory landscape for firms across a host of industries.

This landscape is imperfect, but it has become an important part of the battle for corporate accountability. The stakes are high: from finance and natural resource extraction to manufacturing, big corporations can have a significant social and political impact. As UN Special Rapporteur Philip Alston once asked, "Does Shell's sphere of influence in the Niger Delta not cover everything ranging from the right to health, through the right to free speech, to the rights to physical integrity and due process?" (quoted in Ruggie 2007, 826).

Today, companies such as Facebook, Google, Amazon and Apple have fashioned a global sphere of influence that often begs many of the same questions, and policy makers and some civil society actors have attempted to answer them in a manner similar to other industries: through informal transnational governance.

The EU Approach

In a recent article published in *Internet Policy Review*, I discussed the role of informal regulation for governing online content published on platforms in Europe.

As the technology lawyer Christopher Marsden (2011) has outlined, internet regulation in Europe has used "co-regulation" and other soft-law measures in the technology industry since at least 2005. Child safety was one early frontier. In 2008, the European Union's Social Networking Task Force convened multi-stakeholder meetings with regulators, academic experts, child safety organizations and a group of 17 social networks, including Facebook, MySpace, YouTube, Bebo and others. This process led to the creation of the "Safer Social Networking Principles for the EU," described as a "major policy effort by multiple actors across industry, child welfare, educators, and governments to minimize the risks associated with social networking for children" through more intuitive privacy settings, safety information and other design interventions (Livingstone, Ólafsson and Staksrud 2013, 317).

The European Union rolled out similar techniques to try to minimize the availability of terrorist content. In 2010, the Netherlands, the United Kingdom, Germany, Belgium and Spain sponsored a European Commission project called "Clean IT," which would develop "general principles and best practices" to combatting online terrorist content and "other illegal uses of the internet [...] through a bottom up process where the private sector will be in the lead" (quoted in Gorwa 2019). The Clean IT coalition, which featured significant representation from European law enforcement agencies, initially appeared to be considering some very hawkish proposals (such as requiring all platforms to enact a real-name policy, and to "allow only real pictures of users"), leading to push-back from civil society and the eventual end of the project. However, the project helped set the ideological foundations for the European Union's approach to online terrorist content by advocating for more aggressive terms of service and industry takedowns without formalized legislation.

In 2014, the European Commission laid out its plans for the "EU Internet Forum," which brought together EU governments with Facebook, Google, Microsoft, Twitter and the anonymous question-and-answer website Ask.FM to discuss how platforms should best combat illegal hate speech and terrorist content (Fiedler 2016). These meetings led to the 2016 EU Code of Conduct on online hate speech, signed by the aforementioned big four, and effectively resulted in platforms tweaking their terms of service globally to better reflect EU interests (Citron 2017).

Once civil society-led governance initiatives emerge, firms often create their own competing initiatives.

Civil society groups engage in transnational governance as well, issuing valuable guiding principles and declarations and founding multi-stakeholder transparency and accountability organizations. In 2008, the Global Network Initiative (GNI) was formed following pressure from human rights groups and policy makers over their conduct in China (Maclay 2010). The organization, which features Facebook, Google and Microsoft as members, along with a number of major civil society organizations, developed a set of high-level principles based on international human rights law that each member company says it will internalize; guidelines on how those principles should be implemented in practice, including commitments to engage in human rights assessments and transparency reporting; and an "accountability framework" that outlines the system of oversight, including company self-reporting, independent auditing and various "compliance" mechanisms. The public seems to know little about the organization, perhaps because the public output of the GNI is limited, and because it

Why not try and make platform governance more participatory and democratic rather than purely technocratic?

was formed in another era when governments were perceived as the bad actors of most pressing concern. Relatively little scholarly work has examined its effects and impact, but recent conversations around creating a "Social Media Council" for content moderation follow in the GNI's tradition of attempted civil society platform oversight.

Once civil society-led governance initiatives emerge, firms often create their own competing initiatives. Facebook looks like it will be the first to create a voluntary self-regulatory body for content policy, and recently published an eight-page "Charter" following a consultation period (Darmé and Miller 2019). The body, which will provide some oversight and input into Facebook's content policy process, has been described by legal theorists recently as a form of "structural constitutionalism" with which the company is becoming more governmental and developing a "judicial branch" of sorts (Kadri and Klonick 2019, 38). A more parsimonious explanation might be to conceptualize Facebook's efforts as another example of a private, informal governance arrangement, more akin to the dozens of certification, advisory and oversight bodies that have long been established in the natural resource extraction or manufacturing industries (Gorwa 2019).

Although the "resource" being governed in the case of Facebook's moderation oversight board is novel (user-generated concent, speech) relative to other industries, the process (a company's policies being scrutinized by a body) is not. Regulatory scholars argue that companies pursue these kinds of arrangements for a host of reasons, including to "improve their bargaining position with other actors, to win public relations points, and to evade more costly regulation" (Abbott and Snidal 2009, 71), all of which seem plausible in the Facebook case. Only time will tell how exactly the oversight board pans out, and if it ends up creating meaningful accountability, but one must be clear-eyed about the motives and interests behind such initiatives.

Making Platform Governance More Just

This hodgepodge of transnational governance initiatives, as well as the past efforts of the European Union and other players, have demonstrated what can happen when firms are brought to the bargaining table in a concerted political effort (often, by the threat of regulation and sanctions). For example, EU pressure led to the creation of the Global Internet Forum to Counter Terrorism (GIFCT), which helps coordinate industry takedowns of terrorist content. The GNI helped incentivize companies to publish transparency reporting, now an industry standard practice to at least a certain extent, and to set out best practices around government content removal requests.

However, these forms of soft regulation have also demonstrated a host of serious due process, accountability and freedom of expression concerns. The GIFCT is ultra-secretive, and publishes very little information about its "Shared Industry Hash Database" of "terrorist propaganda" (Llanso 2019). As European Digital Rights (2013) has documented, the EU Internet Forum and subsequent hate speech code of conduct marginalized civil society voices from the get-go and is problematic in a number of ways. Even the GNI is an opaque organization, with members bound by non-disclosure agreements; it reveals frustratingly little public information about the results of its company audits and other activities.

We need to do better. Civil society needs to be included, not marginalized; these organizations should not be secret, and should have as much transparency as possible; and governance efforts need to become far more representative of the hugely diverse users they purport to represent. Why not try and make platform governance more participatory and democratic rather than purely technocratic?

Of course, that is easier said than done: such an effort will require difficult cooperation and compromise, and must acknowledge from the onset that no group can just go it alone. As the political scientists Ken Abbott and Duncan Snidal have argued, in "transnational settings no actor group, even the advanced democratic state, possesses all the competencies needed for effective regulation" (Abbott and Snidal 2009, 68). Meaningful collaboration will be necessary.

Amid today's ambitious proposals for social media councils, platform regulators or oversight bodies — all of which may have significant effects on freedom of expression and other human rights — codes of conduct, self-regulatory bodies and other forms of "soft" regulation may appear to be an increasingly attractive proposition. But we need to be careful, studying the lessons from the past decade of platform regulation, as well as learning from comparable accountability mechanisms that have been introduced in other industries. As the politics of content moderation and intermediary liability become not only more visible but also more consequential, the status quo won't cut it for much longer.

WORKS CITED

Abbott, Kenneth W. and Duncan Snidal. 2009. "The Governance Triangle: Regulatory Standards Institutions and the Shadow of the State." In *The Politics of Global Regulation*, edited by Walter Mattli and Ngaire Woods, 44–88. Princeton, NJ: Princeton University Press.

Citron, Danielle Keats. 2017. "Extremist Speech, Compelled Conformity, and Censorship Creep." *Notre Dame Law Review* 93: 1035–71.

Darmé, Zoe Mentel and Matt Miller. 2019. *Global Feedback & Input on the Facebook Oversight Board for Content Decisions*. Menlo Park, CA: Facebook. https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-1.pdf.

European Digital Rights. 2013. "RIP CleanIT." European Digital Rights, January 29. https://edri.org/rip-cleanit/.

Fiedler, Kirsten. 2016. "EU Internet Forum against terrorist content and hate speech online: Document pool." European Digital Rights, March 10. https://edri.org/eu-internet-forum-document-pool/.

Fransen, Luc W. and Ans Kolk. 2007. "Global Rule-Setting for Business: A Critical Analysis of Multi-Stakeholder Standards." *Organization* 14 (5): 667–684.

Gorwa, Robert. 2019. "The Platform Governance Triangle: Conceptualising the Informal Regulation of Online Content." *Internet Policy Review* 8 (2). https://policyreview.info/articles/analysis/platform-governance-triangle-conceptualising-informal-regulation-online-content.

Kadri, Thomas and Kate Klonick. 2019. "Facebook v. Sullivan: Building Constitutional Law for Online Speech." St. John's Legal Studies Research Paper No. 19-0020. https://papers.ssrn.com/abstract=3332530.

Livingstone, Sonia, Kjartan Ólafsson and Elisabeth Staksrud. 2013. "Risky Social Networking Practices Among 'Underage' Users: Lessons for Evidence-Based Policy." *Journal of Computer-Mediated Communication* 18 (3): 303–20.

Llanso, Emma. 2019. "Platforms Want Centralized Censorship. That Should Scare You." *Wired*, September 26. www.wired.com/story/platforms-centralized-censorship/.

Maclay, Colin Miles. 2010. "Protecting Privacy and Expression Online: Can the Global Network Initiative Embrace the Character of the Net." In *Access Controlled: The Shaping of Power, Rights, and Rule in Cyberspace*, edited by Ronald J. Deibert, John Palfrey, Rafal Rohozinski and Jonathan Zittrain, 87–108. Cambridge, MA: MIT Press.

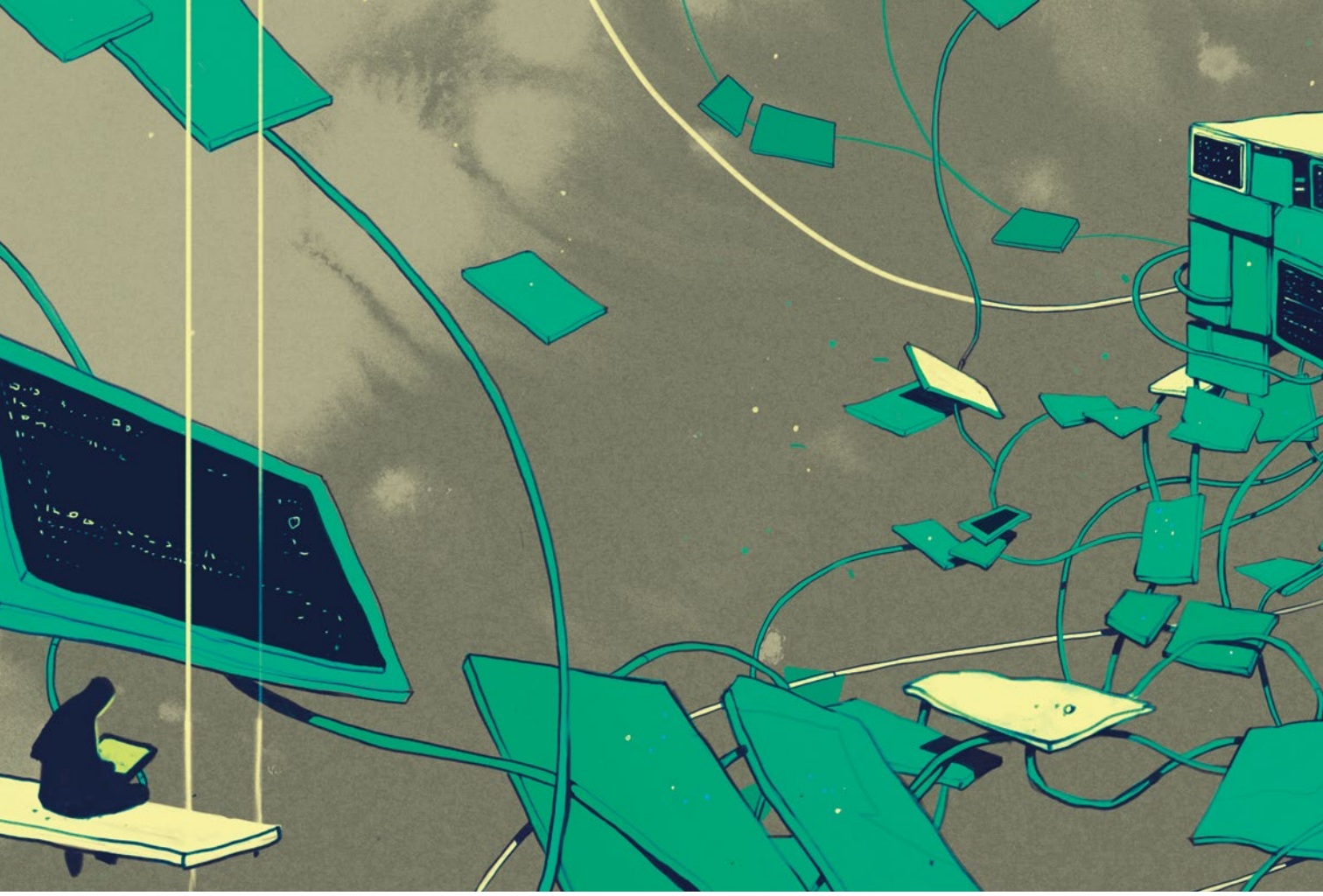
Marsden, Christopher T. 2011. *Internet Co-Regulation: European Law, Regulatory Governance and Legitimacy in Cyberspace*. Cambridge, UK: Cambridge University Press.

Ruggie, John Gerard. 2007. "Business and Human Rights: The Evolving International Agenda." *The American Journal of International Law* 101 (4): 819–84.

Sikkink, Kathryn. 1986. "Codes of Conduct for Transnational Corporations: The Case of the WHO/UNICEF Code." *International Organization* 40 (4): 815–40.

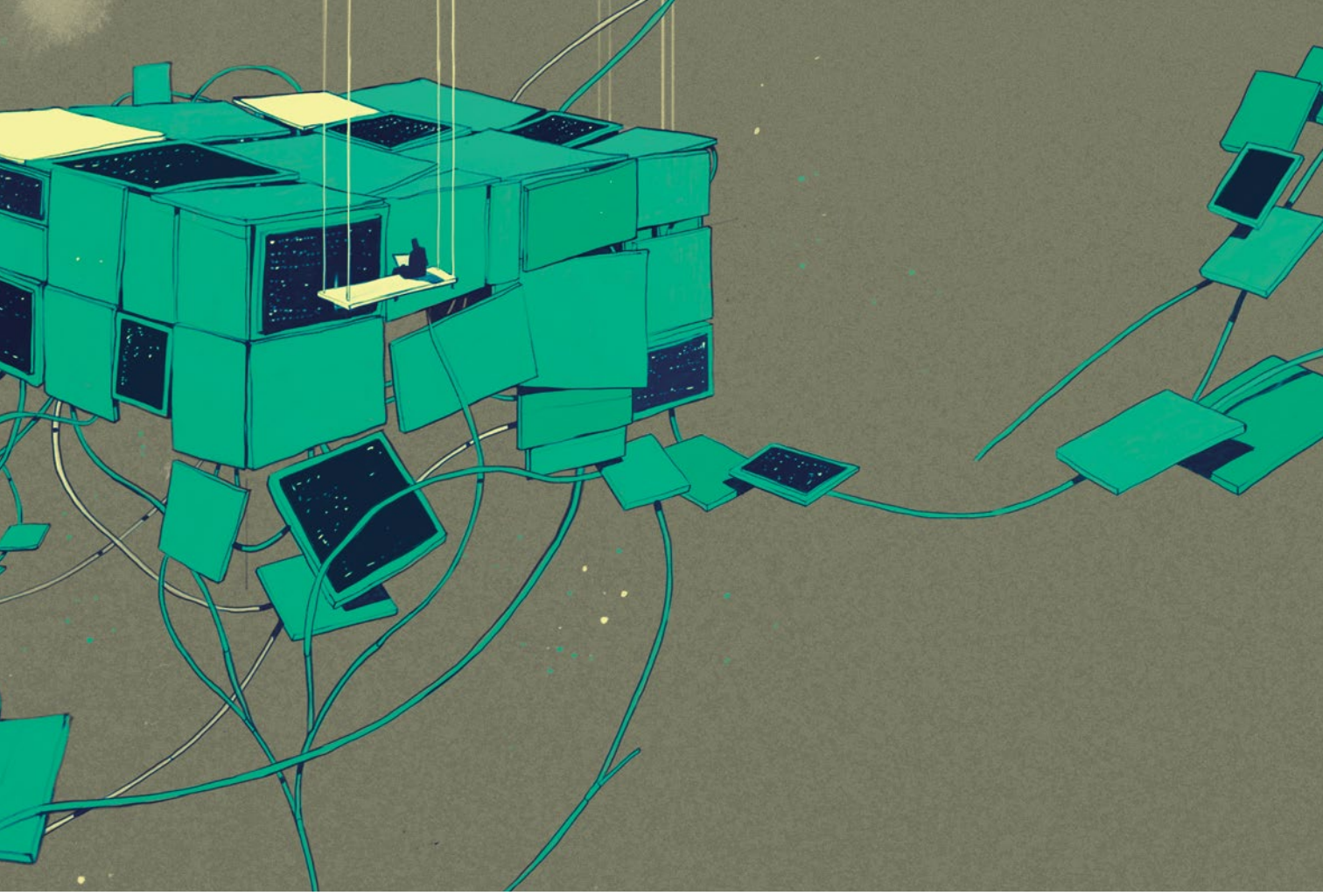
ABOUT THE AUTHOR

Robert Gorwa is a doctoral candidate in the Department of Politics and International Relations at the University of Oxford. Robert's work on platform governance has been recently published in *Information, Communication & Society*, *Internet Policy Review* and other journals. He contributes to the *Los Angeles Review of Books* and has written on tech and politics for *Wired UK*, *The Washington Post*, *Foreign Affairs* and other popular outlets.



Gene Kimmelman

Syncing Antitrust and Regulatory Policies to Boost Competition in the Digital Market



A handful of tech giants have enormous clout in key digital markets, and competition authorities around the globe are concerned about their domination of the market. In many countries, privacy and regulatory policy makers are similarly working to rein in tech power. As it seems clear that neither antitrust nor privacy rules alone are adequate to protect consumers and promote robust competition, countries must find ways to make all their policy interventions support complementary goals. The global nature of digital markets now makes such policy collaboration equally important across national boundaries.

The explosive growth of Google in the search engine and a suite of other popular service markets, of Facebook in social networking and of Amazon in online retailing illustrate the winner-take-all characteristics of many digital markets. Each of these companies has made extensive upfront investments to build platforms characterized by network externalities (for example, that consumers prefer to buy and suppliers to sell where everyone else is congregating), strong economies of scale and scope due to low marginal costs, and increasing profits based on control of data.

This combination of features means that these digital markets feature large barriers to entry. The leaders in search, social networking and other platforms have a large cost advantage with their scale of operations and a large leg-up from the scale of their data. Entrants cannot generally overcome these without either a similar customer base (network effects) or a similar scale (scale economies), both of which are difficult to obtain quickly and cost-effectively.

Consumers' tendencies to seek easy, simple, one-stop-shopping on digital platforms generate more barriers. As the final report of the Stigler Committee on Digital Platforms (2019, 29) noted, consumers tend not to "scroll down to see more search results, they agree to settings chosen by the service [provider], they single-home on [stick with] one platform, and they generally take actions that favor the status quo and make it difficult for an entrant to attract consumers."

In addition, the role of data in the digital sector fuels the advantages of companies such as Google, Facebook and Amazon. These companies' ability to collect massive amounts of personal data of all types allows for targeted advertising to consumers. It appears that the profits generated from gathering more data about and from more people grow larger as more dimensions of data are available to each platform, creating even more advantages for incumbent dominant service providers. Advantages are easily preserved by platforms that require consumers to agree to terms of service that are unclear, difficult to understand and constantly changing, which prevents consumers from understanding how their personal data is being monetized.

behind an independent start-up that would directly or indirectly seek to challenge a dominant platform. Venture capitalists will tend to put money behind companies that seek to be acquired by a dominant platform at an early stage, which reduces opportunities for disruptive investment and innovation.

As Europe presses the limits of antitrust enforcement and other nations slowly follow, we will soon see how much progress antitrust can make on its own to address and unwind examples of market dominance. However, regardless of the outcome of specific cases, antitrust law cannot upend the natural economics that drive digital markets toward a winner-take-all outcome. Eliminating anti-competitive behaviour that tilts the scales and market practices, even asset acquisitions that undercut competition, may not do enough to offset the benefits enjoyed by large networks with declining costs and massive data advantages over all other market players. Coordination among competition authorities should offer enormous opportunities to understand what works and what doesn't across jurisdictions. However, much more policy synchronization is likely necessary to control data-gathering practices and to create opportunities to grow competition.

Antitrust law cannot upend the natural economics that drive digital markets toward a winner-take-all outcome.

A common way for companies that obtain a dominant share of service on a platform to increase profits is to make all necessary complements to platform services themselves or to position themselves as necessary "bottlenecks" between partners and customers. By attempting to maintain complete control over the user relationship, dominant platforms can limit the possibility for independent complementary services to gain meaningful traction and challenge the platforms' power. Similarly, dominant platforms often use exclusive contracts, bundling or technical incompatibilities to restrict entry of competitors. When such practices succeed, investors become wary of putting money

Perhaps the most important change we need is to introduce competition-expanding regulations that address the problems antitrust cannot solve. A new expert regulator equipped with the tools to promote entry and expansion in digital markets could actually expand competition to benefit consumers, entrepreneurship and innovation. The regulatory authority could be housed within an existing agency or, better yet, be a new expert body, focused on digital markets.

The new regulator should also be responsible for consumer protection regulations relating to digital platforms, such as privacy protections for users. These rules may also have

pro-competitive benefits. For example, if the incredibly detailed data dossiers that the large platforms collect on their users are significantly curtailed by data protection legislation that limits collection and use of personal data, it may be easier for smaller or new companies that don't have access to data to compete. But these rules are also crucially important to protect users' rights and people's freedom from the type of control that detailed data collection gives companies.

