Democratic Transparency in the Platform Society

Robert Gorwa and Timothy Garton Ash*

Draft chapter to be published in the forthcoming *Social Media and Democracy: The State of the Field* (Cambridge), edited by Nate Persily and Josh Tucker, in press 2020

Abstract

Following an host of major scandals, transparency has emerged in recent years as one of the leading accountability mechanisms through which the companies operating global platforms for user-generated content have attempted to regain the trust of the public, politicians, and regulatory authorities. Ranging from Facebook's efforts to partner with academics and create a reputable mechanism for third party data access and independent research to the expanded advertising disclosure tools being built for elections around the world, transparency is playing a major role in current governance debates around free expression, social media, and democracy. This article thus seeks to (a) contextualize the recent implementation of transparency as enacted by platform companies with an overview of the ample relevant literature on digital transparency in both theory and practice; (b) consider the potential positive governance impacts of transparency as a form of accountability in the current political moment; and (c) reflect upon the potential shortfalls of transparency that should be considered by legislators, academics, and funding bodies weighing the relative benefits of policy or research dealing with transparency in this area.

*RG was a visiting scholar with the Project on Internet and Democracy at Stanford University while writing this chapter, and is a DPhil student at St. Antony's College, University of Oxford. TGA is Isaiah Berlin Professorial Fellow at St. Antony's and a Senior Fellow at the Hoover Institution. The authors are indebted to Nate Persily, Rob Reich, Eloise Duvalier, the SSRC, and the Project on Internet and Democracy for supporting this work, and to Luke Heemsbergen for providing very helpful comments.

Contents

1	Introduction	2
2	From Bentham to Blockchain: The Historical Evolution of the Transparency Ideal	4
2.1	Transparency and the Corporation	6
3	Platforms and Transparency in Practice	9
3.1	Voluntary Transparency for Content Takedowns	10
3.2	Voluntary Transparency for Content and Advertisements	11
3.3	Mandated Transparency Regimes	15
3.4	Third Party “Audits” and Investigations	16
4	The Future of Platform Transparency	16
4.1	The Facebook Oversight Board	18
5	Conclusion	19
6	References	20

1 Introduction

At a time when social media and social media companies have played a complex, troubled role in recent political events, ranging from the 2016 US election and Brexit to genocide in Myanmar, a growing chorus of scholars, policymakers, and commentators have begun proclaiming that social media — less than ten years ago cast as an emancipatory “liberation technology” — may