The primary goal of the regulator, however, should be to actively *promote* competition, not simply to maintain existing competition. This is an important distinction: given the economic constraints described above, there is not enough competition now to be "maintained" — digital platforms need an extra jolt from a regulator to promote *new* competition. To achieve this goal, the regulator must be equipped with three key tools: interoperability, non-discrimination and merger review.

Interoperability

First, the agency should be authorized to require dominant platforms to be interoperable with other services, so that competitors can offer their customers access to the dominant network. For example, if Facebook, with its dominant position in social networking and ownership of Instagram and WhatsApp, were required to allow Snapchat users and similar alternative platforms to communicate with their Facebook friends easily using these other services, Facebook's network effect advantages would be reduced, and competition could more easily expand.

Of course, a rule requiring the transfer of user data depends on strong privacy protections, either as part of the rule or *guaranteed* by another statute, such as comprehensive privacy legislation. However, it's also important to ensure that privacy improvement efforts don't inadvertently make interoperability harder or impossible, for example, by banning any transfer of data from one company to another. The data protection and data empowerment tools that must be joined with interoperability should be the responsibility of the same regulator or carefully coordinated across two agencies.

Creating open interoperability regimes for the digital economy is a complex task that should be undertaken by an expert regulator, not generalist law enforcers. A regulator is especially useful for a tool like this because it will require technical detail, frequent updates and speedy dispute resolution to make sure the interoperability requirement actually promotes competition effectively. Antitrust enforcers, focused on competition, are not well positioned to effectuate user intent and protect users' personal data.

It's also important to ensure that privacy improvement efforts don't inadvertently make interoperability harder or impossible.

Non-discrimination

Competing against an incumbent digital platform can happen in two ways: head-to-head platform entry and expansion from one vertical to many. Recently, there have been few market entries into areas where one platform has gained enormous market share, such as with Google in search and Facebook in social networking. Therefore, it is important to assess other ways in which competition may grow.

Online platforms know that companies that use their platform can "disintermediate" them by connecting directly with the consumer, effectively cutting out the platform middleman. Online platforms know that a company that competes with them in one vertical can expand to compete in other verticals, becoming stronger as it takes advantage of synergies from the multiple verticals. This means that for platforms, the companies that use the platform are also potential competitors. Because of this competitive dynamic, some platforms have the incentive and ability to discriminate in ways that may harm competition. The platform has a variety of mechanisms it can

use to disadvantage companies that pose a competitive threat, including its access to transaction data, its prioritization of search results and its allocation of space on the page. In the most extreme versions of this behaviour, antitrust can prevent abuse, but it is less useful to prevent many subtle discriminatory practices.

The new regulator should monitor and ban discrimination by digital platforms with bottleneck power that favours their own services and disadvantages their competitors who rely on their platform to reach customers. Non-discrimination is another tool that particularly requires speedy adjudication and an expert regulator. It is difficult to identify which aspects of business are features of a platform, and which are products competing on the platform. For example, an app store may be an essential part of a smartphone operating system, so preferencing the operating system's own app store by having it preloaded on the phone may not be appropriately understood as "discrimination." In contrast, a grocery store is probably not an inseparable part of an e-commerce platform, so preferencing Whole Foods, an Amazon acquisition, over a competing grocery retailer on the Amazon Marketplace might be a good example of discrimination. The slow pace and complexity of antitrust litigation does not lend itself to fast-paced digital markets where discrimination can quickly make or break a competitive outcome.

Antitrust can prevent abuse, but it is less useful to prevent many subtle discriminatory practices.

Similarly, the agency should be authorized to ban certain "take it or leave it" contract terms that require companies doing business with a dominant digital platform to turn over customer data for the dominant platform to use however it pleases. Such terms effectively bundle the service the companies need with data sharing that could undermine their competitive market position. By prohibiting these practices, we can give potential competitors a fighting chance.

Merger Review

Another major concern with digital platforms is their acquisition of potential competitors. Acquisitions of potential or nascent competitors are often small, even falling below the value threshold for pre-merger notification of the competition authorities under the Hart-Scott-Rodino Act in the United States and similar thresholds in other countries. It is difficult to assess the likelihood that such companies in adjacent markets will be potential competitors. The small size or lack of pre-existing direct competition of these mergers can make it harder for antitrust enforcement agencies to block them, even if there are indications the merger may be anti-competitive. Markets move quickly, and a competitor's window of opportunity to gain traction against the incumbent is narrow, not only making mergers an even more effective tactic at preventing competition, but also making effective merger enforcement that much more important.

Thus, the regulator should also have the power to review and block mergers, concurrently with the existing antitrust agencies. For particularly important industries, such as communications, energy and national security, the United States has an additional merger review structure on top of antitrust. Similarly, digital platforms that have become essential in our economy and society, and that face inadequate competition, require merger review under a new and different standard, besides traditional antitrust review.

The new regulator would have a different standard than the antitrust agencies. This different standard should place a higher burden on dominant platforms to demonstrate their overall benefits to society, which antitrust enforcers do not have the tools to thoroughly measure. It should assess mergers involving platforms with bottleneck power, and it should only allow those mergers that actually *expand* competition. Also, there should be no size limit for mergers to warrant pre-merger review by the agency. Any acquisition by a platform with bottleneck power should be reviewed for its competitive impact. This would prevent increased concentration of power when the

This type of regulatory power, alongside strong antitrust enforcement, stands the best chance to expand opportunities for competition...

company being purchased is too small or the competitive consequences are too uncertain. Mergers that provide no clear competitive benefit would be blocked. The standard also must take account of the particular ways that competition happens in digital platforms. For example, non-horizontal mergers may be particularly harmful here due to the economies of scope in data-driven platforms, as well as the importance of interoperability between complementary products.

Jurisdiction

To which types of companies should these regulations apply? Some of the regulations, such as limits on data collection and use, are not related to levels of competition and therefore must apply broadly to be effective. Some others, like the requirement of non-discrimination, need only apply to especially powerful digital platforms that have the incentive and opportunity to disadvantage competitors. Identifying which platforms are powerful enough to be subject to those rules will require some additional work by the agency. Using the definitions of market power from the jurisprudence of antitrust is likely not sufficient. Instead, the regulator would need to make a determination of which companies hold important bottlenecks in the marketplace. This might be because they hold the buying power of so many customers that anyone who wants to sell must be on their platform to reach those customers. Or it might be because they have a monopoly on a key product that's complementary to many others, creating lock-in for a suite of related products as well. The Stigler Center report referenced above describes "bottleneck power" as a situation where buyers or sellers "primarily single-home and rely upon a single-service provider, which makes obtaining access to those [buyers or sellers] for the relevant activity…prohibitively costly" for other companies.

This type of regulatory power, alongside strong antitrust enforcement, stands the best chance to expand opportunities for competition on and across digital platforms, while also securing and limiting data gathering to protect consumers' privacy. However, restrictions such as non-discriminatory contracting or transparency requirements in one country may not succeed if those dependent on dominant platforms need similar treatment in other markets to make such protections profitable. Similarly, data protections that limit data portability and competition in one country may undermine another jurisdiction's effort to expand competition through broader access to data. It is therefore important for competition authorities to work with their own regulators and those across the globe to ensure that policy tools designed to promote competition and to protect consumers can truly achieve their goals.

WORK CITED

Stigler Committee on Digital Platforms. 2019. *Final Report*. Chicago, IL: Stigler Center. https://research.chicagobooth.edu/-/media/research/stigler/pdfs/digital-platforms--committee-report--stigler-center.pdf.

ABOUT THE AUTHOR

After serving as president and CEO of Public Knowledge for five years, Gene Kimmelman handed over that leadership role in 2019 and is now the senior adviser. Previously, Gene served as director of the Internet Freedom and Human Rights project at the New America Foundation, and as chief counsel for the US Department of Justice's Antitrust Division. Prior to joining the Department of Justice, Gene served as vice president for Federal and International Affairs at Consumers Union. Gene has also served as chief counsel and staff director for the Antitrust Subcommittee of the Senate Judiciary Committee and as legislative director for the Consumer Federation of America. Gene began his career as a consumer advocate and staff attorney for Public Citizen's Congress Watch.

Kate Klonick

Does Facebook's Oversight Board Finally Solve the Problem of Online Speech?

During the summer of 2019, the mass shooting in El Paso, Texas — the gunman apparently spurred on by online hate and troll community 8chan (Roose 2019) — reignited calls for online social media platforms to take down harmful content. In response to pressure, the intermediary site Cloudflare dropped its protection of 8chan (Newman 2019), only the second time in the service's questionable history that it has deplatformed one of its clients (the first being white supremacy site the Daily Stormer, following the fatal protests in Charlottesville in August 2017) (Klonick 2017). A few months later, in a dance that has become depressingly familiar over the last four years, conservative members of the US Congress demanded that platforms uphold free speech and not be allowed to wantonly "censor" certain types of content (Padhi 2018).

This tension is nothing new. The difficulty of preserving private companies such as Facebook, Twitter and YouTube as open platforms that support freedom of expression, while also meeting real-world concerns and removing harmful content, is as old as the internet itself. And while activists, scholars, government and civil society have called on platforms to be more accountable to their users for decades,[1] the feasibility of creating some kind of massive global stakeholder effort has proved an unwieldy and intractable problem. Until now.

On September 17, 2019, after nine months of consultation and planning, Facebook announced that it would be jettisoning some of its power to regulate content to an independent external body called the "Oversight Board" (Harris 2019). Described unofficially by Facebook CEO Mark Zuckerberg as a "Supreme Court" of Facebook in interviews in early 2018, the board is now imagined as a body to serve in an oversight capacity for Facebook's appeals process on removal of user content, as well as in a policy advisory role (Klein 2018). Having concluded their period of consultation and planning, Facebook says that the board will consist of 11–40 members, who will issue binding decisions on users' content removal appeals for the platform, and issue recommendations on content policy to Facebook (ibid.).

The creation of the board comes after more than a decade of agitation from civil society, users and the press to create a more accountable mechanism for platform content moderation. Platform policies around content moderation have high stakes because they involve the private governance of an essential human right — speech. Trying to build a more "publicly" accountable system means not only developing an appellate procedure through which users can seek re-review of removed content, but also finding a way for them to voice displeasure with existing policies and their enforcement.

But creating legitimacy around such a board — given that it originates with the entity it is supposed to be checking — is no easy task. It requires careful and deliberate thought around how to actually design and implement an independent adjudicatory board, how to ensure meaningful transparency, how to generate reasoned explanations for decisions, and how to perpetuate such a system into the future and guard against its capture by Facebook or other outside groups — or worse, impotence.

The board is not perfect, but it is a potentially scene-changing leverage point for accountability.

Essential to the board's gaining legitimacy as a public-private governance regime is others' perception of it as independent. Right, now the board's independence comes from two places. Financially, a third-party foundation will be given an irrevocable initial endowment by Facebook; that trust will operate independently to pay salaries and organizational costs for the board. Substantively, the board is in charge of its own "docket" or case selection (except in extreme circumstances where Facebook asks for expedited review on a piece of content), investigations and decisions. Moreover, a board member can't be removed for their vote on a particular content question, only for violating a code of conduct. These factors to preserve independence certainly aren't perfect, but they're more robust firewalls than Facebook has implemented in the past.

One of the other elements in creating legitimacy is developing a public process that incorporates listening to outside users and stakeholders into developing what the board would look like. In a series of six large-scale workshops around the world, dozens of smaller town halls, hundreds of meetings with stakeholders and experts, and more than 1,000 responses from an online request for public input, the Facebook team creating the infrastructure and design of the board listened to 2,000 external stakeholders and users in 88 languages about what they wanted from it, how it should look and how it should be built. Then the team published all that feedback in a 44-page report, with 180 pages of appendices, in late June 2019 (Facebook 2019a; 2019b). The report spelled out the panoply of questions inherent in making such an oversight body.

The release of the board's charter[2] reflects the work of this massive global consultancy that took place over six months. It also finally provided some answers and details on the trade-offs Facebook has decided to take in constructing the board.

Listening to outside voices and preserving the board's independence in both financial structure and subject matter jurisdiction help to ensure legitimacy, but perhaps one of the most vital pieces of establishing that the board would not be merely a toothless body was specifying how the board's decisions would interact with the decisions of Facebook.

This dynamic is one of the most significant points covered in the charter — how the board's decisions will be binding on Facebook — a major point of contention throughout the six-month consultancy with outsiders. Many in workshops and feedback fora thought the board's views ought to have a binding effect on Facebook's policy about what speech stayed up or came down on the platform. Others argued that the board having that much power was no better than Facebook having it — and that enacting those decisions online would be impractical to implement from an engineering perspective. Ultimately, the compromise made was that Facebook is to be bound by all decisions by the board on individual users' appeals. However, and perhaps most significantly, although the board can only *recommend* policy changes to Facebook, the company is *required* to give a public explanation as to why it has or has not decided to follow that recommendation.

Although Facebook is not under a mandate to take up the board's recommendations, neither can the platform just autocratically avoid transparency and accountability in deciding not to follow such policy recommendations from the board. Instead, it must furnish reasons — and, one assumes, good reasons — for deciding not to take up a recommendation of the board, and publish those reasons. The mandate of transparency creates an indirect level of accountability through public pressure

— or "exit, voice, and loyalty"[3] — traditional measures of popular opposition to power structures private or public. While not perfect, this arrangement allows the public much more access and influence over content moderation policy than users have ever had before.

The charter has answered a lot of questions, but also highlights some key ones remaining, among them who will serve. The subject was actively debated at workshops and drew feedback from outside stakeholders. Perhaps unsurprisingly, generally people felt that the board should be diverse, but in making specific recommendations, they seemed to believe that the board should reflect themselves. International human rights experts felt all board members should have a background in international human rights; lawyers felt that all board members should be legally trained or even all be judges. Ultimately, the charter envisions a board of 40 people at most who will meet certain professional qualifications — "knowledge," "independent judgment and decision making" and "skill[s] at making and explaining decisions" — while also bringing diverse perspectives.

Precisely who is on the first oversight board will no doubt be incredibly important: inaugural members will set the norms for what the board does and how it functions. But names alone will not secure the board's success or establish its legitimacy. While legitimacy can surely be helped along by a board whose members are judicious, even-handed, fair-minded and well-reasoned, those individuals must be guided by binding documents such as the charter and principles that structurally and procedurally incorporate transparency, independence and accountability. All these things together will determine the legitimacy of the board, and that process will not begin until users see Facebook make decisions recommended or mandated by the board that are in users' interests but against the company's immediate best interests. Those choices to act are up to Facebook, but the second part is up to all of us. This is where users expressing their exit, voice or loyalty responses to Facebook becomes more important than ever. The board is not perfect, but it is a potentially scene-changing leverage point for accountability. As we hurtle forward in this new land of private platform governance, we as users can't afford to let such opportunities for accountablility languish.

NOTES

1 As a few representative examples, see Electronic Frontier Foundation (www.eff.org), the Center for Democracy & Technology (https://cdt.org) and the work of Rebecca MacKinnon (2012).

2 See https://fbnewsroomus.files.wordpress.com/2019/09/oversight_board_charter.

3 A model of consumer responses created by Albert O. Hirschman (1970).

WORKS CITED

Facebook. 2019a. *Global Feedback & Input on the Facebook Oversight Board for Content Decisions*. Oversight Board Consultation Report, June 27. https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-2.pdf.

———. 2019b. *Global Feedback and Input on the Facebook Oversight Board for Content Decisions: Appendix*. Appendices A-F, June 27. https://fbnewsroomus.files.wordpress.com/2019/06/oversight-board-consultation-report-appendix.pdf.

Harris, Brent. 2019. "Establishing Structure and Governance for an Independent Oversight Board." Facebook news release, September 17. https://newsroom.fb.com/news/2019/09/oversight-board-structure/.

Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.

Klein, Ezra. 2018. "Mark Zuckerberg on Facebook's hardest year, and what comes next." *Vox*, April 2. www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge.

Klonick, Kate. 2017. "The Terrifying Power of Internet Censors." *The New York Times*, September 13. www.nytimes.com/2017/09/13/opinion/cloudflare-daily-stormer-charlottesville.html.

MacKinnon, Rebecca. 2012. *Consent of the Networked: The Worldwide Struggle for Internet Freedom*. New York, NY: Basic Books.

Newman, Lily Hay. 2019. "Cloudflare Diteches 8chan. What Happens Now?" *Wired*, August 5. www.wired.com/story/cloudflare-8chan-support-ddos/.

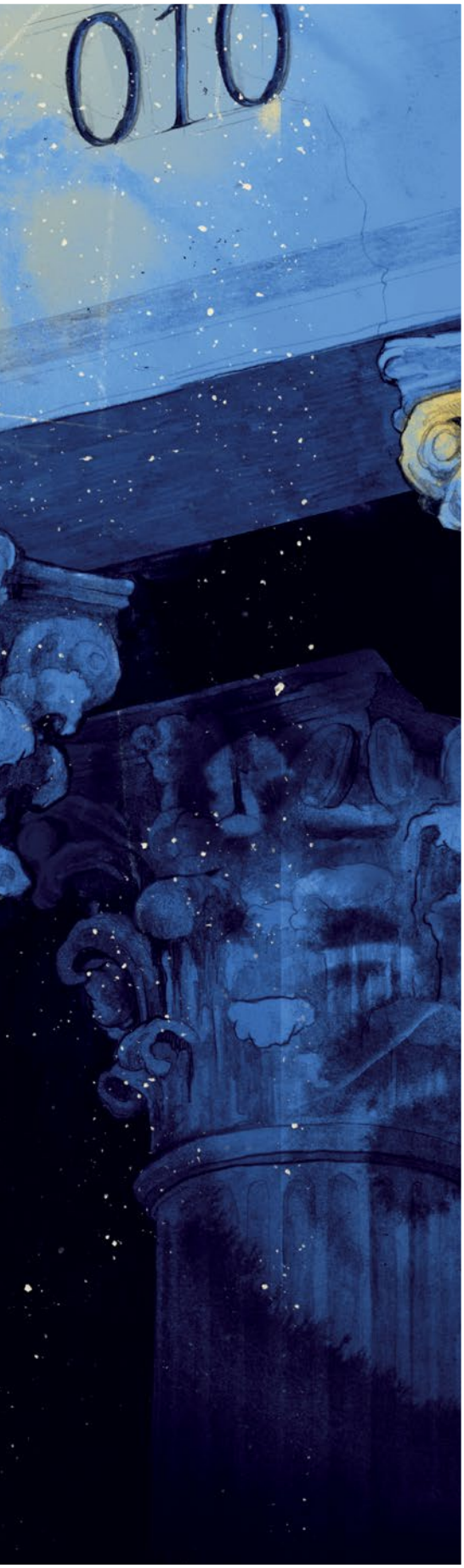
Padhi, Catherine. 2018. "Ted Cruz vs. Section 230: Misrepresenting the Communications Decency Act." *Lawfare* (blog), April 20. www.lawfareblog.com/ted-cruz-vs-section-230-misrepresenting-communications-decency-act.

Roose, Kevin. 2019. "'Shut the Site Down,' Says the Creator of 8chan, a Megaphone for Gunmen." *The New York Times*, August 4. www.nytimes.com/2019/08/04/technology/8chan-shooting-manifesto.html.

ABOUT THE AUTHOR

Kate Klonick is an assistant professor of law at St. John's University School of Law and an affiliate fellow at the Information Society Project at Yale Law School. She holds a J.D. from Georgetown University Law Center — where she was a senior editor at *The Georgetown Law Journal* and the founding editor of *The Georgetown Law Journal Online* — and a Ph.D. from Yale Law School, where she studied under Jack Balkin, Tom Tyler and Josh Knobe. Between law school and her time at Yale, Kate clerked for the Hon. Richard C. Wesley of the Second Circuit and the Hon. Eric N. Vitaliano of the Eastern District of New York.


10110 10010 010



Sean McDonald

The Fiduciary Supply Chain

Ironically, the recent, politically grandiose calls to break up big technology platform companies come just as the companies are already busily unbundling. Whether through independent advisory boards, outsourcing core trust and safety functions, or actual restructuring, technology companies are atomizing their core operations and offshoring liability. More important than how a platform company structures itself is how that structure defines user rights and corporate accountability. Relying on legislation or regulation to define platform governance standards through punishment avoids the problem: companies can just restructure. Recent headlines offer numerous examples of mergers, investments and bankruptcies that have been used to manipulate corporate liability (see, for example, Witt and Pasternack 2019). The defining policy question for the digital era isn't how we regulate company size, it's how we ensure that digital and platform governance standards persist across these companies' supply chains. How do we build data rights we can trust, and ensure that companies can't use corporate restructuring to avoid accountability?

While antitrust investigations and pressure gain momentum, there's already significant expert criticism of antitrust's ability to cope with fluid investment interests and complex data and digital sharing. The problem isn't just company size, it's that companies weren't designed to keep promises to the public, but to create, distribute and dispose of value and liability. And because that's their purpose, companies are exceptionally good at using incorporation and contracting to make meaningful accountability almost impossible.

One alternative to prevalent practice is the common law trust — a legal instrument that creates purpose-built governance over a set of assets and, importantly, creates fiduciary and contractual duties. *Digital trusts* — trusts that manage digital assets, such as data, code or the right to represent a data subject — may offer a more reliable way to ensure that platform companies are practically and legally accountable for their impact.

Instead of focusing our public investments in platform governance on breaking up the platform companies, we should focus on improving the ways that companies make important promises and increasing the prevalence of legal tools designed to uphold public interests, values and loyalties, such as trusts.

Unbundling Platform Governance

The term platform governance often hides a significant amount of complexity around fundamental questions — starting with what it means, but also including "by whom," "to what end" and "enforced how," among others. By framing policy discussions around platform governance, we risk focusing on an *instrumental* debate and missing the much larger — and more concerning — political economy of global regulatory enforcement, in addition to the specific complications posed by technology platform companies.

Platform governance can refer to a range of decisions and structures, but this analysis focuses on governmental and corporate governance of the companies that own technology platform companies, and the business decisions that ultimately determine user rights. It's important to be specific about the frame because no matter how commendable, ethical or well-designed a governance mechanism may be, the fundamental challenge remains that our institutions struggle to effectively regulate global companies. If any type of platform governance is to be effective, it will have to grapple with unbundling — substantively and structurally.

Globalization, Platform Regulation and Arbitrage

At the most basic level, technology markets are global and the laws that protect our fundamental rights are national. Both sovereign governments and multinational companies understand the impacts of that disconnect and are racing to leverage the resulting political economy to their own benefit. While that's predictable, it's also predictably raising questions about the effectiveness of strictly sovereign approaches to setting platform and data governance standards.

Historically, companies started local and grew, gradually, into international markets through careful negotiations with a range of related stakeholders at each step. Technology platform companies, however, launch globally overnight — leapfrogging market-access negotiations, often only responding to public and user concerns when they become scandals.

In addition to obvious scale issues, the shift in dynamic also decouples market access and government authority — whereas companies used to have to proactively earn government approval to reach their citizens, governments now have to proactively take steps to limit access to their markets, or find other ways to punish companies for abuse. That's a relatively weak stance for governmental regulators — and is the current footing for most data rights and governance protections.

The "global first" nature of technology platforms also means that companies no longer need to have offices everywhere they offer services, which enables them to focus on other drivers, such as access to skilled labour, favourable market conditions or suppliers. That flexibility also enables platform companies to manage their structure to minimize their regulatory burden, a practice called arbitrage. For nearly every type of government regulation, there are jurisdictions that market themselves to large companies through beneficial regulation.

The most obvious and highest-profile example of the tension between governmental authority and corporate arbitrage is taxation. For measure, a joint study between the University of Copenhagen and the International Monetary Fund found that 40 percent of all foreign direct investment is, in fact, multi-national companies avoiding taxes through shell corporations (Damgaard, Elkjaer and Johannesen 2019, 13). Even more damning, according to research from the Council on Foreign Relations, US companies report nearly seven times more profit from known international tax havens than from large, commercial markets (cited by Wolf 2019). Tax justice is a systemic issue, and a particularly acute problem among technology platforms, who are holding at least US$500 billion offshore to avoid US tax.

Ireland is one of the most important tax havens in the world; US-based multinationals represent 50 percent of the largest companies in Ireland, and 80 percent of its domestic corporate tax revenue (Quell 2019). Ireland markets itself to multinational technology companies based on favourable tax rates and government-sanctioned access to European markets. Ireland's US$14 billion in tax breaks

to Apple prompted a European Commission order demanding that Apple pay the bill, over Irish objection (Cao 2019).

The example of Ireland also demonstrates the shifting power of companies and countries in setting regulatory standards through arbitrage. If platform governance standards are to be effective, they'll need to grapple with the ways that sovereign competition may create a "race to the bottom" in corporate accountability.

What's missing is accountability to the people and groups that taxes and regulations are meant to protect — exactly the people whose trust these platforms need.

Antitrust, Unbundling and Supply Chain Governance

Most approaches to platform regulation focus on punishing companies for business practices that exploit users or cause large, negative social outcomes. The challenge with this focus is that the nature of company formation has changed (OpenCorporates 2018). Globalization and automation make it easier for companies to evolve from single entities into supply chains or "service-oriented incorporation" (McDonald 2019a). While holding companies and supply chains aren't new, technology enables platform companies to manage corporate structure with unprecedented speed, geographic spread and operational granularity. As a result, companies are unbundling into supply chains, both increasing the potential for arbitrage and limiting the effectiveness of regulation.

economy, the links are competing with each other. In order for platform governance to be effective, it will need to incorporate approaches to accountability that extend to entire supply chains. So, while antitrust dominates the political discourse because it cathartically promises to punish platform companies, building legitimate platform governance requires a constructive approach, ensuring that digital supply chains can credibly, accountably uphold standards and duties of care.

There's a significant amount of public and international pressure for the companies that own the world's largest platforms to unbundle: Alphabet, Facebook, Amazon and Apple are all under antitrust investigation in the United States, as well as by various authorities across Europe. Google, specifically, is under antitrust investigation by all 50 US attorneys general (McKinnon and Kendall 2019), after receiving US$9.4 billion in antitrust fines from the European Union (Lomas 2019). There are countless op-eds, political speeches and academic theories for how to break up big tech companies.

Yet, whether it's to minimize national and regional tax burdens, offshore liability for the risky new ventures or comply with data sovereignty and localization requirements, big tech companies are unbundling as fast as they can. As companies grow and mature, they often manipulate the way they're incorporated to minimize the cost and burden of regulatory compliance. The problem is, "regulatory compliance" is anodyne, executive-speak for

Our digital rights are defined by the standards upheld across entire supply chains.

Framing platform governance regulation around corporate accountability only enables authorities to focus on the behaviour of individual companies — individual links in the supply chain — whereas our digital rights are defined by the standards upheld across entire supply chains. As the saying goes, a chain's only as good as its weakest link — and in the digital

"avoiding public protections of workers, the environment and governmental authority." Mature, global companies use company structure differently than most people intuitively expect, using individual companies as shells that contain unbundled parts of their operation, to manage obligations to governments and the public.

Alphabet, the primarily Google-funded holding company run by Google's founders, is the highest-profile example of this approach. Google created Alphabet in 2015, amid investor and European antitrust pressures (Sharma 2018) and spun out a number of companies dedicated to individual lines of business, such as Nest (home thermostats), Google Capital (investment) and Sidewalk Labs (urban technology). While these units are technically separate, Alphabet has a well-documented history of merging elements of those separate companies in and out of Google, often in ways that fundamentally alter the company's previous statements about data privacy or use. Many of these companies share data, leverage the same advertising products and co-invest in joint ventures, intertwining them financially. Google unbundling into Alphabet has done something far more important than temporarily provide antitrust cover — it perfectly, publicly illustrated how corporate structures can be used to manipulate accountability to the public, customers and governments.

The weakest link in digital governance supply chains isn't any specific company, but the way we design the links themselves.

There are, of course, a range of trends and pressures compelling the transition from single companies to supply chains. At the operational level, most platform companies weren't designed for the kind of politically and socially complex governance that their businesses require. As a result, platform companies build, outsource and partner with a growing range of corporate structures, ostensibly independent oversight bodies and third-party vendors to rebuild public trust. Whether it's recognizing the social influence of the interdependent technologies involved in internet infrastructure, resolving disputes between users or convincing governments to allow them to operate in public spaces, platform companies are acting in recognition that their long-term sustainability is contingent on their ability to build functional supply chains — and that trust is a system requirement.

Supply chain governance has become an increasingly prevalent vector for improving the social impact of industrial practice, with several high-profile successes in improving labour conditions and environmental impact. And social activists are increasingly using supply chain advocacy to achieve social impact ends such as Google's banning of advertising for predatory loans (Sydell 2016) and Cloudflare's deplatforming of online hate sites the Daily Stormer and 8chan (Wong 2017; Elfrink 2019).

While these approaches have been novel and effective, they've also been anecdotal and opportunistic, rather than clear or systemic. The public reaction to platform governance has been to push for transparency, consistency and accountability. Structural approaches to building trust across variably aligned companies, linked by a supply chain of services and customers, aren't necessarily new, but they are fundamentally different to antitrust enforcement. And they start, not at the supply chain level, but where all chains fail: with the weakest link.

The weakest link in digital governance supply chains isn't any specific company, but the way we design the links themselves. In order for us to build trustworthy platform governance, we'll need to build corporate forms and contracting structures that are designed explicitly to accountably uphold common standards and duties.

Fiduciary Supply Chain Governance

Sometimes governance systems aren't defined by how well they achieve their core goals, but by how effectively they prevent bad actors from exploiting their core goals. One of the most important considerations for designing credible platform governance is creating operational, accessible and legally enforceable approaches to accountability and governance processes over time. Those questions are contributing to a growing field of digital political science (McDonald 2019b), which

focuses on designing governance standards and systems that blend public and private infrastructure to build equity.

In common law, the oldest and most established approach to creating shared standards of duty — especially during times of rapid legal transition — is fiduciary duties. Fiduciary duties are legally enforceable promises to act on someone else's behalf based on a type of law called equity, which relies on fairness to resolve disputes when there isn't binding or applicable law. Fiduciary duties typically include duties to loyalty (representing a person or group's best interests, to the best of a fiduciary's ability) and care (upholding an appropriate standard in that representation). There's already quite a bit of exploration of the idea of fiduciaries in digital and platform economies (for example, Balkin 2016), largely because they offer a key, unique benefit: they enable parties to negotiate broadly defined, legal accountability in ways that regulation and more traditional contracts can't.

The trajectory of fiduciary law scholarship has moved from its base in individual representation toward complex and multi-party governance, creating ample foundations to apply to digital economies. The recent focus on data trusts by the Canadian and UK governments, as well as by large corporate actors such as Alphabet, Microsoft and Mastercard, suggests that we've moved past asking "if" there is a role for digital and platform fiduciaries to the "how" of adapting those structures to platform governance.

Proposals about how to ensure that platform companies work toward, and take responsibility for, their social impact range from broadly defined duties of care, to setting industrial and engineering professional standards, to imposing fiduciary duties on platform companies, to using trusts to govern aspects of the data economy. The defining difference between these approaches isn't "what" they might accomplish — it's "who" has the power to mandate an approach, and "how" it might work in practice. Each differs in its approach to accountability, mitigating power asymmetry through sovereign regulation, industrial self-regulation, specialist professional services and collective self-governance, respectively.

Social activists are increasingly using supply chain advocacy to achieve social impact ends.

While there are a lot of small differences between the proposed approaches to platform governance, they break down according to "who" gets to hold platforms accountable. The current and traditional approach in many places is for government regulators to receive complaints, investigate abuse and then issue punitive or compensatory fines. Standards-based approaches rely on industry to self-regulate, with the potential for the public to hold individual companies accountable in civil court, for breach of industry standard. And, fiduciary and trust-based models create a dedicated steward, with a defined mandate, that can also be held liable in court, in cases where they ignore or underperform their duties. The key difference between these approaches is whether they focus on empowering public rights of action (government action) or private rights of action (public accountability).

The primary strength of public rights of action and statutory approaches to fiduciary duties is that governments typically have an established set of mature tools and infrastructure to regulate markets and companies. One central criticism of public rights of action is that they're inherently political, based on the influence and jurisdiction of individual governments over a company or industry. The most obvious example of this is the way that technology companies, and their stock prices, react to news about investigations by different authorities. The International Grand Committee on Big Data, Privacy and Democracy, whose most recent meetings in May 2019 were attended by representatives from 12 governments, has had three consecutive requests to testify ignored by Facebook CEO Mark Zuckerberg (Kates 2019). By contrast, technology stocks all lost significant (temporary) value when the US Department of Justice announced its antitrust investigations into the big five American

tech platforms (Savitz 2019). So, if a single government defines or imposes a model of fiduciary duties on tech platforms, they could become global standards, effectively imposing their cultural norms on global platform users. Beyond that, and more concerningly, statutory fiduciary duties could also restrict the accessibility of enforcement mechanisms, for people outside the imposing country. In other words, the process of deciding "who" gets to define standards around platform governance could prevent, or at least undermine, the effectiveness of that governance.

By contrast, centring fiduciary platform governance on individual rights of action relies on private law, focusing on enabling members of the public to resolve disputes with platforms without government intervention. The strongest argument against private rights of action are the large power asymmetries involved in access to justice, digital literacy and capacity. That being said, technology platforms have also been responsible for building governance mechanisms that scale quite effectively — eBay famously pioneered "online dispute resolution" to settle 60 million disputes per year (Rule 2008). That's not to suggest more technology is always the solution, but rather that there are lots of ways for technology platforms to build credible, scalable and trustworthy governance processes. More importantly, focusing on individual agency, which can be assigned or allocated to fiduciaries, advocates and other collective action models, centres the conversation around public equity, instead of focusing on political economy of large institutions. Unlike statutory approaches to fiduciary law, data trusts enable people to design and negotiate for their own priorities and values in the way that they're represented in digital systems.

The term "trust" doesn't inherently earn public trust.

Thankfully, statutory (public) and data trust (private) approaches to creating fiduciary duties aren't mutually exclusive — and are likely to complement each other as the field of practice develops professional standards. These trends point to the public's desire for smaller, more accountable technology platform companies, especially as they relate to user-facing data governance. They also point to the practical complexity of competing public and private authorities, designing digital rights across cultures and legal jurisdictions, and balancing competing, valid interests. Those challenges aren't specific to technology platforms, but they are significantly complicated by their global reach and domestically incorporated supply chains. Ultimately, these questions are not novel technology issues, but foundational political science questions. Given the prominence of digitization, their solutions are as likely to be engineered in state houses as in Silicon Valley.

Conclusion

Ultimately, platform governance is an almost infinitely complex challenge because of the scale of negotiation involved. The simple existence of the term "trust" doesn't inherently earn public trust, and the use of the legal instrument doesn't inherently ensure good governance. That said, trusts are a clear, established, tested legal vehicle for articulating, consolidating and stewarding the public interest — especially in contexts without well-established laws or rights.

The opportunity that data trusts offer to platform governance is a credible legal container to use as we start experimenting with new approaches, without risking that a failed experiment will make it easier to exploit the underlying data or its subjects.

Whether platform companies continue to unbundle to avoid liability, or because governments figure out how to make them, their component pieces will need what their aggregate lacked: a clearly articulated and operationalized duty to protect the public. It's possible that some sovereign, or group of sovereigns, will be able to compel that articulation and operationalization — but it's far more likely that those involved will figure it out first, in practice. Rather than try to drive deterministic approaches to platform governance, which are framed by a government's legitimacy and leverage, policy authorities should focus on building an enabling environment for principled, accountable experimentation around data governance that clearly articulates standards for accountability and redress.

Policy makers looking for practical approaches to advancing platform governance should prioritize de-risking the enabling environment for data trusts, including by harmonizing international fiduciary laws, in parallel to their investments in antitrust. While governments and companies continue to wrestle over who has the authority to take companies apart, data trusts are a critical element of laying the foundation for the future — not because they inherently solve our trust problems but because, unlike most other legal tools, they clearly establish duties and accountabilities, often to specific groups and social causes. No matter how we decide to fix the platform economy we have, we'll need to build the foundations of the future with different legal tools than the ones that got us here.

Data trusts are a new version of an old tool, and one that provides continuity during big transitions. Importantly, they also do the one thing that our current institutions do not: clearly and directly create actionable accountability. Rather than search for a perfect, silver bullet tool — the authorities pushing for platform governance should start with the powerful legal tools that we have and focus on ways to maximize their utility. No matter how we approach fixing platform governance, antitrust can only be part of the solution. In order to build the future of platform governance, policy makers will also need to maximize the value of the legal and governance tools we have to build trust. Data trusts are one place to start.

WORKS CITED

Balkin, Jack M. 2016. "Information Fiduciaries and the First Amendment." Faculty Scholarship Series 5134. https://digitalcommons.law.yale.edu/fss_papers/5154.

Cao, Sissi. 2019. "Apple Refuses to Pay Ireland $14 Billion in Back Taxes — And the Irish Don't Want It." *Observer*, September 17. https://observer.com/2019/09/apple-ireland-tax-lawsuit-european-union-corporate-tax-dodging/.

Damgaard, Jannick, Thomas Elkjaer and Niels Johannesen. 2019. "The Rise of Phantom Investments." *Finance & Development* 56 (3): 11–13.

Elfrink, Tim. 2019. "'A cesspool of hate': U.S. web firm drops 8chan after El Paso shooting." *The Washington Post*, August 5. www.washingtonpost.com/nation/2019/08/05/chan-dropped-cloudflare-el-paso-shooting-manifesto/.

Kates, Graham. 2019. "Facebook's Mark Zuckerberg declines latest invite to appear before international lawmakers." CBS News, September 9. www.cbsnews.com/news/facebooks-mark-zuckerberg-declines-latest-invite-to-questioning-by-international-lawmakers/.

Lomas, Natasha. 2019. "Google fined €1.49BN in Europe for antitrust violations in search ad brokering." *Techcrunch*, March 20. https://techcrunch.com/2019/03/20/google-fined-1-49bn-in-europe-for-antitrust-violations-in-search-ad-brokering/.

McDonald, Sean. 2019a. "How Regulations Are Reshaping Digital Companies." Cigionline, April 15. www.cigionline.org/articles/how-regulations-are-reshaping-digital-companies.

———. 2019b. "What Is Stalling Better Data Governance?" Cigionline, June 7. www.cigionline.org/articles/what-stalling-better-data-governance.

McKinnon, John D. and Brent Kendall. 2019. "States to Move Forward With Antitrust Probe of Big Tech Firms." *The Wall Street Journal*, August 19. www.wsj.com/articles/attorneys-general-to-move-forward-with-antitrust-probe-of-big-tech-11566247753.

OpenCorporates. 2018. "Fireflies and algorithms — the coming explosion of companies." https://medium.com/@opencorporates/fireflies-and-algorithms-the-coming-explosion-of-companies-9d53cdb8738f.

Quell, Molly. 2019. "Apple and Ireland Fight Against EU War on Corporate Tax Deals." Courthouse News Service, September 17. www.courthousenews.com/apple-and-ireland-lead-charge-in-eu-war-on-corporate-tax-deals/.

Rule, Colin. 2008. "Resolving Disputes in the World's Largest Marketplace." *ACResolution: The Quarterly Magazine of the Association for Conflict Resolution*, Fall. http://colinrule.com/writing/acr2008.pdf.

Savitz, Eric J. 2019. "Facebook and Other Big Tech Stocks Are Barely Moving on the DoJ's New Probe. Here's Why." *Barron's*, July 24. www.barrons.com/articles/doj-investigation-tech-stocks-51563989279.

Sharma, Rakesh. 2018. "Why Google Became Alphabet." Investopedia, January 2. www.investopedia.com/articles/investing/081115/why-google-became-alphabet.asp.

Sydell, Laura. 2016. "Google To Ban Payday Loan Ads." National Public Radio, May 11. www.npr.org/2016/05/11/477693475/google-to-ban-payday-loan-ads.

Witt, Jesse and Alex Pasternack. 2019. "The strange afterlife of Cambridge Analytica and the mysterious fate of its data." *Fast Company*, July 26. www.fastcompany.com/90381366/the-mysterious-afterlife-of-cambridge-analytica-and-its-trove-of-data.

Wolf, Martin. 2019. "Martin Wolf: why rigged capitalism is damaging liberal democracy." *Financial Times*, September 18. www.ft.com/content/5a8ab27e-d470-11e9-8367-807ebd53ab77.

Wong, Julia Carrie. 2017. "The far right is losing its ability to speak freely online. Should the left defend it?" *The Guardian*, August 28. www.theguardian.com/technology/2017/aug/28/daily-stormer-alt-right-cloudflare-breitbart.

ABOUT THE AUTHOR

Sean Martin McDonald is a CIGI senior fellow and the co-founder of Digital Public, which builds legal trusts to protect and govern digital assets. He is a lawyer and the CEO of FrontlineSMS, an award-winning global technology social enterprise, a fellow at the Duke Center on Law & Technology, a visiting fellow at Stanford's Digital Civil Society Lab and a former affiliate at Harvard's Berkman Klein Center. Sean is an adviser to Digital Democracy and the IEEE's Ethics and AI Committee and a researcher and writer whose work has been published by the *International Review of the Red Cross*, *Foreign Policy*, *Stanford Social Innovation Review*, Cornell's Legal Informatics Institute, IRIN and *Innovations*, to name a few. He holds a J.D./M.A. from American University, with specialization in international law and alternative dispute resolution, and is a member of the New York State Bar Association. Sean's research focuses on civic data trusts as vehicles that embed public interest governance into digital relationships and markets.

Nanjala Nyabola

Platform Governance of Political Speech

In May 2017, a few months before the Kenyan general election, the *Star* newspaper in Nairobi reported that British data analytics and public relations firm Cambridge Analytica was working with the ruling Jubilee coalition (Keter 2017). Then, days before Kenyans went to the polls, global privacy protection charity Privacy International told the BBC that the company had been paid $6 million in the contract (Bright 2017). Stories had long been simmering about Cambridge Analytica's association with illicit data harvesting, and it later emerged that the firm had been involved in the US presidential and Brexit campaigns. Why would a company associated with a US presidential election or a vote on the future of the European Union be involved in a general election in an African country?[1]

In fact, this wasn't the first time the public relations firm had been working in the area of African politics. They had been active in South Africa and Nigeria, and according to the firm's website at the time, they had worked in Kenya during the previous election runup, in 2013, building a profile of the country's voters through a survey of more than 47,000 Kenyans and developing a campaign strategy "based on the electorate's needs (jobs) and fears (tribal violence)" (cited in Nyabola 2018, 160). Cambridge Analytica's research identified the issues that resonated most with each constituent group, measuring their levels of trust, how they voted and how they digested political information; on the basis of this data, Cambridge Analytica recommended that the campaign focus on the youth as a segment that "could be highly influential if mobilised" (ibid., 161). Accordingly, they concentrated on developing an online social media campaign for the Jubilee presidential campaign to "generate a hugely active following" (ibid.).

Fears grew that these platforms could be used to mobilize ethnic hatred and even worse violence.

In Kenya in 2013, politics was paid greater attention on the internet than it had been given in the past. Monitoring political conversation online amid broader efforts to curb hate speech demanded significant organization. In the aftermath of the post-election violence in 2007-2008, social media had been named in passing as playing a role; by 2013, as the power of these spaces had become more recognized by the public, fears grew that these platforms could be used to mobilize ethnic hatred and even worse violence. The non-profit technology company Ushahidi originated around the 2008 violence to "map reports of violence in Kenya after the post-election violence in 2008"[2] and would eventually become one of Kenya's digital success stories. In 2013, Ushahidi expanded its mandate to study patterns of hate speech online, launching an initiative called "Umati" — Swahili for crowd — to monitor social media channels for any kind of hate speech and alert authorities before it escalated. These developments, combined with the shift toward digital reporting of election results, suggest the scale of changes that made 2013 Kenya's first digital election (Hopkins 2013).

Kenya's experience with Cambridge Analytica and other data analytics firms raises the question of political speech on internet platforms, and specifically, what platform governance in relation to political speech could and should look like. For Kenya, the challenge was complex. First, the country had a history of violent political processes, and there was every reason to fear that the election in 2017 would also be violent if poorly managed. Second, the firms implicated in the analytics operations were all foreign-owned, which raises the spectre of foreign manipulation of democratic processes. Third, there is the question of capital and money, given that none of the platforms that were connected to this experience were Kenyan, nor were the implicated analytics firms. Should foreign companies be able to meddle in political processes where they have no skin in the game, and in particular, when they are doing so for profit?

And, in fact, Kenya did live with the consequences of the manipulation of political speech on the internet in the 2007, 2013 and 2017 elections. In 2007, social media was identified as one of the platforms — radio and text messages were also identified — on which were distributed hate speech and incitement to violence, fuelling the deadly 2007-2008 post-election violence that resulted in at least 1,500 deaths and much more displacement. In 2013, manipulation of political speech online was part of a broader campaign to shift public opinion away from demands for accountability for that violence and toward portraying the International Criminal Court as

a tool of Western interference. Finally, in 2017, the distribution of hate speech on political platforms not only skewed political behaviour, but also contributed to one of the most virulent election campaigns in the country's history.

Nor has Kenya been the most grievous example of this phenomenon. A disproportionate amount of time has been spent debating the outcomes of the Brexit vote and the 2016 US presidential elections, because these countries exert a large influence on the global political economy. Yet, the impact of these issues is hardest felt in poor countries, where authoritarianism is a constant threat and regulations on speech are somewhat more fluid. These factors may perhaps explain why even Facebook itself acknowledges that its entry into Myanmar was connected to hate speech that fuelled the genocide against the Rohingya (Warofka 2018), or why the platform is increasingly the preferred space for disseminating hate speech in local languages in countries such as South Sudan (Reeves 2017).

Evidently, manipulation of political messaging online will be one of the major platform governance challenges of the coming years, given the frequency at which this style of manipulation is happening around the world. In the analogue world, political speech has always been recognized as a special category of speech governed by a subset of general norms and regulations on media and speech. Most countries have tight restrictions on what political candidates can and cannot say in traditional media, and publications are required to distinguish between political advertorials and organic journalism. These lines are starkest in countries with a long journalism tradition where the absence of these demarcations has led to serious social and political issues. Germany, for example, has severe restrictions on how political speech can be represented in the media, precisely because the media was such a focal point for Nazi mobilization in the lead-up to World War II. Yet, only after the Cambridge Analytica scandal are we seeing more attention paid to the developing of similar regulations online.

At heart, the methods used to create and disseminate hate speech online are strikingly similar to those used offline — playing on people's vaguest yet most potent fears; elevating the spectre of "The Other" as an existential threat to the dominant way of life or world view. These are all tricks that are as old as political speech itself. What is different is that internet platforms allow such political speech to be highly fragmented and targeted. The same ability to buy advertisements that allow you to sell jeans only to 18-to-30-year-olds living in Nairobi who are online between 7:00 p.m. and 9:00 p.m., as opposed to sending blanket advertisements to an unsegmented audience, is now being used for political speech. What this means is that the individual voters within a single state are not consuming the same political information; it is difficult to consider a public sphere truly representative when public discourse is reduced to highly inward-looking, fragmented groups talking past each other, rather than a true rational-critical exchange involving debate and deliberation.

The West is experiencing a fraction of what this looks like in the rest of the world.

At the same time, shifting so much of our political discourse online makes it hard to regulate. The West is experiencing a fraction of what this looks like in the rest of the world, where pockets of people are becoming increasingly radicalized on highly targeted or specialized websites. But in other parts of the world, the issue is compounded by the language problem. For example, in Ethiopia, even though the official language is Amharic, much of the content online is generated in one of the other hundreds of languages other than Amharic, because the internet makes it possible for people to do that in a way that a national newspaper or television station cannot. This is a double-edged sword — on the one hand, the internet has been great for creating space for the preservation of smaller languages, but on the other hand, this very

characteristic makes it difficult to effectively moderate what people are saying online, and some of it is really nasty (Fick and Dave 2019).

Underpinning all this is the question of accountability — who should be held responsible when all of this falls apart? The case of Myanmar has starkly demonstrated that just because a platform is ready to penetrate a new market doesn't mean it should. In 2018, the United Nations identified Facebook as a vector for hate speech calling for the genocide of the Rohingya.[3] That report was widely circulated, and Facebook made some strong statements acknowledging its recommendations and pledging to work toward them. And then — nothing. As it stands, there is practically nothing that can be done to make sure that Facebook is held accountable — or even to define what that accountability would look like — because the main interested parties in that genocide were the government itself. So, what does platform regulation look like when the platform that is upending the political space is run elsewhere, governed by laws outside your jurisdiction, and ultimately answerable to shareholders in a country other than your own?

At this stage, it should be very clear that platform governance with regards to political speech should be an urgent priority. But the imperatives must still be balanced out by the need to preserve free speech and the characteristics of a digital commons that make these platforms such powerful spaces for political action. Nor can the goals of such governance be transferred wholesale to states without consideration of the nature of the states in question. In many of the countries given as examples in this essay, the main perpetrator or the main interested party in misusing platforms to disseminate hate speech is the state itself.

One idea toward this end would be enhancing multilateral governance, headed up by an international norm-setting organization that identifies principles for political speech online and works with regulatory agencies in states, to both promote the norms and develop standards for implementation. There are already a wide number of internet governance dialogues out there — can they be better coordinated into a harmonized system that builds on what exists for the regulation of political speech in the analogue domain? Such coordination would also enhance the conversation between those working in political analysis, who recognize this problem as a growing problem, and those working in technology, who might not see how connected this issue is to what has gone before.

Self-regulation is also an option. The first stage would be transparency, namely, that political actors would be forced to disclose their identities when disseminating political speech online. But transparency may also involve public monitoring and observation. Crucially, content moderation needs to be systematized and made more transparent, and a platform should not be allowed to operate in a market if it is not prepared to take full responsibility for the political consequences it might trigger. This means, for example, that platforms should not be allowed to operate in a country until full content moderation in all the major languages of that country is operational. Yes, it's expensive and burdensome, but the alternative is Myanmar.

Ultimately, there is nothing truly new under the sun. Each iteration of media throws up the same uncertainties and challenges that have gone before and requires a renewed effort to address these complications for everyone's benefit. History records the panic that surrounded the development of the printing press — the idea that common people

It's a reminder that political speech is treated as a special class of speech on other platforms for a reason.

would be able to access potentially explosive information sent the church, in particular, into a tailspin. The popularization of tracts inflamed political behaviour across Europe and led to a variety of unexpected outcomes, including, arguably, what would become the Protestant Reformation. Eventually, regulations were developed to control what could be said in these tracts, media ethics became normalized, and laws on libel and defamation were much more strictly enforced.

None of this undercuts the fact that the internet has changed the way political information is generated and travels, but it does put things into perspective. It's a reminder that political speech is treated as a special class of speech on other platforms for a reason. It is a reminder that platform governance has been tricky before, but it was accomplished, even though it was not always timely. It is a reminder that while the challenge seems large and complex, it is not insurmountable.

NOTES

1 The author's description of this episode draws on text from her book *Digital Democracy, Analogue Politics: How the Internet Era Is Transforming Politics in Kenya* (Zed Books, 2018).

2 See www.ushahidi.com/about.

3 See generally, United Nations (2018).

WORKS CITED

Bright, Sam. 2017. "After Trump, 'big data' firm Cambridge Analytica is now working in Kenya." *BBC Trending* (blog), August 3. www.bbc.com/news/blogs-trending-40792078.

Fick, Maggie and Paresh Dave. 2019. "Facebook's Flood of Languages Leave it Struggling to Monitor Content." Reuters, April 23. www.reuters.com/article/us-facebook-languages-insight/facebooks-flood-of-languages-leave-it-struggling-to-monitor-content-idUSKCN1RZ0DW.

Hopkins, Curt. 2013. "How technology is shaping the decisive Kenyan elections." *The Daily Dot*, February 13. www.dailydot.com/layer8/kenyan-election-2013-technology-umati/.

Keter, Gideon. 2017. "Uhuru hires data firm behind Trump, Brexit victories." *Star*, May 10. www.the-star.co.ke/news/2017-05-09-uhuru-hires-data-firm-behind-trump-brexit-victories/.

Nyabola, Nanjala. 2018. "Politics, predators and profit: ethnicity, hate speech and digital colonialism." In *Digital Democracy, Analogue Politics: How the Internet Era Is Transforming Politics in Kenya*, 157–78. London, UK: Zed Books.

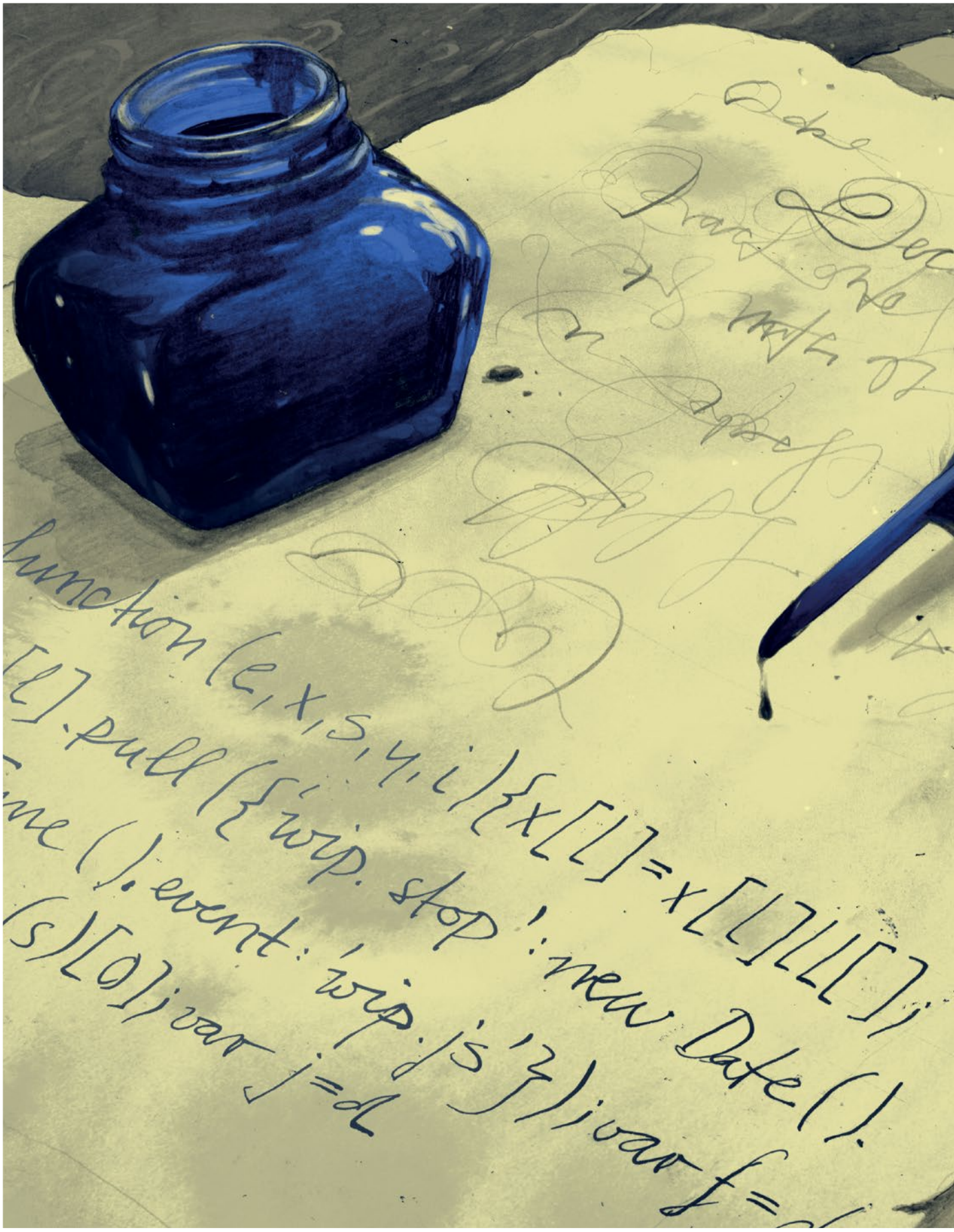
Reeves, Benjamin. 2017. "Online Fake News and Hate Speech are Fueling Tribal 'Genocide' in South Sudan." Public Radio International, April 25. www.pri.org/stories/2017-04-25/online-fake-news-and-hate-speech-are-fueling-tribal-genocide-south-sudan.

Warofka, Alex. 2018. "An Independent Assessment of the Human Rights Impact of Facebook in Myanmar." Facebook Newsroom, November 5. https://newsroom.fb.com/news/2018/11/myanmar-hria/.

United Nations. 2018. *Report of the Detailed Findings of the Independent Fact-Finding Mission on Myanmar*. A/HRC/39/CRP.2, September 17. Geneva, Switzerland: United Nations. https://digitallibrary.un.org/record/1643079.

ABOUT THE AUTHOR

H. Nanjala Nyabola is a writer, political analyst and activist based in Nairobi, Kenya. She writes extensively about African society and politics, technology, international law and feminism for academic and non-academic publications. Her first book, *Digital Democracy, Analogue Politics: How the Internet Era Is Transforming Politics in Kenya* (Zed Books, 2018), was described as "a must read for all researchers and journalists writing about Kenya today." Nanjala held a Rhodes Scholarship at the University of Oxford in 2009 and in 2017 was among the inaugural cohort of Foreign Policy Interrupted's fellows and a Logan Nonfiction Program Fellow at the Carey Institute for Global Good.


function(e,x,s,y,i){x[[]]=x[[]][[]];
[[]].pull({'wip.stop':new Date().
ne().event:'wip./s'});var f=
(s)[0];var j=d

Jonathon W. Penney

Protecting Information Consumers

Consumers today are confronted by a host of digital threats to their rights, safety and information environment: fake news; disinformation and the bots and automated networks amplifying it; online hate and harassment; mass data breaches; election interference by hostile foreign states; and algorithmic and big-data-driven targeting and manipulation, just to name a few. And the social media platforms where most users encounter these challenges — those with global reach such as Google, Facebook, Twitter, Instagram and YouTube — have utterly failed to take adequate steps to address to them.

A central part of the challenge is that most of these platforms are based in the United States and thus possess broad First Amendment protection — which limits content restrictions and other forms of speech regulation — while also enjoying blanket immunity from tort liability under Section 230 of the Communications Decency Act. This framework provides little incentive for them to act. With public opinion now turning, and the threat of tougher regulations, social media companies are finally beginning to act (Hirsh 2019) — but slowly, unevenly and still with a tendency to paralysis when competing claims arise. As recent examples, Facebook reduced the platform distribution of the widely shared fake video of House Speaker Nancy Pelosi but refused to remove it, citing the need to balance authenticity concerns with freedom of expression (Waterson 2019). And Pinterest aggressively pursued the takedown of harmful vaccine conspiracies (Sky News 2019), while ignoring other kinds of politically focused conspiracies and disinformation.

In short, the present law and policy framework governing these platforms is wholly inadequate. But what should replace it? This essay proposes a new comprehensive regulatory framework — one outlining information consumer protection — to hold these companies accountable and to better address the threats and challenges of the information and digital age.

The Evolution of Consumer Protection

Consumer protection is ancient. There are elements of it in Hammurabi's Code and evidence of consumer protective ordinances in ancient Greece and the Bible (Geis and Edelhertz 1973). To meet each era's different social, economic and technological challenges, consumer protection has evolved. The English common law's tort of deceit and doctrine of caveat emptor — let the buyer beware — suited consumers who mostly dealt with small merchants face to face (Pistis 2018). The consumer of the late eighteenth and nineteenth centuries, however, required greater protection from new manufacturing processes developed during the Industrial Revolution, such as food adulteration — the use of harmful preservatives in food — and from the lack of safety standards in increasingly large and impersonal industries. These changes led to new product liability laws. And Upton Sinclair's 1905 novel *The Jungle*, which chronicled the unsavoury conditions of Chicago's then meat-packing industry, famously helped foster the passage of the Pure Food and Drug Act in 1906.[1] Following World War II, the public's perception that industry had become too impersonal and powerful led to a strong mid-century consumer protection movement. This was exemplified in the United States by the expansion of federal consumer agencies and President John F. Kennedy declaring a Consumer Bill of Rights[2] in 1962, based on consumer rights to safety, to be informed and not deceived, to have choices among competitive options and to be heard and represented in policy making and related administrative processes. It would later be expanded in 1985 to include rights to basic consumer needs, to redress against businesses for wrongs, to consumer education and a healthy environment.[3]

In one sense, information consumer protection is simply a continuation of this evolution: a new consumer protection regulatory framework largely based on these same values — quality and safety, transparency, anti-deception, antitrust/consumer choice and accountability — but updated and refocused on today's data-driven information sphere. However, this framework requires an essential caveat: consumers of information, as well as the environment in which they are embedded, are fundamentally different from consumers of previous eras.

The Information Consumer

The digital age is driven by data and information, and so a renewed consumer protection movement should focus predominantly on information consumers — people who generate, share and consume news, information and content in the social media and digital spheres, not just for commerce, but also socially and democratically. The information consumer faces several unique threats and challenges.

First, the consumer is also the product. A traditional definition of a consumer is someone who engages in a transaction for a product or service. In the age of surveillance capitalism, people are both consumers and the consumed — information and data, predominantly *about consumers themselves*, is collected, analyzed and monetized to drive the digital economy (Srnicek 2017; Zuboff 2015). Consumers have always been the targets of deception and manipulation, but a combination of big data, powerful analytics driven by and targeting artificial intelligence (AI), platforms with entrenched market monopolies, and multiple public and private sector actors seeking to influence them, distort the information

Consumers of information, as well as the environment in which they are embedded, are fundamentally different from consumers of previous eras.

environment and create profound new possibilities for digital manipulation and "systematic consumer vulnerability" (Calo 2014; Zuboff 2019).

Second, the digital sphere in which the information consumer exists is not just a business or consumer environment, but a democratic one. Social media platforms are businesses with corporate aims, and people certainly use them to do business or to obtain goods or services. But that is not the primary reason people use them. Most do so for news and information and to connect with friends, families and others in their community. They are also now important sites for citizen engagement and democratic engagement. People obtain and share news and information on these platforms, and debate and deliberate on politics. Social media platforms are the new "quasi-public sphere" (York 2010) or, to use danah boyd's (2011) term, "networked publics," defined by a blurring of public and private. But they are also defined by a unique combination of digital consumerism and democracy — where the most important democratic spaces for the information consumer are owned, operated, shaped and controlled by private sector interests.

Third, the information consumer exists in an era of unparalleled distrust. Corporate and governmental failures to address many of the earlier noted threats — fake news, mass data breaches, online hate and abuse, and digital manipulation — have created a corrosive information environment. The quasi-public sphere is now an "information-industrial complex" (Powers and Jablonski 2015), where people have little choice but to endure these harms in order to engage socially or democratically.

Not surprisingly, these failures have deeply eroded public trust in social media, governments and the integrity of the broader information environment. In a recent Pew Internet Study, more than half of Americans cited false news and misinformation as a greater threat than terrorism, with majorities indicating the issue has reduced their trust in both government (68 percent) and other citizens (54 percent) (Siddiqui 2019). Another recent poll found a majority believed social media does more to "spread lies and falsehoods" (55 percent) and "divide the country" (57 percent) than it does to spread actual news (Murray 2019). Most also distrusted social media companies — 60 percent did not trust Facebook "at all" to protect their information — yet seven in 10 report using social media daily (ibid.).

More than half of Americans cited false news and misinformation as a greater threat than terrorism.

Americans are not alone. CIGI's recent poll of 25,229 internet users in 26 countries found an average of 86 percent of people internationally reported falling for fake news (quoted in Thompson 2019). Canadians were fooled at an even higher rate (90 percent) and also cited social media as their top source of distrust online (89 percent), more than even cybercriminals (85 percent). Yet Canadians do not stay away. According to the Canadian Internet Registration Authority's 2019 Internet Factbook, an annual online survey of Canadian internet use, 60 percent of the 2,050 Canadians polled in March 2019 indicated that they use social media daily.

Any new consumer protection paradigm designed for the digital era must address, beyond traditional consumer concerns, these realities — ensuring protection for consumers whose information environment and the democratic activities therein, is targeted, surveilled, manipulated and distorted by public and private sector forces in ways no other previous era has experienced. It must also be driven by an agenda that rebuilds trust in the information environment and is sensitive to its importance to — and impact on — the democratic activities of citizens.

Why Information Consumer Protection?

These are complex challenges without simple solutions. Why might information consumer

protection offer the right regulatory framework to address them?

First, social media platforms' capacity to manipulate, deceive and mistreat users derives in no small part from their powerful monopolies in the information environment. Countering the power, leverage and abusive practices of such entrenched industries has been a core driver of consumer protection law and policy for a half century. The right to consumer choice, heralded by Kennedy in 1962, provides the link between antitrust and consumer protection law. The former aims to ensure competition so that consumers have a range of choices, while the latter aims to enable consumers to make their choices effectively (Lande and Averitt 1997). An information consumer protection paradigm, attuned to and reoriented for the digital era, offers a solid regulatory framework with which to tackle the market monopolies and abuses of unresponsive platforms.

Social media platforms' capacity to manipulate, deceive and mistreat users derives in no small part from their powerful monopolies in the information environment.

Second, another central focus of consumer protection has been addressing and mitigating information asymmetries — another threat to effective consumer choice, and a pervasive and entrenched dimension in the data-driven economy (Ciuriak 2018). Information asymmetry refers to the uneven balance of information between parties in a transaction, giving the party possessing power leverage over the other (Akerlof 1970). Such imbalances underpin many digital era harms and democratic challenges. They exist, for example, between platforms and information consumers, with platforms possessing vast resources, technical expertise and power to experiment on users and shape the information environment for corporate aims with little transparency. But these are far from the only information asymmetries in the digital sphere. Others include information imbalances between *any* state or corporate interests and the regular users they are seeking to influence online. These state and corporate actors increasingly tailor, amplify and spread their targeted messages or disinformation through other mechanisms, such as data-driven profiling, promoted or paid content, automated accounts/botnets and coordinated troll networks. And with the emergence of big data, algorithms and AI, these imbalances may only deepen. An information consumer protection framework, informed by countless historical successes in overcoming such asymmetries — including in the technology context — is well positioned to offer a path forward (Morgner, Freiling and Benenson 2018).

Third, other core consumer protection principles — such as quality and safety, transparency, accountability and consumer representation — are broad enough to encompass the wide range of market-based and democratic threats in the digital information environment. Consumer protection laws have long provided people with assurance of quality and safety (Klein, n.d.) and thus can provide a framework to regulate typical user concerns about information and content quality, such as fake news, disinformation and content moderation, or health and safety concerns relating to cyberbullying, harassment and online child safety. Their transparency and accountability principles likewise provide a regulatory foundation for concerns about algorithmic accountability and information and data protection audits. Consumer protection laws against food adulteration also provide a regulatory framework to prevent what might be called "information adulteration" — the addition of harmful additives to a person's information environment, such as fake news, that reduce its quality and integrity. An information consumer protection paradigm is both normatively and practically broad to cover a wide range of information consumer interests.

Information consumer protection cannot be left solely to users and consumers to enforce.

Finally, a new information consumer protection movement can be a global move and also hit the ground running by taking advantage of existing legal, regulatory and government infrastructure around the world. Consumer protection is international in scope and origins, with historical precedents in the United States, Europe and Asia (Hilton 2012). And international consumer protection regimes and dedicated governmental agencies now exist (Corradi 2015; Micklitz and Saumier 2018). This legal and governmental infrastructure at both the national and the international levels can be immediately built upon, reshaped or repurposed, and then deployed. This is important, as information consumer protection cannot be left solely to users and consumers to enforce — one lesson learned in European data protection. Powerful agencies, such as the Federal Trade Commission (FTC), will be essential to success in exercising new antitrust and information consumer protection powers. And, in the United States, building information consumer protection on the existing consumer protection framework — given its long legal history and application — has a good chance of withstanding First Amendment scrutiny.

Putting Information Consumer Protection into Practice

Although a comprehensive treatment is beyond the scope of this essay, this section offers some concrete ideas for reform under a new information consumer protection framework.

Information Quality and Adulteration

An information consumer protection regulatory framework would empower users to better judge the quality and integrity of information on platforms. First, following past consumer protection quality and safety measures such as food labelling or energy product disclosure laws, social media platforms could be subject to mandatory "account labelling." For example, labels could include a verification of the account's identity, location and whether the account is real or automated, and whether the account or its content has been sponsored or promoted presently or in the past. These labels would immediately allow users to better judge the quality and integrity of an information source. Second, regulatory measures can be taken to prevent what I call "information adulteration," that is, the reduction in quality of the information environment through harmful additives such as fake news. Here, platforms could be required by law to, on notice, remove or reduce the visibility and/or distribution of patently false information or deceptive media (such as the faked Nancy Pelosi video). They also could be mandated to issue standardized disinformation corrections, recalls or warnings, like safety warnings on consumer packaging. A combination of strict liability, reasonable due diligence requirements and conditional safe harbour from penalties and legal liability would ensure compliance.[4] During election periods, these duties could be heightened with additional transparency — as with political ads — required.

Content Transparency, Accountability and Representation

An information consumer protection would also seek to improve existing platform content moderation practices through greater transparency, by requiring disclosure of content or account removals (by notice or placeholders), with information as to the basis for removal and the specific law or term of service violated. Following consumer protection safety inspection regimes in other contexts, platforms could be subject to information and content moderation audits, for example, inspection of algorithms and moderator teams. Stronger consumer or user representation would also be warranted in any

There is new research suggesting such warnings reduce polarizing behaviour and promote engagement.

new procedural content moderation solutions proposed by platforms, such as Facebook's internal review panel — presumably based in part on institutional review boards found at universities — established in 2014 to ethically review internal studies (Facebook 2014), or its forthcoming "Oversight Board for Content Decisions" to review the platform's content moderation practices (Facebook 2019).

Information Environment Safety

Safety for information consumers would be another central focus. One regulatory model here would be *preventative* protection measures that aim to prevent unsafe content or other harmful activities on platforms *before* being introduced. Such preventative measures have been central to consumer protection in the past, such as drug testing and approval laws that require companies to establish the safety of drugs before being introduced into the market. Examples of such protections for information consumers would be mandated content warnings, so that users receive a warning when content they seek to post violates terms of services *before* they share harmful content. There is new research suggesting such warnings reduce polarizing behaviour and promote engagement (Matias 2019). Another example would be requiring a "cooling off" period for new users, so that new accounts on platforms would be restricted in functionality until they had "proven" themselves to be safe, through good behaviour. This period would help undercut spammers, trolls and harassers from circumventing bans or propagating spam and false information through new or multiple "sock puppet" (fictional) account identities.

Information Consumer Rights Enforcement

A new information consumer protection framework would empower governmental consumer protection agencies such as the FTC with new powers to enforce information consumer rights. Efforts to enforce these rights would include working to rebuild trust in the information environment by aggressively pursuing new forms of deceptive and unfair platform practices — for example, distorted content moderation, information consumer manipulation, and data mishandling and misappropriation. Another measure that would build on traditional consumer protections against deceptive and unfair business practices — but be redefined for information consumer challenges such as distorted content moderation and fake news — would be information consumer audits, carried out by the FTC or equivalent agencies, to investigate forms of user experimentation and manipulation on platforms, such as the Facebook contagion study (McNeal 2014).

Conclusion

The proposed new regulatory framework to protect the unique vulnerabilities of the information consumer may be an appropriate solution to meet challenges we are aware of today, but the public and private threats and democratic challenges of the digital age are complex and constantly evolving. There is no silver bullet. And platform and governmental inaction have made things worse in this space, creating a corrosive information environment, rife with distrust. This is a problem: extensive literature speaks to the importance of trust (Rainie and Anderson 2017). It is a fundamental social, political and economic "binding agent," key to social capital, commerce, democracy and overall public satisfaction. Consumer protection laws have historically been employed to rebuild and sustain public trust in business, government and social institutions (Klein, n.d.). A strong information consumer protection framework would be an important step in this worthwhile direction.

NOTES

1 See https://history.house.gov/Historical-Highlights/1901-1950/Pure-Food-and-Drug-Act/.

2 See https://hoofnagle.berkeley.edu/2015/05/07/president-kennedy-consumer-bill-of-rights-march-15-1962/.

3 See www.encyclopedia.com/finance/encyclopedias-almanacs-transcripts-and-maps/consumer-bill-rights.

4 In the United States such measures would be contrary to blanket immunity for platforms under section 230 of the Communications Decency Act. However, section 230 has come under increasing criticism in recent years, with experts calling for reforms to its legal immunity. See, for example, Citron and Wittes (2017). Protection for information consumers could constitute an exception to its blanket protections.

WORKS CITED

Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84 (3): 488–500.

boyd, danah. 2010. "Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications." In *Networked Self: Identity, Community, and Culture on Social Network Sites*, edited by Zizi Papacharissi, 39–58. New York, NY: Routledge.

Calo, Ryan. 2014. "Digital Market Manipulation." *The George Washington Law Review* 82 (4): 995–1051.

Citron, Danielle Keats and Benjamin Wittes. 2017. "The Internet Will Not Break: Denying Bad Samaritans §230 Immunity." *Fordham Law Review* 86 (2): 401–23. https://ir.lawnet.fordham.edu/flr/vol86/iss2/3/.

Ciuriak, Dan. 2018. "The Economics of Data: Implications for the Data-driven Economy." Cigionline.org, March 5. www.cigionline.org/articles/economics-data-implications-data-driven-economy.

Corradi, Antonella. 2015. "International Law and Consumer Protection: The history of consumer protection." Hauser Global Law School Program. www.nyulawglobal.org/globalex/International_Law_Consumer_Protection.html.

Facebook. 2014. "Research at Facebook." Facebook Newsroom, October 2. https://newsroom.fb.com/news/2014/10/research-at-facebook/.

———. 2019. "Draft Charter: An Oversight Board for Content Decisions." Facebook Newsroom, January 28. https://fbnewsroomus.files.wordpress.com/2019/01/draft-charter-oversight-board-for-content-decisions-2.pdf.

Geis, Gilbert and Herbert Edelhertz. 1973. "Criminal Law and Consumer Fraud: A Sociolegal View." *American Criminal Law Review* 11: 989–1010.

Hilton, Matthew. 2012."Consumer Movements." In *The Oxford Handbook of the History of Consumption*, edited by Frank Trentmann, 505–20. Oxford, UK: Oxford University Press.

Hirsh, Jesse. 2019. "Why Social Platforms Are Taking Some Responsibility for Content." Cigionline.org, September 11. www.cigionline.org/articles/why-social-platforms-are-taking-some-responsibility-content.

Klein, Daniel B. n.d. "Consumer Protection." The Library of Economics and Liberty. www.econlib.org/library/Enc/ConsumerProtection.html.

Lande, Robert H. and Neil W. Averitt. 1997. "Consumer Sovereignty: A Unified Theory of Antitrust and Consumer Protection Law." *Antitrust Law Journal* 65: 713–45.

Matias, J. Nathan. 2019. "Preventing harassment and increasing group participation through social norms in 2,190 online science discussions." *PNAS* 116 (20): 9785-89.

McNeal, Gregory S. 2014. "Facebook Manipulated User News Feeds To Create Emotional Responses." *Forbes*, June 28. www.forbes.com/sites/gregorymcneal/2014/06/28/facebook-manipulated-user-news-feeds-to-create-emotional-contagion/#5291200139dc.

Micklitz, Hans-W. and Geneviève Saumier, eds. 2018. *Enforcement and Effectiveness of Consumer Law*. New York, NY: Springer.

Morgner, Philipp, Felix Freiling and Zinaida Benenson. 2018. "Opinion: Security Lifetime Labels — Overcoming Information Asymmetry in Security of IoT Consumer Products." In *Proceedings of the 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks*, 2018–11. Stockholm, Sweden: Association for Computing Machinery.

Murray, Mark. 2019. "Polls: Americans give social media a clear thumbs-down." NBC News, April 5. www.nbcnews.com/politics/meet-the-press/poll-americans-give-social-media-clear-thumbs-down-n991086.

Pistis, Marco. 2018. "Italy: From Caveat Emptor to Caveat Venditor — a Brief History of English Sale of Goods Law." *Mondaq*, January 26.

Powers, Shawn M. and Michael Jablonski. 2015. *The Real Cyber War: The Political Economy of Internet Freedom*. Urbana, IL: University of Illinois Press.

Rainie, Lee and Janna Anderson. 2017. "The Fate of Online Trust in the Next Decade." Pew Research Center, August 10. www.pewinternet.org/2017/08/10/the-fate-of-online-trust-in-the-next-decade/.

Siddiqui, Sabrina. 2019. "Half of Americans see fake news as bigger threat than terrorism, study finds." *The Guardian*, June 7. www.theguardian.com/us-news/2019/jun/06/fake-news-how-misinformation-became-the-new-front-in-us-political-warfare.

Sky News. 2019. "Pinterest teams up with health bodies to tackle false vaccine information." Sky News, August 29. https://news.sky.com/story/pinterest-teams-up-with-health-bodies-to-tackle-false-vaccine-information-11796591.

Srnicek, Nick. 2016. *Platform Capitalism*. Hoboken, NJ: Wiley.

Thompson, Elizabeth. 2019. "Poll finds 90% of Canadians have fallen for fake news." CBC, June 11. www.cbc.ca/news/politics/fake-news-facebook-twitter-poll-1.5169916.

Waterson, Jim. 2019. "Facebook refuses to delete fake Pelosi video spread by Trump supporters." *The Guardian*, May 24. www.theguardian.com/technology/2019/may/24/facebook-leaves-fake-nancy-pelosi-video-on-site.

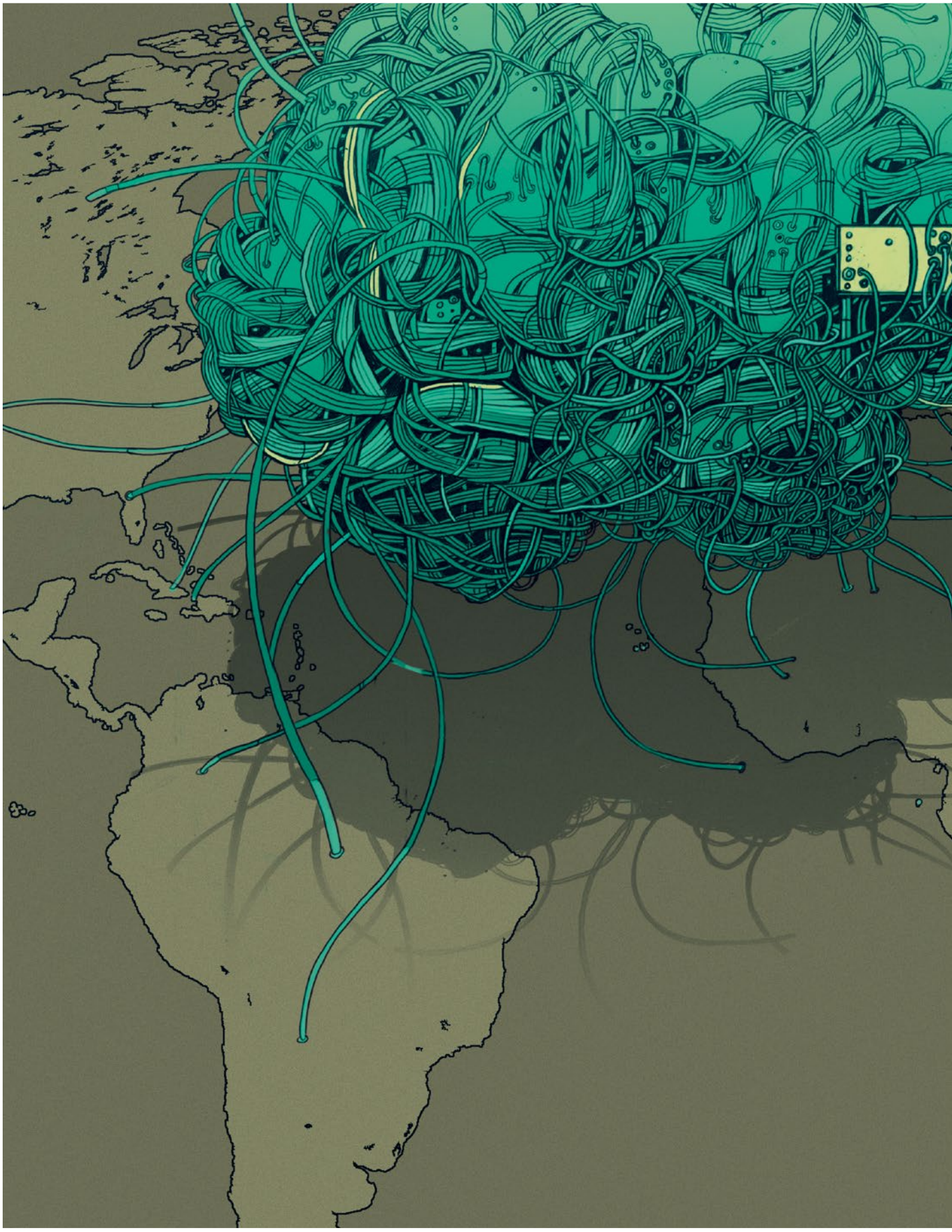
York, Jillian C. 2010. "Policing Content in the Quasi-Public Sphere." OpenNet Initiative bulletin, September. https://opennet.net/policing-content-quasi-public-sphere.

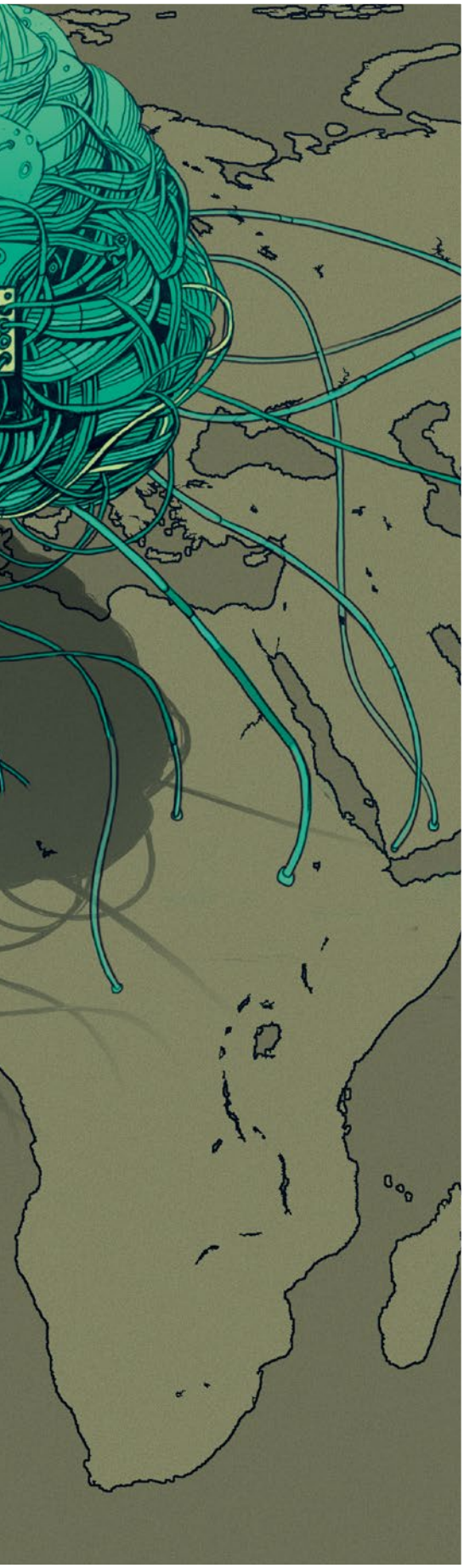
Zuboff, Shoshana. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30 (1): 75–89.

———. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York, NY: PublicAffairs.

ABOUT THE AUTHOR

Jon Penney is a legal scholar and social scientist who does research at the intersection of law, technology and human rights, with strong interdisciplinary and empirical dimensions. He is presently a visiting scholar at Harvard Law School and a research affiliate of Harvard's Berkman Klein Center for Internet & Society. As well, he is a long-time research fellow at the Citizen Lab based at the University of Toronto's Munk School of Global Affairs and Public Policy, and a research associate of the CivilServant project based at Cornell University's Department of Communication. Starting in July 2020, he will be an associate professor at Osgoode Hall Law School at York University in Toronto.





Karine Perset, Jeremy West, David Winickoff and Andrew Wyckoff

Moving "Upstream" on Global Platform Governance

Platforms are hardly a new phenomenon[1] and include newspapers, stock markets and credit card companies. And although more recent than traditional platforms, *online* platforms have been around since the early days of the World Wide Web, with early examples such as America Online (or AOL, as it later became known), Netscape and Myspace typifying these "as digital services that facilitate interactions between two or more distinct but interdependent sets of users (whether firms or individuals) who interact through the service via the Internet" (Organisation for Economic Co-operation and Development [OECD] 2019a, 11). Such platforms are said to be two-sided or, in some cases, multi-sided, with each side having its own distinct set of services that appeal to different audiences.

What differentiates today's online platforms are a number of economic properties fuelled by the digital transformation that have led to a new cadre of platforms with a global reach, trillion-dollar valuations and huge profits that feed large research and development (R&D) efforts, as well as significant merger and acquisition activity, and ever-growing engagement in public affairs. These characteristics include powerful network effects; cross-subsidization; scale without mass, which enables a global reach; panoramic scope; generation and use of user data to optimize their services; substantial switching costs; and, in some markets, winner-take-all or winner-take-most tendencies. Skillfully exploiting these properties, online platforms have become very popular, and we now depend on them for everything from entertainment and news to searching for jobs and employees, booking transportation and accommodation, and finding partners.

Their popularity underscores the fact that online platforms have brought benefits to economies and societies, including considerable consumer welfare. They also raise a new set of important policy challenges, ranging from the classification of workers (for example, contractors versus employees) to the misuse of user data to the adverse impacts of tourists in city centres. Likewise, some online platforms have drawn the attention of competition authorities

That inadequacy has prompted leaders of some technology companies to ask for a new regulatory scheme.

and other regulatory bodies to issues ranging from abuse of dominance to taxation.

The growing prominence of these issues on policy agendas at both the national and the international level provokes a more fundamental question about how to craft an appropriate model of governance that strikes the balance between, on the one hand, promoting online platform innovation and productivity and, on the other hand, achieving basic policy objectives, such as ensuring sufficient competition, protecting consumers and workers, and collecting tax revenue. It is clear that many of the regulations designed for traditional businesses are not a good fit for online platforms. In the last few years, that inadequacy has prompted leaders of some technology companies to ask for a new regulatory scheme. Testifying before the US Congress, Facebook's CEO Mark Zuckerberg stated: "My position is not that there should be no regulation.…I think the real question, as the Internet becomes more important in people's lives, is what is the right regulation, not whether there should be or not" (CBC 2018). Microsoft President Bradford Smith has stated a similar position regarding facial recognition software, saying that "we live in a nation of laws, and the government needs to play an important role in regulating facial recognition technology" (Singer 2018).

But regulating these businesses is far more complex than the political debate would suggest, largely because platforms vary significantly, rendering any omnibus "platform regulation" unsuitable. The OECD recently undertook in-depth profiles of 12 successful online platform firms and found that they differed widely on a number of different axes (such as size, functionality, income and profitability) and cannot be compartmentalized into just a few categories, let alone a single sector (OECD 2019a, 12). They differ in how they generate income, with some drawing revenue from advertisers, others from transaction fees, still others from subscriptions and some from a combination of those. Online platforms serve different needs of different customers looking for different things. Indeed, it is striking how many different economic activities online platforms encompass.

Another factor that tends to be overlooked in the current debate about online platforms is the rise of Chinese platforms, which yet again add to the diversity. Policy makers' attention, understandably, has been largely focused on the big Western platform companies, in particular, Amazon, Apple, Facebook and Google. Discussions are under way about placing new regulations on them and even about breaking them up. Conversely, sparse attention is being devoted to preparing for the expansion of the major Chinese platform firms: Baidu, Alibaba and Tencent (ibid.). While their platforms have been seen as being largely confined to China, this is quickly changing, due in large part to the integration of mobile payment apps into the platforms. For example, Tencent's WeChat Pay is already available in 49 countries outside of China (Wu 2018). This access will only grow as merchants move to serve Chinese tourists, already the number-one tourist demographic both by number and money being spent (OECD 2019a, 77-78). Given the global reach of the online platforms, it is important that a debate about governance schemes be forward-looking and take into account all the big global players.

As eager as policy makers are to act, effective policies need careful deliberation. *Ex ante* regulation can be problematic if it's not fully grounded in experience and insight, characteristics that can be difficult to achieve in periods of rapid transformation awash in steady streams of new technologies. *Ex post* regulation may face resistance from embedded interests, including dedicated user bases. At the same time, when policy intervention is either too frequent or absent altogether, uncertainty arises, which may limit investments and innovation. Countries with a well-established and elaborated policy framework and constituencies built up over hundreds of years

may be disadvantaged relative to emerging economies or countries that have recently switched from one system (for example, communism) to another, providing them with "leapfrog" opportunities. To add to this list of considerations, online platforms with two- or multi-sided markets frequently have product scope (providing information services as well as, say, financial payments and the provision of media) that crosses traditional policy domains segmented by government ministries and agencies, requiring a joined-up approach to policy.

The traditional "end-of-pipe" regulation that focuses on a single final product and tries to fit that to an existing policy framework is ill-suited to highly innovative and dynamic online platform businesses with global reach. Instead, a new, more anticipatory and upstream approach is needed,[2] one that uses the multi-stakeholder model to collectively shape developments so that innovation is encouraged and productivity-boosting disruption enabled, but within a set of publicly defined policy objectives.

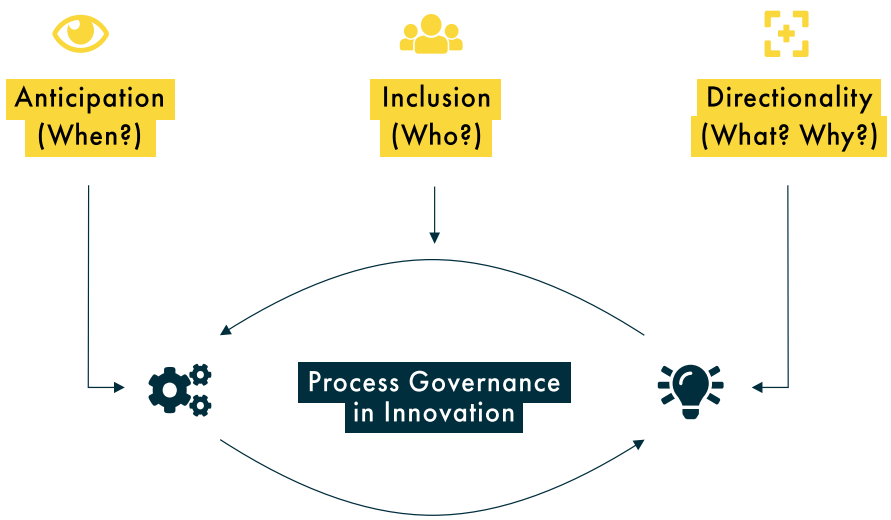
Moving Upstream

The governance of emerging technologies poses a well-known puzzle: the so-called Collingridge dilemma holds that early in the innovation process — when interventions and course corrections might still prove easy and cheap — the full consequences of the technology — and hence the need for change — might not be fully apparent (Collingridge 1980).

Conversely, when the need for intervention becomes apparent, changing course may become expensive, difficult and time-consuming. Uncertainty and lock-ins are at the heart of many governance debates (Arthur 1989; David 2001) and continue to pose questions about "opening up" and "closing down" development trajectories (Stirling 2008).

Several emerging approaches in science policy seek to overcome the Collingridge dilemma by engaging concerns with technology governance "upstream." Process governance shifts the locus from managing the risks of technological products to managing the innovation process itself: who, when, what and how. It aims to anticipate concerns early on, address them through open and inclusive processes, and steer the innovation trajectory in a desirable direction. The key idea is to make the innovation process more anticipatory, inclusive and purposive (Figure 1), which will inject public good considerations into innovation dynamics and ensure that social goals, values and concerns are integrated as they unfold.

Figure 1: Three Imperatives of a Process-based Approach to Governance



Source: Adapted and reproduced with permission from the OECD (Winickoff and Pfotenhauer 2018, 224).

Artificial Intelligence as a Test Case

While governments have experimented with policies that seek to move upstream by adopting a process-based approach to governance, the global nature of online platforms, which are currently some of the largest actors in artificial intelligence (AI), demands an international approach — a requirement all the more challenging to meet when the relevancy of many multilateral institutions is being questioned.[3]

AI, which owes some of its recent advances to innovations from online platforms (whose R&D spending on AI dwarfs that of any countries' investment in it), presents itself as a classic case warranting a process-based approach to governance. AI is reshaping economies, promising to generate productivity gains, improve efficiency and lower costs. It contributes to better lives and helps people make better predictions and more informed decisions. At the same time, AI is also fuelling anxieties and ethical concerns. There are questions about the trustworthiness of AI systems, including the dangers of codifying and reinforcing existing biases, or of infringing on human rights and values such as privacy. Concerns are growing about AI systems exacerbating inequality, climate change, market concentration and the digital divide.

There are questions about the trustworthiness of AI systems, including the dangers of codifying and reinforcing existing biases, or of infringing on human rights and values such as privacy.

AI technologies, however, are still in their infancy. At a Digital Ministers Group of Seven (G7) meeting in Takamatsu, Japan, in 2016, the Japanese G7 presidency proposed the "Formulation of AI R&D Guidelines" and drafted eight principles for AI R&D. Japan began to support OECD work on AI, multi-stakeholder consultations and events, and analytical work (OECD 2019b). G7 work on AI was furthered in 2017, 2018 and 2019 under the Italian, Canadian and French presidencies, with ministers' recognition that the transformational power of AI must be put at the service of people and the planet. In this sense, the G7 had begun to espouse the "anticipation" imperative of process governance.

In May 2018, the OECD's Committee on Digital Economy Policy (CDEP) established an Expert Group on Artificial Intelligence (AIGO) to explore the development of AI principles. This decision effectively echoed the G7 call for *anticipation* and took a multi-stakeholder *inclusive* approach. The AIGO consisted of more than 50 experts, with representatives from each of the stakeholder groups, as well as the European Commission and the UN Educational, Scientific and Cultural Organization. It held meetings in Europe, North America and the Middle East, and produced a proposal containing AI principles for the responsible stewardship of trustworthy AI. This work was the basis for the OECD Council's eventual recommendation on AI, which the CDEP agreed to at a special meeting in March 2019, and recommended the OECD Council, meeting at ministerial level, adopt. That adoption occurred in May 2019 when 40 countries adhered to the recommendation.[4]

In June 2019, the Group of Twenty (G20) Ministerial Meeting on Trade and Digital Economy in Tsukuba adopted human-centred AI principles that were informed by the OECD AI principles.[5] This action vastly improved the global reach of the principles.

The OECD AI principles focus on features that are specific to AI and set a standard that is implementable and sufficiently flexible to stand the test of time in a rapidly evolving field. In addition, they are high-level and context-dependent, leaving room for different implementation mechanisms, as appropriate to the context and consistent with the state of

Another option to spur faster and more effective decisions is the use of digital tools to design policy, including innovation policy, and to monitor policy targets.

art. As such, these principles effectively provide *directionality* for the development of AI. In this way, they provide an example of a possible new governance model applicable to platforms as they signal a shared vision of how AI should be nurtured to improve the welfare and well-being of people, sustainability, innovation and productivity, and to help respond to key global challenges. Acknowledging that the nature of future AI applications and their implications are hard to foresee, these principles aim to improve the trustworthiness of AI systems — a key prerequisite for the diffusion and adoption of AI — through a well-informed whole-of-society public debate, to capture the beneficial potential of AI, while limiting the risks associated with it.

Importantly, the principles also recommend a range of other actions by policy makers that seek to move governance upstream, including governments using experimentation to provide controlled environments for the testing of AI systems. Such environments could include regulatory sandboxes, innovation centres and policy labs. Policy experiments can operate in "start-up mode." In this case, experiments are deployed, evaluated and modified, and then scaled up or down, or abandoned quickly. Another option to spur faster and more effective decisions is the use of digital tools to design policy, including innovation policy, and to monitor policy targets. For instance, some governments use "agent-based modelling" to anticipate the impact of policy variants on different types of businesses. Governments can also encourage AI actors to develop self-regulatory mechanisms, such as codes of conduct, voluntary standards and best practices. These can help guide AI actors through the AI life cycle, including for monitoring, reporting, assessing and addressing harmful effects or misuse of AI systems. Finally, governments can also establish and encourage public and private sector oversight mechanisms of AI systems, as appropriate. These could include compliance reviews, audits, conformity assessments and certification schemes.

By coupling flexible, upstream interventions across and between levels, international and national, a new paradigm of governance for online platforms and for AI is beginning to emerge. This new approach strives both to enable innovation and the benefits that these platforms and actors deliver to society, but also to provide a means for channelling this "creative destruction" toward societal public policy goals.

AUTHORS' NOTE

The authors are all members of the OECD Secretariat — follow their work at www.oecd.org/sti and on Twitter (@OECDinnovation). The views and opinions here are those of the authors and should not be attributed to the OECD or its member countries.

NOTES

1 This section is drawn from *An Introduction to Online Platforms and Their Role in the Digital Transformation*, published by the OECD (2019a).

2 The discussion of such an approach in the section "Moving Upstream" draws on and is indebted to *OECD Science, Technology and Innovation 2018: Adapting to Technological and Societal Disruption* (OECD 2018), in particular, chapter 10.

3 This section draws on the OECD's recent report *Artificial Intelligence in Society* (OECD 2019b).

4 See https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

5 See www.oecd.org/going-digital/ai/principles/.

WORKS CITED

Arthur, W. Brian. 1989. "Competing Technologies, Increasing Returns, and Lock-In by Historical Events." *The Economic Journal* 99 (394): 116–31. http://dx.doi.org/10.2307/2234208.

CBC. 2018. "Zuckerberg sees regulation of social media firms as 'inevitable.'" CBC News, April 11. www.cbc.ca/news/technology/facebook-zuckerberg-users-privacy-data-mining-house-hearings-1.4614174.

Collingridge, David. 1980. *The Social Control of Technology*. London, UK: Frances Pinter.

David, P. A. 2001. "Path dependence, its critics and the quest for 'historical economics.'" In *Evolution and Path Dependence in Economic Ideas: Past and Present*, edited by Pierre Garrouste and Stavros Ioannides, 15–40. Cheltenham, UK: Edward Elgar Publishing.

OECD. 2018. *OECD Science, Technology and Innovation 2018: Adapting to Technological and Societal Disruption*. Paris, France: OECD. https://doi.org/10.1787/sti_in_outlook-2018-en.

———. 2019a. *An Introduction to Online Platforms and Their Role in the Digital Transformation*. Paris, France: OECD Publishing. https://doi.org/10.1787/53e5f593-en.

———. 2019b. *Artificial Intelligence in Society*. Paris, France: OECD Publishing. https://doi.org/10.1787/eedfee77-en.

Singer, Natasha. 2018. "Microsoft Urges Congress to Regulate Use of Facial Recognition." *The New York Times*, July 13. www.nytimes.com/2018/07/13/technology/microsoft-facial-recognition.html.

Stirling, Andy. 2008. "'Opening Up' and 'Closing Down': Power, Participation, and Pluralism in the Social Appraisal of Technology." *Science, Technology & Human Values* 33 (2): 262–94. doi:10.1177/0162243907311265.

Winickoff, David E. and Sebastian M. Pfotenhauer. 2018. "Technology Governance and the Innovation Process." In *OECD Science, Technology and Innovation 2018: Adapting to Technological and Societal Disruption*, 221–39. Paris, France: OECD. https://doi.org/10.1787/sti_in_outlook-2018-en.

Wu, Julianna. 2018. "A surprising number of countries now accept WeChat Pay or Alipay." Tech in Asia, December 10. www.techinasia.com/surprising-number-countries-accept-wechat-pay-alipay.

ABOUT THE AUTHORS

Karine Perset is an economist and policy analyst on AI policy at the OECD, in the Division for Digital Economy Policy in Paris. She focuses on trends in development and diffusion of AI and on opportunities and challenges that AI raises for public policy and is developing OECD.AI, the OECD's AI Policy Observatory. She was previously senior director and adviser to the Internet Corporation for Assigned Names and Numbers' Governmental Advisory Committee in Los Angeles, and before that was the counsellor of the OECD's Directors for Science, Technology and Innovation. She has a dual master's degree in telecommunications and international economics from the University of Paris Dauphine.

Jeremy West is senior policy analyst in the Division for Digital Economy Policy at the OECD. Jeremy is a lawyer with experience at the OECD, the US Department of Justice's Antitrust Division, a Washington, DC law firm and the New Zealand Commerce Commission. His OECD papers have been quoted and cited in scholarly journals, by the US Antitrust Modernization Commission, in *The Financial Times*, *Yomiuri Shimbun* and *Kyodo News*, and by an advocate general of the European Court of Justice. He currently serves on the editorial board of the online journal *Oxford Competition Law* and was an assistant, associate and senior editor for the *Antitrust Law Journal* from 2009 to 2017. After working in the OECD's Competition Division for nine years, he moved to the Digital Economy Policy Division in 2013 to lead a multi-disciplinary project on intellectual property's economic impact.

David E. Winickoff is senior policy analyst and secretary of the Working Party on Bio-, Nano- and Converging Technologies at the OECD in the Directorate for Science, Technology and Innovation. He holds a J.D. from Harvard University and was a fellow for two years at the Harvard Kennedy School. Formerly a tenured professor at University of California, Berkeley, David currently works on innovation policies for and the governance of emerging technologies in an international context. He is also an affiliated professor of law at the Paris Institute of Political Studies, where he teaches biotechnology policy and global governance.

Andrew W. Wyckoff is the director of the OECD's Directorate for Science, Technology and Innovation where he oversees OECD's work on innovation, business dynamics, science and technology, and information and communication technology (ICT) policy, as well as the statistical work associated with each of these areas. His experience prior to the OECD includes positions at the US Congressional Office of Technology Assessment, the US National Science Foundation and the Brookings Institution. He has served as an expert on various advisory groups and panels, which includes joining the Digital Future Society Global Board of Trustees, being a member of the Research Advisory Network of the Global Commission on Internet Governance and the International Advisory Board of the Research Council of Norway, and heading the OECD's delegation at the G20 and G7 meetings on ICT and the digital economy.

75%

say social media companies are responsible for distrust in the internet.

CIGI-Ipsos Gobal Survey on

Internet Security and Trust

The CIGI-Ipsos Global Survey, now in its fifth year, is the world's largest and most comprehensive survey of internet security and trust, involving more than 25,000 internet users in over two dozen countries across North America, Latin America, Europe, the Middle East, Africa and the Asia-Pacific region.

See the full survey results: **www.cigionline.org/internet-survey**

Centre for International Governance Innovation

Ipsos

Internet Society

United Nations UNCTAD

Victor Pickard

Public Investments for Global News

Commercial journalism is failing in many countries around the world.[1] Numerous factors contribute to this crisis, but at its root lies a "systemic market failure" in which for-profit news institutions, in particular those that depend on advertising revenue, are increasingly unviable. Symptomatic of this broader decline, traditional news organizations' cost-cutting measures are depriving international news operations of the considerable resources they need to survive. In recognition of this systemic failure, a healthy global civil society requires structural alternatives — specifically, non-market models — to support adequate levels of journalism. In particular, we need a large global trust fund to finance international investigative reporting. Given their role in hastening the decline of journalism and proliferating misinformation, platform monopolies should contribute to this fund to offset social harms. While setting up such a fund will be novel in many respects, there are models and policy instruments we can use to imagine how it could support the public service journalism that the market no longer sustains.

The Broader Context

As commercial journalism collapses around the world, a glaring lesson comes into focus: no commercial model for journalism can adequately serve society's democratic needs. To be more specific, no purely profit-driven media system can address the growing news deserts that are sprouting up all over the United States and beyond, or fill the various news gaps in international coverage. If we come to see systemic market failure for what it is, we will acknowledge that no entrepreneurial solution, no magical technological fix, no market panacea lies just beyond discovery. While subscription and membership models might sustain some relatively niche outlets and large national and international newspapers such as *The New York Times* and *The Guardian*, they do not provide a systemic fix. In particular, they cannot support the public service journalism — the local, policy, international and investigative reporting — that democratic society requires. Commercial journalism's structural collapse has devastated international news at a time when global crises are worsening — from climate change to growing inequalities to fascistic political movements. At perhaps no other time has the need for reliable, fact-based, and well-resourced journalism been more acute.

(Silverman 2016; Vaidhyanathan 2018). It has also played a key role in destabilizing elections in places such as the Philippines and in facilitating ethnic cleansing in Myanmar (Stevenson 2018). In short, Facebook has hurt democracy around the world. Considering the accumulating damage it has wreaked and the skewed power asymmetry between Facebook and its billions of users, we need a realignment.

The social harms of Facebook's monopoly power is especially apparent in how it corrupts the integrity of our news and information systems. As an algorithm-driven gatekeeper over a primary information source for more than two billion users, Facebook wields tremendous political economic power. In the United States, where Americans increasingly access news through the platform, Facebook's role in the 2016 presidential election has drawn well-deserved scrutiny. Moreover, along with Google, Facebook is devouring the lion's share of digital advertising revenue and starving the traditional media that provide quality news and information — the same struggling news organizations that these platforms expect to help fact-check against misinformation (Kafka 2018). Journalism's financial future is increasingly threatened by the Facebook-Google duopoly, which in recent years took a combined 85 percent of all new

The social harms of Facebook's monopoly power is especially apparent in how it corrupts the integrity of our news and information systems.

Coinciding with the precipitous decline of traditional news media, platform monopolies — Facebook most notably — are attaining levels of media power unprecedented in human history. So far, this power has been largely unaccountable and unregulated. As Facebook extracts profound wealth across the globe, it has generated tremendous negative externalities by mishandling users' data, abusing its market power, spreading dangerous misinformation and propaganda, and enabling foreign interference in democratic elections

US digital advertising revenue growth, leaving only scraps for news publishers (Shields 2017). According to some calculations, these two companies control 73 percent of the total online advertising market.[2] Given that these same companies play an outsized role in proliferating misinformation, they should be bolstering — not starving — news outlets.

Despite the lack of silver-bullet policy solutions, this moment of increased public scrutiny offers a rare — and most

likely fleeting — opportunity to hold an international debate about what interventions are best suited to address these informational deficits and social harms. Ultimately, these problems necessitate *structural* reforms; they cannot be solved by simply shaming digital monopolies into good behaviour or by tweaking market incentives. With platform monopolies accelerating a worldwide journalism crisis, a new social contract is required that includes platform monopolies paying into a global public media fund.

Taxing Digital Monopolies and Redistributing Media Power

Platform monopolies have not single-handedly caused the journalism crisis — overreliance on market mechanisms is the primary culprit — but they have exacerbated and amplified communication-related social harms. Beyond regulating and penalizing these firms, we should require that they help undo the damage they have caused. Not only do these firms bear some responsibility, but they also have tremendous resources at their disposal. Despite a general unease about policy interventions in this arena — especially in the United States where a combination of First Amendment absolutism and market fundamentalism render many policy interventions off-limits — scholarship has long established that media markets produce various externalities (see, for example, Baker 2002). It is the role of government policy to manage them — to minimize the negative and maximize the positive externalities for the benefit of democratic society. Even the relatively libertarian United States redistributes media power with public access cable channels, the universal service fund and subsidized public broadcasting, to name just a few examples. Policy analysts also have proposed various schemes for taxing platforms in the US context.[3]

Internationally, policy makers and advocates have proposed a number of similar models. For example, in the United Kingdom, the Media Reform Coalition and the National Union of Journalists proposed allocating capital raised from taxes on digital monopolies to support public service journalism. Jeremy Corbyn echoed this plan by calling for digital monopolies to pay into an independent "public interest media fund" (Corbyn 2018). Similarly, the *Cairncross Review*, a detailed report on the future of British news media, called for a new institute to oversee direct funding for public-interest news outlets (Waterson 2019).

While these various proposals for national-level media subsidies are encouraging, the international scope of this problem requires a *global* public media fund. One proposal has called for establishing a $1 billion[4] international public interest media fund to support investigative news organizations around the world, protecting them from violence and intimidation (Lalwani 2019). According to this plan, the fund would rely on capital from social media platforms as well as government agencies and philanthropists.

According to some calculations, these two companies control 73 percent of the total online advertising market.

Platform monopolies should not be solely responsible for funding global public media, but the least they could do is support the investigative journalism, policy reporting and international news coverage that they are complicit in undermining. Thus far, Google and Facebook have each promised $300 million over three years for news-related projects. Google has pledged this money toward its News Initiative,[5] and Facebook has sponsored several projects, including its $3 million journalism "accelerator" to help 10–15 news organizations build their digital subscriptions using Facebook's platform (Ingram 2018). Another program, Facebook's "Today In" app section, aggregates local news in communities across the United States, but it ran into problems when Facebook found many areas already entirely bereft of local news (Molla and Wagner 2019). More recently, Google has announced that it would tailor its algorithms to better promote original reporting (Berr 2019), and Facebook has promised to

offer major news outlets a license to its "News Tab" that will feature headlines and article previews (Fussell 2019). Nonetheless, these initiatives are insufficient given the magnitude of the global journalism crisis — efforts that one news industry representative likened to "handing out candy every once in a while" instead of contributing to long-term solutions (quoted in Baca 2019).

Redistributing revenue toward public media could address the twin problems of unaccountable monopoly power and the loss of public service journalism. Facebook and Google (which owns YouTube) should help fund the very industry that they both profit from and eviscerate. For example, these firms could pay a nominal "public media tax" of one percent on their earnings, which would generate significant revenue for a journalism trust fund (Pickard 2018). Based on their 2017 net incomes, such a tax would yield $159.34 million from Facebook and $126.62 million from Google/Alphabet. Together, this $285.96 million, if combined with other philanthropic and government contributions over time, could go a long way toward seeding an endowment for independent journalism. A similar, but more ambitious, plan proposed by the media reform organization Free Press calls for a tax on digital advertising more broadly, potentially yielding $2 billion per year for public service journalism (Karr and Aaron 2019, 8). These firms can certainly afford such expenditures, especially since they pay precious little in taxes.

Redistributing revenue toward public media could address the twin problems of unaccountable monopoly power and the loss of public service journalism.

In countries around the world, there is a growing consensus that digital monopolies should be sharing their wealth, conceding to public oversight and taking on more responsibilities for the social harms they have caused. Increasingly, these presumed responsibilities include protecting sources — and providing resources — for reliable information. Although calls for taxing these firms have yet to succeed in any significant way, they reflect rising awareness about the connections between digital monopolies' illegitimate wealth accumulation, the continuing degradation of journalism and the rise of misinformation. If we are to grant platform monopolies such incredible power over our vital communication infrastructures, we must have a new social contract to protect democratic society from social harms.

From Theory to Action

Creating an independent public international media fund does not end with procuring sufficient resources. Once the necessary funding is in place, we have to ensure these new journalistic ventures operate in a democratic manner. We must establish structural safeguards mandating that journalists and representative members of the public govern them in a transparent fashion in constant dialogue with engaged, diverse constituencies. These funds should go to the highest-need areas, with key allocations decided democratically via public input through an international body. These resources might support already-existing organizations such as the Organized Crime and Corruption Reporting Project and the International Consortium of Investigative Reporting, or they might help create entirely new outlets. Whether this body is housed at the United Nations or some new institution, we must start pooling resources and imagining what this model might look like.

No easy fix will present itself for journalism — or for the tremendous social problems around the world — but a well-funded, international and independent public media system is a baseline requirement for tackling the global crises facing us today. It is urgent that we begin the serious conversations needed to create this fund now and move quickly into action.

NOTES

1 The analysis in this essay derives from the author's book *Democracy without Journalism? Confronting the Misinformation Society* (Oxford University Press, 2019).

2 Calculations vary somewhat; a report from eMarketer (2017) places the duopoly's share at slightly less. Amazon is gradually becoming a third significant player in digital advertising.

3 For earlier articulations of this idea, see Pickard (2016; 2018), Waldman (2017) and Bell (2017).

4 All dollar values in US currency.

5 See https://newsinitiative.withgoogle.com/.

WORKS CITED

Baca, Marie C. 2019. "Google and Facebook's latest efforts to 'save' journalism are sparking debate." *The Washington Post*, September 20. www.washingtonpost.com/technology/2019/09/13/google-facebooks-latest-efforts-save-journalism-are-already-getting-eye-rolls/.

Baker, C.E. 2002. *Media, Markets, and Democracy*. New York, NY: Cambridge University Press.

Bell, Emily. 2017. "How Mark Zuckerberg could really fix journalism." *Columbia Journalism Review*, February 21. www.cjr.org/tow_center/mark-zuckerberg-facebook-fix-journalism.php.

Berr, Jonathan. 2019. "Is Google's Embrace Of 'Original Reporting' Good News For Publishers?" *Forbes* (blog). www.forbes.com/sites/jonathanberr/2019/09/12/is-googles-embrace-of–original-reporting-good-news-for-publishers/.

Corbyn, Jeremy. 2018. Alternative MacTaggart Lecture, August 23. https://labour.org.uk/press/full-text-jeremy-corbyns-2018-alternative-mactaggart-lecture/.

eMarketer. 2017. "Looking Beyond the Facebook/Google Duopoly." eMarketer, December 12. www.emarketer.com/content/exploring-the-duopoly-beyond-google-and-facebook.

Fussell, Sidney. 2019. "Facebook Wants a Do-Over on News." *The Atlantic*, August 22. www.theatlantic.com/technology/archive/2019/08/facebooks-news-tab-will-be-run-humans-and-algorithms/596554/.

Ingram, Mathew. 2018. "The media today: Facebook tosses a dime at local journalism." *Columbia Journalism Review*, February 28. www.cjr.org/the_media_today/facebook-local-news-funding.php.

Kafka, Peter. 2018. "These Two Charts Tell You Everything You Need to Know about Google's and Facebook's Domination of the Ad Business." *Vox*, February 13. www.vox.com/2018/2/13/17002918/google-facebook-advertising-domination-chart-moffettnathanson-michael-nathanson.

Karr, Timothy and Craig Aaron. 2019. *Beyond Fixing Facebook: How the multibillion-dollar business behind online advertising could reinvent public media, revitalize journalism and strengthen democracy*. February 25. Florence, MA: Free Press. www.freepress.net/policy-library/beyond-fixing-facebook.

Lalwani, Nishant. 2019. "A free press is the lifeblood of democracy — journalists must not be silenced." *The Guardian*, July 5. www.theguardian.com/global-development/2019/jul/05/a-free-press-is-the-lifeblood-of-democracy-journalists-must-not-be-silenced.

Molla, Rani and Kurt Wagner. 2019. "Facebook Wants to Share More Local News, but It's Having Trouble Finding It." *Vox*, March 1. www.vox.com/2019/3/18/18271058/facebook-local-news-journalism-grant.

Pickard, Victor. 2016. "Yellow Journalism, Orange President." *Jacobin*, November 25. www.jacobinmag.com/2016/11/media-advertising-news-radio-trump-tv/.

———. 2018. "Break Facebook's Power and Renew Journalism." *The Nation*, April 18. www.thenation.com/article/break-facebooks-power-and-renew-journalism/.

Shields, Mike. 2017. "CMO Today: Google and Facebook Drive 2017 Digital Ad Surge." *Wall Street Journal*, March 14. www.wsj.com/articles/cmo-today-google-and-facebook-drive-2017-digital-ad-surge-1489491871.

Silverman, Craig. 2016. "This Analysis Shows How Viral Fake Election News Stories Outperformed Real News on Facebook." BuzzFeed News, November 16. www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook.

Stevenson, Alexandra. 2018. "Facebook Admits It Was Used to Incite Violence in Myanmar." *The New York Times*, November 6. www.nytimes.com/2018/11/06/technology/myanmar-facebook.html.

Waldman, Steve. 2017. "What Facebook Owes to Journalism." *The New York Times*, February 21. www.nytimes.com/2017/02/21/opinion/what-facebook-owes-to-journalism.html.

Waterson, Jim. 2019. "Public funds should be used to rescue local journalism, says report." *The Guardian*, February 11. www.theguardian.com/media/2019/feb/11/public-funds-should-be-used-to-rescue-local-journalism-says-report.

Vaidhyanathan, Siva. 2018. *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy*. New York, NY: Oxford University Press.

ABOUT THE AUTHOR

Victor Pickard is an associate professor at the University of Pennsylvania's Annenberg School for Communication where he co-directs the Media, Inequality, and Change Center. Victor's research focuses on the history and political economy of media institutions, media activism, and the politics and normative foundations of media policy. He has published dozens of scholarly articles, book chapters and policy reports, as well as essays for *The Guardian*, *The Nation*, *Jacobin* and *The Atlantic*. He has authored or edited five books, including *America's Battle for Media Democracy*; *Media Activism in the Digital Age* (with co-editor Guobin Yang); *Will the Last Reporter Please Turn Out the Lights* (with co-editor Robert W. McChesney); and *After Net Neutrality* (with co-author David Elliot Berman). His book *Democracy without Journalism? Confronting the Misinformation Society* will be published with Oxford University Press in November 2019.

Damian Tambini

Rights and Responsibilities of Internet Intermediaries in Europe: The Need for Policy Coordination

A handful of dominant platform companies that combine data gathering, media distribution and advertising now have immense power to affect our actions. By processing our personal data and monopolizing our attention, companies such as Facebook and Google can deploy artificial intelligence (AI) to optimize and target messages to us based on sophisticated profiles of our weaknesses, wants and means. In short, having our attention and data enables them to modify our behaviour — from purchasing decisions, to dating, to voting. With this power comes responsibility — to work for the public interest rather than against, and to deal with the negative outcomes that such activities can create.

Power also needs checks and balances. Countries around the world are working toward a new "settlement" on the rights and responsibilities of platforms. This settlement is not just about punitive laws. It will look at the responsibilities of the platforms themselves to protect their users and whether existing businesses have any incentives to deal with the social problems they may be associated with, as well as consider how the law will provide the right incentives to curb online harm without overbearing regulation.

Dangers of Speech Control

Economists often refer to pollution as an example of a negative externality — that is, an unintended impact of business, with undesired or damaging effects on people not directly involved. Policy makers across the political spectrum agree that digital platforms' engagement-driven business models are polluting the public realm with harmful, hateful, misleading and anti-social speech and other externalities. Self-regulation by companies such as Facebook and YouTube is seen as ineffective, and there is now a consensus in favour of more formal and transparent regulation by independent regulators or courts. From the point of view of free speech, such a consensus is dangerous. Censorship and chilling of free speech is a constant danger when considering governance of such platforms.

Censorship and chilling of free speech is a constant danger when considering governance of such platforms.

Even before this new regulatory mood, North American commentators looked to Europe to solve the problems that arose with the phenomenal success of US platform companies. The European Commission has levied eye-catching antitrust fines,[1] and in 2018 passed the General Data Protection Regulation, which rapidly became the standard for the world. The hope is that as a continent with a reputation for effective regulation and strong protection of free speech, Europe will also provide the answer for online harm.

Most recent European reforms have taken place at the member-state level. The United Kingdom and France have each published legislative proposals for social media regulation and Germany passed the Network Enforcement Act in 2017.[2] Such policies have in common an attempt to tighten the liability obligations on online intermediaries to ensure that they have strong incentives — including fines — to deal with harmful content and conduct more promptly and effectively, and reduce its prevalence online.

European Commission activity to date has focused mainly on encouraging self-regulation: corralling platforms to observe conduct guidelines on hate speech, misinformation and terrorism, without major change to legislation. The overall policy settlement established in the 1990s — under which platforms were immune from liability for hosting illegal content until notified of it, and free to develop their own rules for content that is harmful but not illegal — remains intact, although the European Commission has published more guidance on how platforms can deal with socially questionable content.[3]

Alongside the debate about illegal and harmful content, proposals have been developed to update competition and fiscal policy. Multiple European countries have implemented or made proposals for new taxes on digital platforms. The proposed "digital services tax" in the United Kingdom, for example, would be calculated as a percentage of advertising revenue in the United Kingdom. Competition regulators are looking into what new forms of antitrust are necessary to deal with multi-sided data and advertising markets. The next phase is the difficult part: coming up with pan-European rules in areas that require it, standardizing this thicket of new regulation, yet enabling national policies and forms of accountability where citizens expect them. It is inevitable that such policy developments will eventually come into conflict with the existing European legal settlement on online intermediaries, namely the directive on electronic commerce.[4]

Challenges and Contradictions

As a policy area that has implications for economic development, competition, free speech, national security, democracy, sovereignty and the future role of AI, we should not expect the negotiation of a new regulatory settlement to be simple. The initial phase in this policy cycle has revealed the following challenges and contradictions, among others.

The need for subsidiarity and accountability of standard setting, but also the pressures for regional and global standards of procedure on efficiency grounds: Sensitive issues of media policy and democratic accountability — such as those raised by, for example, the Cambridge Analytica scandal and so-called disinformation — are best resolved close to consumers in a transparent fashion. For issues such as speech considered to be insulting, threatening or even inciting violence, context and culture are paramount. Any new law on speech or definition of harmful speech will be instantly controversial, and rightly subject to calls for transparency, subsidiarity, due process and open accountability. The balance between law and self-regulation will, inevitably, be difficult to strike: platform self-monitoring schemes are often seen as preferable in that they invest power in the users themselves, but such systems have been criticized as ineffective in hindering hatred and misinformation online. In this context, the present period is characterized by struggle between platforms, parliaments and the European institutions over who sets the rules of speech. Whatever new standards and norms for speech are involved, it is clear they will be multi-levelled, in line with other European rules on speech, such as the Audiovisual Media Services Directive.[5] At some point, efficiency will dictate a Europe-wide baseline of legal standards of acceptable speech, and a layer of opt-in standards that vary by company or platform.

Controversies about who is responsible for protecting whose free speech: As platforms are ever more called upon to regulate speech to prevent negative outcomes, they are inevitably criticized as unaccountable private censors. European and international standards on freedom of expression are increasingly held up as relevant to these "censorship" activities of platforms,[6] but there is uncertainty about what duties states have to regulate platforms to ensure they protect free speech of users.[7] As fundamental rights are asserted online, the focus of activity has been to protect rights other than free expression — such as privacy[8] — but assertion of platform regulation will inevitably conflict with speech rights. The United Kingdom's *Online Harms White Paper* has been criticized, for example, for undermining the European Court of Human Rights principle that restrictions on free expression need to be "enshrined in law" (and therefore subject to parliamentary and public scrutiny) rather than hidden somewhere in standards developed in the shadowy space between platforms and government.[9] There will be a need for clear, positive assertion of rights and obligations of users, platforms and governments, and these are not provided by current standards.

Assertion of platform regulation will inevitably conflict with speech rights.

The need to bring together disparate policy instruments into a coherent overall framework and regulatory architecture: In the United Kingdom, and other European countries, internet intermediaries face parallel policy processes in the areas of taxation, competition and content regulation. It is inevitable that these discussions will be brought together on grounds of regulatory efficiency. There are concrete attempts at the national and European level to redefine competition policy: there is a consensus that competition law has failed in the light of platform business models to check new forms of market power, such as dominance of data in complex multi-sided markets. Yet competition proposals, such as the Furman Review,[10] and current proposals for fiscal reform, such as the digital services tax, have no social component: they do not attempt to use fiscal policy to alter incentives for companies to mitigate negative social externalities. Although it is true that there may be freedom of expression concerns with such a foray into the speech field, sooner or later the role of these fiscal levers will come into play, particularly if the public purse benefits from the business models of the tech giants, while doing nothing to shape their behaviour. In the old world of broadcasting, specialist regulators such as the UK Office of Communications dealt with matters both of competition and social regulation. It is likely that regulators — perhaps the same ones — will do the same with social media platforms.

Fiscal, market structure and competition issues entangled with issues of fundamental rights, media accountability and responsibility: It would be hard to sustain the claim that a small social media platform prohibiting or removing a post is a "censor" of speech or that, with its handful of subscribers, it could have a profound effect on the right to impart and receive ideas in the way that Facebook or Google would. A political viewpoint or an artistic expression that breached Facebook's guidelines on extremism and hate, or even one that failed to meet a threshold for "trusted" or "quality" journalism on YouTube, could be consigned to the sidelines or even silenced. So, competition law and regulation, as they have done in the past in relation to media pluralism, will in some sense combine general social interests and matters of special public interest in new ways through behavioural competition policy and also merger rules. Law makers can have a profound influence on how the digital media business model develops. The key question is the division of labour between European and nation-state regulators.

The limitations of platform regulation in preventing online harms will also have to be faced.

The general problem with this flurry of policy activity is, therefore, fragmentation — at the national level between fiscal, competition and content regulation, and at the European level between separate complex nation-state-level regulatory policies, which could in time create problems for the operation of the single market. Clearly, there will be pressure to simplify and standardize across Europe: there are strong economic imperatives for a pan-European solution that would reduce platform costs in national tailoring and governance of services. Whatever the Brexit outcome, there will likely be strong imperatives for the United Kingdom to align.

Toward a New European Settlement on Platforms' Rights and Responsibilities

It is unfortunate that such a key challenge for European governance coincided with the dual crises of the euro zone and Brexit. The response will have to be pragmatic, addressing the challenges of the current impasse with platform regulation and developing careful and creative answers for the challenges of multi-levelled governance. The European Commission's recommendation on tackling illegal content online,[11] which permits member states to require a "duty of care" by platforms on the model of the UK proposals, will need to be reviewed along with the electronic commerce directive, which offers platforms a holiday from liability, and a much more solid evidence base about consumer harms gathered at both the pan-European and member-state level. What is being attempted is a gradual "responsabilization" of platforms: by changing the incentive structures, regulation and liability will seek to encourage a new ethics of responsible platforms, which can provide certainty, fairness and accountability of enforcement of speech rules, but ensure that speech control is strictly autonomous from the state. Policy makers will use a broad tool kit to achieve the right incentives: tax breaks, distribution rules and competition policy, as well as regulation.

There are numerous challenges in this multi-levelled policy game. The first is the power of the actors involved: US companies have funded a huge lobbying exercise that attempts to deploy new notions of internet freedom to stymie calls for accountability, often using the opinion-shaping power and public reach of the platforms themselves to oppose regulation. One strong factor in the favour of progressive reform, however, is that vested interests in the legacy media can be deployed in support of policy reform. Legal reforms in Germany and the United Kingdom have enjoyed firm support from the press.

The limitations of platform regulation in preventing online harms will also have to be faced. The platforms, fortunately or unfortunately, are not the internet. Attempts to regulate them may lead to a tendency for content to migrate to other, less-regulable places on the internet. This will ensure the whole process is disciplined

by consumer power. If consumers feel the rules are inappropriate or illegitimate, and if competition policy successfully creates the opportunity to switch, they will vote with their clicks and go elsewhere.

In the next policy cycle, it is inevitable that the current framework for intermediary liability will come up for scrutiny, and we may see a new understanding whereby governments and regulators are more active in monitoring self-regulation by social media and other intermediaries. But in an environment characterized by an understandable lack of faith in democratic institutions, the argument for subsidiarity in media accountability is overwhelming. Censorship, whether through the targeting or subsidy of speech by law or the opaque fiat of powerful private actors, would not be trusted. Whatever the eventual architecture of speech control, whatever the eventual settlement on the rights and duties of platforms, it must be rooted in civil society and legitimate in the eyes of users.

NOTES

1 In a series of antitrust cases since 2017, the European Commission has fined Google more than eight billion euros for abuse of dominance in the market for online advertising. See EC, Press Release, IP/19/1770, "Antitrust: Commission fines Google €1.49 billion for abusive practices in online advertising" (20 March 2019), online: *EC Press Release Database* <https://europa.eu/rapid/press-release_IP-19-1770_en.htm>.

2 The NetzDG (Netzwerkdurchsetzungsgesetz, or Network Enforcement Act) was passed in Germany in 2017 and came fully into force in January 2018. See *Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken*, BGBl IS 3352, 1 September 2017. It provides for enhanced standards to encourage platforms to remove illegal hate speech and fines if they do not. In France, Law No. 2018-1202 of December 22, 2018, on combatting manipulation of information, came into force in 2019 and provides for courts to apply fines for deliberate disinformation during election campaigns. See *LOI n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information*, JO, 23 December 2018, no 2. A more comprehensive legislative proposal was published in 2019. The UK government published proposals for a new regulatory framework on online harms in 2019: UK, Secretary of State for Digital, Culture, Media & Sport and Secretary of State for the Home Department, *Online Harms White Paper* (CP 57, 2019) at 41, online: <www.gov.uk/government/consultations/online-harms-white-paper>.

3 EC, *Commission Recommendation (EU) 2018/334 of 1 March 2018 on measures to effectively tackle illegal content online*, [2018] OJ, L 63/50.

4 EC, *Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services in particular electronic commerce, in the Internal Market ("Directive on electronic commerce")*, [2000] OJ, L 178/1.

5 EC, *Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) (Text with EEA relevance)*, [2010] OJ, L 95/1.

6 See the Joint Declaration on Freedom of Expression on the Internet, online: <www.osce.org/fom/78309?download=tru.e>.

7 This is the question under the European Convention on Human Rights of whether platforms have a positive duty to promote free speech. See UN Human Rights Committee, *General comment No 34 on Article 19: Freedoms of opinion and expression*, UN Doc CCPR/C/GC/34, 12 September 2011 at para 7; see also Dominika Bychawska Siniarska, *Protecting the right to freedom of expression under the European Convention on Human Rights — A handbook for legal practitioners* (Strasbourg, France: Council of Europe, 2018).

8 Council of Europe, Committee of Ministers, *Recommendation CM/Rec(2012)4 of the Committee of Ministers to member States on the protection of human rights with regard to social networking services, 4 April 2012*, online: *Council of Europe* <https://go.coe.int/lL7Uv>.

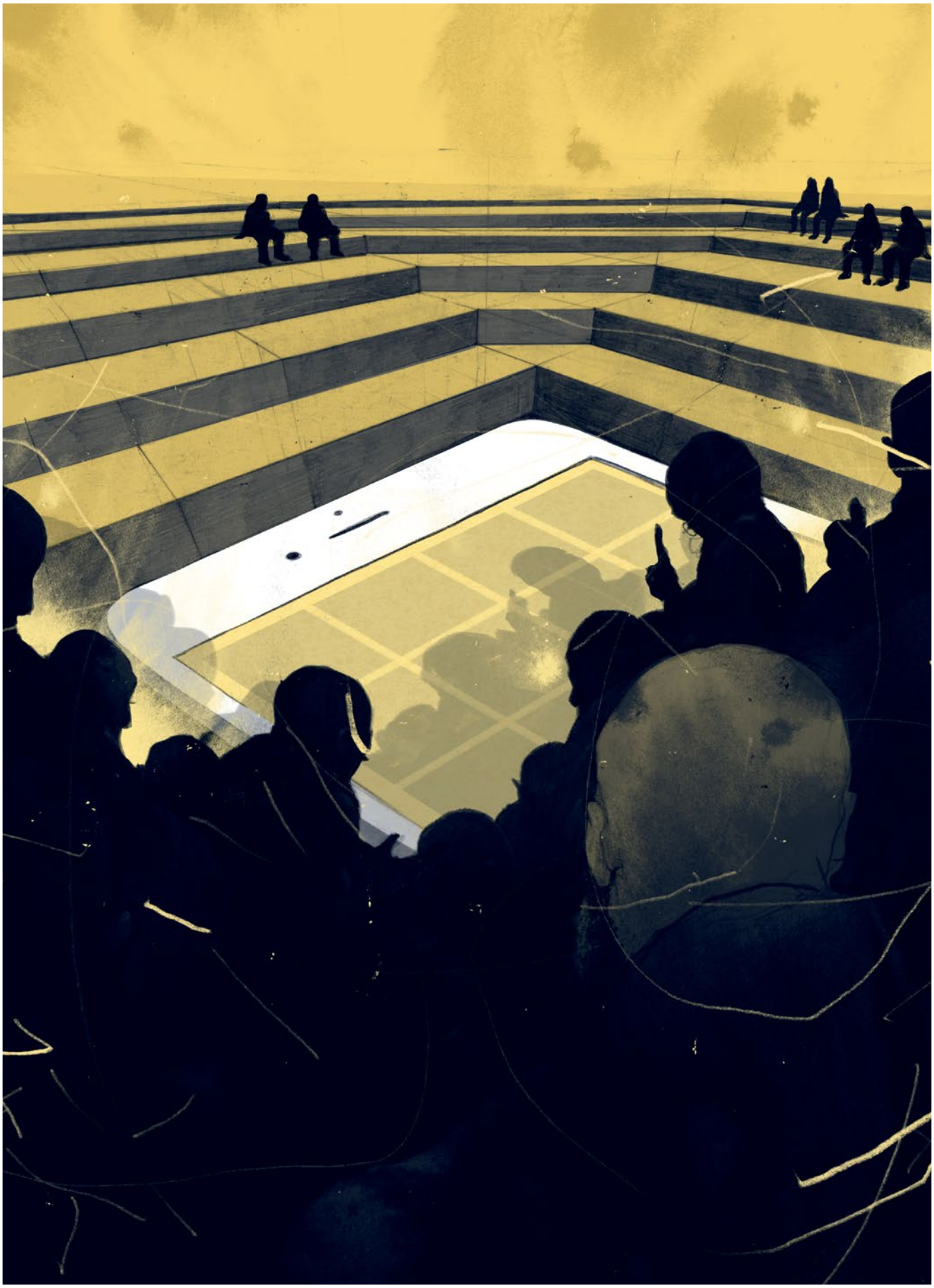
9 Damian Tambini, "Regulating Online Harms through a Differentiated Duty of Care: A Response to the Online Harms White Paper" (2019), online (pdf): *Foundation for Law, Justice and Society* <www.fljs.org/search-publications>.

10 UK, Digital Competition Expert Panel, *Unlocking digital competition: Report of the Digital Competition Expert Panel* (PU2242) (London: HM Treasury, March 2019), online: <www.gov.uk/government/publications/unlocking-digital-competition-report-of-the-digital-competition-expert-panel>.

11 *Supra* note 3.

ABOUT THE AUTHOR

Damian Tambini is an academic researcher specializing in media and communications policy and law. He is currently associate professor at the London School of Economics (LSE), and also associate fellow at the Institute for Public Policy Research and the Oxford Internet Institute. He is also a fellow of the Royal Society of Arts and founder of the LSE Media Policy Project, and the Oxford Media Convention. He has acted as an adviser to the Government of the United Kingdom and to the Council of Europe and the European Commission. From 2002 to 2006, he served as head of the Programme in Comparative Media Law and Policy at Oxford University. Before that he was at Nuffield College, Oxford (post-doctoral fellow, 1998); Humboldt University, Berlin (lecturer, 1997); and the European University Institute, Florence, Italy (Ph.D., 1996). His recent publications include numerous articles on media freedom and regulation and the power of the tech giants.



Heidi Tworek

Social Media Councils

In May 2019, shortly before the International Grand Committee on Big Data, Privacy and Democracy met in Ottawa, a doctored video of US Democratic Speaker of the House Nancy Pelosi went viral. The video had slowed Pelosi's words, making her appear intoxicated. Later, the Facebook representatives testifying before the committee, Neil Potts and Kevin Chan, were asked about their company's policies on altered material (House of Commons 2019). They responded that Facebook would not remove the Pelosi video, but would label it to ensure that users knew it was doctored. (Journalists later found postings of the video without the label.) After further probing from British MP Damian Collins, Potts clarified that Facebook would remove pages masquerading as a politician's official page. Altered videos could stay; pages impersonating a politician could not.

Questioning of social media representatives continued, in a marathon session lasting three hours. Eventually, the floor was handed to the Speaker of the House of Assembly from St. Lucia, MP Andy Daniel. He had found Facebook's answers intriguing: for some time, the prime minister of St. Lucia had been trying to get Facebook to remove a page impersonating him, but Facebook had not responded. At the hearing, Potts and Chan promised to look into the case; it is not clear if they did.

The episode illustrates many of the problems with our current content moderation landscape. First, social media companies' processes are often opaque, their rules arbitrary and their decisions lacking in due process.

Second, companies have global terms of service, but do they enforce them evenly around the world? Or do they pay more attention to larger markets such as the United States and Germany, where they are worried about potential legislative action?

Third, the episode showed a lack of accountability from top employees at major social media companies. Mark Zuckerberg has testified before the US Congress and the European Parliament. Zuckerberg even returned to Congress in September 2019 to meet with senators and US President Donald Trump (Lima 2019). But he seems determined not to testify anywhere else. Between testimonies to Congress, Zuckerberg has declined to testify in the United Kingdom and Canada, despite subpoenas.

The Central Challenge of Content Moderation

It has been more than a decade since internet companies cooperated with human rights and civil-liberties organizations to create the Global Network Initiative (GNI), a set of voluntary standards and principles to protect free expression and privacy around the world. Since the GNI's creation in 2008, questions have only grown more urgent about how to coordinate standards and whether voluntary self-regulation suffices. The discussion has also shifted to see content moderation as one of the central challenges for social media. Content moderation has become a key area of regulatory contestation, as well as concern about how to ensure freedom of expression.

Content moderation is already subject to various forms of regulation, whether the Network Enforcement Law (NetzDG) in Germany (Tworek 2019a), self-regulatory company-specific models, such as the oversight board under development at Facebook, or

Social media companies' processes are often opaque, their rules arbitrary and their decisions lacking in due process.

Removing a page impersonating a politician is fairly straightforward, yet this simple question raised a host of complicated issues. It also showed that one hearing barely scratched the surface. Clearly, we need institutions to provide regular, consistent meetings to answer pressing questions, discuss content moderation standards and push for further transparency from social media companies.

This essay examines one idea to improve platform governance — social media councils. These councils would be a new type of institution, a forum to bring together platforms and civil society, potentially with government oversight. Social media councils would not solve the structural problems in the platforms' business models, but they could offer an interim solution, an immediate way to start addressing the pressing problems in content moderation we face now.

subject-specific voluntary bodies such as the Global Internet Forum to Counter Terrorism (GIFCT).

The Facebook model is a novel example of independent oversight for content moderation decisions. In 2018, Mark Zuckerberg announced that he wanted to create "an independent appeal [body]" that would function "almost like a Supreme Court" (Klein 2018). After consultations in many countries, Facebook announced in mid-September 2019 the charter and governance structure of their independent oversight board (Facebook 2019a). A trust will operate the board to create independence from the company. It also seems possible that Facebook may open up the board to other companies in the future.

Some legal scholars, Evelyn Douek (2019) for example, have praised Facebook's plans as a

step in the right direction, because the board will at least externalize some decisions about speech. Yet, the board foresees having only 40 members, who will have to deal with cases from around the world. It is inevitable that much context will be lost. We do not know which cases will be chosen or how quickly decisions will be made. Decisions that take a long time to make — as they do at the Supreme Court — could be made too late for the Rohingya whose violent expulsion from Myanmar was accelerated through hate-filled Facebook posts. Facebook plans to have their board running by early 2020. Even if it were to function seamlessly, questions remain about whether models based on one company can work or if there should be more industry-wide supervision.

Unlike the single-company model of Facebook's oversight board, the GIFCT enables coordination between major social media companies (Facebook, Twitter, YouTube and Microsoft) who want to remove terrorist content. Since its emergence in summer 2017, the GIFCT has facilitated a form of industry-wide cooperation, but without public oversight. The GIFCT published its first transparency report in July 2019 (GIFCT 2019). Among other things, the GIFCT houses a "common industry database of 'hashes' — unique digital fingerprints — for violent terrorist imagery or terrorist recruitment videos or images that the member companies have removed from their services" (Heller 2019, 2), although it is unclear exactly how the GIFCT defines terrorism. For now, the GIFCT remains mostly a mystery to those outside the companies or specific civil society organizations and governments who cooperate with the forum. To give a few examples, we do not even know where the database is housed. The GIFCT has no provision for third-party researcher access. We do not know if additions to the database by one company are ever disputed by another or if there are even mechanisms to resolve such a dispute.

A More Inclusive Solution

Social media councils would take a different approach than the GIFCT or Facebook's oversight board. They would be multi-stakeholder fora, convened regularly to address content moderation and other challenges. Civil society organizations would participate from the start. Ironically, although social media promised to connect everyone, platforms' users have far too often been excluded from regulatory conversations. The Christchurch Call, for example, was initially only signed by companies and governments.[1]

Social media councils would include representatives from marginalized groups and civil society organizations across the political spectrum to ensure that any solutions incorporate the experiences and suggestions of those people often most affected by content moderation decisions. The general idea is supported by others, including the UN Special Rapporteur on the Right to Freedom of Opinion and Expression (Global Digital Policy Incubator, ARTICLE 19 and David Kaye 2019). By mandating regular meetings and information-sharing, social media councils could become robust institutions to address problems that we have not yet even imagined.

The exact format and geographical scope of social media councils remain up for debate. One free speech organization, ARTICLE 19, currently has an open consultation about social media councils; the consultation will close on November 30, 2019.[2] The survey includes questions about councils' geographical scope (international or national) and the remit (whether the council would review individual cases or decide general principles).

If social media councils do emerge, there are ways of combining the national and the international. A national social media council might operate on the basis of an international human rights framework, for example. National social media councils could be set up simultaneously and designed to coordinate with each other. Or, an international council could exist with smaller regional or national chapters. If these councils address content, it is worth remembering that international cooperation will be difficult to achieve (Tworek 2019b).

The regulatory structure of social media councils is another foundational question. Governance need not always involve government intervention. Broadly speaking, there are four different models of regulation with different levels of government involvement, as explained in a paper on classifying media content from the Australian Law Reform Commission (2011, 17):

- **Self-regulation:** Organizations or individuals within an industry establish codes of conduct or voluntary standards. Self-regulation may be spurred by internal decisions within the industry, public pressure for change or concern about imminent government action. Self-regulating associations are generally funded by membership fees. The industry itself enforces rules and imposes punishments for violations.

- **Quasi-regulation:** Governments push businesses to convene and establish industry-wide rules and compliance mechanisms. Governments do not, however, determine the nature of those rules nor their enforcement.

- **Co-regulation:** An industry creates and administers its own rules and standards. However, governments establish legislation for enforcing those arrangements. Broadcasting is a classic example. The online advertising industry in Israel attempted to self-regulate and failed; although the industry wished to self-regulate, the public wanted to see co-regulation (Ginosar 2014).

- **Government or statutory regulation:** Governments create legislation for an industry and implement a framework and its rules. Governments may also implement international regulation on a national level. Enforcement mechanisms may include councils or arm's-length regulatory bodies, such as data protection authorities.

Some regulatory forms can be hybrid and incorporate elements from several of these categories. There could also be differentiated membership and expectations of council members, based upon the size of the social media company, with higher expectations placed on the major players.

Social media councils could fulfill a wide range of functions, including all or some of those listed here:

- **Give a voice** to those who struggle to gain attention from social media companies. Social media have enabled underrepresented voices to be heard, but those voices are often left out of conversations about how to improve social media. Social media council meetings could be a place to hear from those who have been abused or affected by online harassment. Or, it could be a place to hear from those who have had difficulties appealing against takedowns.

- **Provide a regular forum** to discuss standards and share best practices. These discussions need not standardize content moderation between companies. Different content moderation practices are one way of providing users with choices about where they would like to interact socially online. But sharing best practices could help companies to learn from each other. Might Twitter and Facebook have learned something about how to combat anti-vaxxer disinformation from Pinterest, for example (Wilson 2019)?

- **Create documented appeals processes.**

- **Create codes of conduct**, avoiding the issues that have plagued the European Union's Code for Countering Illegal Hate Speech Online, when civil society organizations walked out over frustration at the process (Bukovská 2019).

- **Establish frameworks to enable access** to third-party researchers, including protocols to protect users' privacy and companies' proprietary information.

- **Develop standards and share best practices** around transparency and accountability, especially in the use of artificial intelligence (AI) and algorithms to detect "problematic" speech.

- **Establish industry-wide ethics boards** on issues such as AI. This might avoid public relations debacles such as the short-lived Google AI Ethics Board (Piper 2019).

- **Discuss labour issues around content moderators.**

- **Address issues of jurisdiction and international standards.**

- **Establish fines and other powers** for government or other enforcement agencies to support the council's authority and decisions.

A social media council may be more or less proactive in different places. That could depend upon a country or region's own cultural policies. Canada, for example, might like to ensure the council examines the implications of the Multiculturalism Act for content moderation (Tenove, Tworek and McKelvey 2018).

Regardless of scope, there should be coordination on international principles for social media councils. International human rights standards and emerging international best practices would provide a helpful baseline and smooth coordination among countries and social media companies. One source of standards is the "Santa Clara Principles" on content moderation.[3] A more established source is international human rights law.

In 2018, the UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, David Kaye, suggested a framework for applying international human rights to content moderation. "A human rights framework enables forceful normative responses against undue State restrictions," he proposed, so long as companies "establish principles of due diligence, transparency, accountability and remediation that limit platform interference with human rights" (UN General Assembly 2018, para. 42). Social media councils can push companies to establish those principles.

When companies try to have global terms of service, international human rights law on speech makes sense as a starting point. Facebook's charter for its new oversight board now explicitly mentions human rights (Facebook 2019b). One clear basis in international speech law is Article 19 of the International Covenant on Civil and Political Rights (ICCPR) (UN General Assembly 1966). Created in 1966, the ICCPR was ratified slowly by some countries and entered into force in 1976. The United States only ratified the Covenant in 1992 (although with some reservations).

Although individual countries differ in their approaches to speech law, international covenants such as the ICCPR provide common ground for transatlantic and transnational cooperation on freedom of expression (Heller and van Hoboken 2019). The UN Human Rights Committee functions as an oversight body for the ICCPR.

A Possible Path Forward

Creating an industry-wide regulatory body is "easier said than done" (Ash, Gorwa and Metaxa 2019). There are many obvious questions. What is "the industry"? What would be the geographical scope of such a body? How could it be structured to balance national, regional and international concerns? How should it ensure freedom of expression and enshrine compliance with international human rights law, while allowing relevant services to operate in different jurisdictions with different legal standards? How could social media councils fit into dispute resolution mechanisms? There are multiple civil society organizations such as ARTICLE 19 and Stanford's Global Digital Policy Incubator working on these questions. Many of them are represented on the High-Level Transatlantic Working Group on Content Moderation and Freedom of Expression Online, which will be exploring such dispute resolution mechanisms over the next few months and considering how transatlantic cooperation might function in this space.

They could put the societal back into social media.

Social media councils may not be a panacea, but they are one possible path forward. Further discussions will help to clarify the jurisdictional and institutional parameters of such a council. The specific set-up may differ for each state, depending upon its communications policies and extant institutions. There are also questions about how to coordinate councils internationally or whether to create cross-border councils with participation from willing countries, such as those participating in the International Grand Committee on Big Data, Privacy and Democracy.

There are issues that councils do not resolve, such as tax policy or competition. Still, they offer a potential solution to many urgent problems. They could put the societal back into social media. They could establish fair, reliable, transparent and non-arbitrary standards for content moderation. At a time when decisions by social media companies increasingly structure our speech, councils could offer a comparatively swift method to coordinate and address pressing problems of democratic accountability. Creating a democratic, equitable and accountable system of platform governance will take time. Councils can be part of the solution.

NOTES

1 See www.christchurchcall.com/call.html.

2 See www.article19.org/resources/social-media-councils-consultation/ for more information and to access the survey.

3 See https://newamericadotorg.s3.amazonaws.com/documents/Santa_Clara_Principles.pdf.

WORKS CITED

Ash, Timothy Garton, Robert Gorwa and Danaë Metaxa. 2019. *GLASNOST! Nine ways Facebook can make itself a better forum for free speech and democracy*. An Oxford-Stanford Report. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-01/Garton_Ash_et_al_Facebook_report_FINAL_0.pdf.

Australian Law Reform Commission. 2011. "Designing a regulatory framework." In *National Classification Scheme Review Issues Paper 40*, 17–32. Sydney, Australia: Australian Law Reform Commission. www.alrc.gov.au/wp-content/uploads/2019/08/IP-40-Whole-pdf.pdf.

Bukovská, Barbora. 2019. "The European Commission's Code of Conduct for Countering Illegal Hate Speech Online: An analysis of freedom of expression implications." Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression Series, May 7. www.ivir.nl/publicaties/download/Bukovska.pdf.

Douek, Evelyn. 2019. "Finally, Facebook Put *Someone* in Charge." *The Atlantic*, September 18. www.theatlantic.com/ideas/archive/2019/09/facebook-outsources-tough-decisions-speech/598249/.

Facebook. 2019a. "Establishing Structure and Governance for an Independent Oversight Board." Facebook Newsroom, September 17. https://newsroom.fb.com/news/2019/09/oversight-board-structure/.

———. 2019b. "Oversight Board Charter." Facebook Newsroom, September 19. https://fbnewsroomus.files.wordpress.com/2019/09/oversight_board_charter.pdf.

GIFCT. 2019. "Transparency Report." https://gifct.org/transparency/.

Ginosar, Avshalom. 2014. "Self-Regulation of Online Advertising: A Lesson from a Failure." *Policy & Internet* 6 (3): 296–314.

Global Digital Policy Incubator, ARTICLE 19 and David Kaye. 2019. *Social Media Councils: From Concept to Reality*. Conference report, February. Stanford, CA: Global Digital Policy Incubator. https://cyber.fsi.stanford.edu/gdpi/content/social-media-councils-concept-reality-conference-report.

Heller, Brittan. 2019. "Combating Terrorist-Related Content Through AI and Information Sharing." Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression Series, April 26. www.ivir.nl/publicaties/download/Hash_sharing_Heller_April_2019.pdf.

Heller, Brittan and Joris van Hoboken. 2019. "Freedom of Expression: A Comparative Summary of United States and European Law." Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression Series, May 3. www.ivir.nl/publicaties/download/TWG_Freedom_of_Expression.pdf.

House of Commons. 2019. "Meeting No. 153 ETHI — Standing Committee on Access to Information, Privacy and Ethics." Televised proceedings, Tuesday, May 28. http://parlvucloud.parl.gc.ca/Harmony/en/PowerBrowser/PowerBrowserV2/20190528/-1/31627?Language=English&Stream=Video.

Klein, Ezra. 2018. "Mark Zuckerberg on Facebook's hardest year, and what comes next." *Vox*, April 2. www.vox.com/2018/4/2/17185052/mark-zuckerberg-facebook-interview-fake-news-bots-cambridge.

Lima, Cristiano. 2019. "Zuckerberg meets with Trump, Republican senators." *Politico*, September 19. www.politico.com/story/2019/09/19/facebook-mark-zuckerberg-congress-meetings-1505445.

Piper, Kelsey. 2019. "Exclusive: Google cancels AI ethics board in response to outcry." *Vox*, April 4. www.vox.com/future-perfect/2019/4/4/18295933/google-cancels-ai-ethics-board.

Tenove, Chris, Heidi J. S. Tworek and Fenwick McKelvey. 2018. *Poisoning Democracy: How Canada Can Address Harmful Speech Online*. Ottawa, ON: Public Policy Forum. https://ppforum.ca/wp-content/uploads/2018/11/PoisoningDemocracy-PPF-1.pdf.

Tworek, Heidi. 2019a. "An Analysis of Germany's NetzDG Law." Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression Series, April 15. www.ivir.nl/publicaties/download/NetzDG_Tworek_Leerssen_April_2019.pdf.

———. 2019b. "Looking to History for Lessons on Platform Governance." Cigionline, July 31. www.cigionline.org/articles/looking-history-lessons-platform-governance.

UN General Assembly. 1966. *International Covenant on Civil and Political Rights*. December 16. www.ohchr.org/en/professionalinterest/pages/ccpr.aspx.

———. 2018. *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*. A/HRC/38/35, April 6. https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/35.

Wilson, Mark. 2019. "The tech giant fighting anti-vaxxers isn't Twitter or Facebook. It's Pinterest." *Fast Company*, February 26. www.fastcompany.com/90310970/the-tech-giant-fighting-anti-vaxxers-isnt-twitter-or-facebook-its-pinterest.

ABOUT THE AUTHOR

Heidi Tworek is assistant professor of international history at the University of British Columbia (UBC), Vancouver, Canada. She works on media, international organizations and transatlantic relations. She is a member of the Science and Technology Studies program, the Language Science Initiative, and the Institute for European Studies at UBC. She is a visiting fellow at the Joint Center for History and Economics at Harvard University as well as a non-resident fellow at the German Marshall Fund of the United States and the Canadian Global Affairs Institute.

Centre for International Governance Innovation

About CIGI

We are the Centre for International Governance Innovation: an independent, non-partisan think tank with an objective and uniquely global perspective. Our research, opinions and public voice make a difference in today's world by bringing clarity and innovative thinking to global policy making. By working across disciplines and in partnership with the best peers and experts, we are the benchmark for influential research and trusted analysis.

Our research initiatives focus on governance of the global economy, global security and politics, and international law in collaboration with a range of strategic partners and have received support from the Government of Canada, the Government of Ontario, as well as founder Jim Balsillie.

À propos du CIGI

Au Centre pour l'innovation dans la gouvernance internationale (CIGI), nous formons un groupe de réflexion indépendant et non partisan doté d'un point de vue objectif et unique de portée mondiale. Nos recherches, nos avis et nos interventions publiques ont des effets réels sur le monde d'aujourd'hui car ils apportent de la clarté et une réflexion novatrice pour l'élaboration des politiques à l'échelle internationale. En raison des travaux accomplis en collaboration et en partenariat avec des pairs et des spécialistes interdisciplinaires des plus compétents, nous sommes devenus une référence grâce à l'influence de nos recherches et à la fiabilité de nos analyses.

Nos projets de recherche ont trait à la gouvernance dans les domaines suivants : l'économie mondiale, la sécurité et les politiques internationales, et le droit international. Nous comptons sur la collaboration de nombreux partenaires stratégiques et avons reçu le soutien des gouvernements du Canada et de l'Ontario ainsi que du fondateur du CIGI, Jim Balsillie.

cigionline.org

The debate about digital technology's role in society has gone through a remarkable transformation. Following two decades of techno-optimism, whereby digital technology — social media, in particular — was left to be governed in a laissez-faire environment, we are now in the midst of a "techlash."

Google, Facebook and Amazon serve billions of users around the globe and increasingly perform core functions in society. The private gains are clear to see — these are among the most profitable companies in history. But in spite of myriad benefits offered by platforms, the costs are clear as well: platforms threaten our social fabric, our economy and our democracy.

Despite growing calls for global platform governance, no solution has been found. To begin to address this, CIGI has convened leading thinkers to explore new models for governing digital platforms. Given platforms' unprecedented influence on democracy and the global economy alike, a cohesive framework for platform governance is crucial.

cigionline.org/platforms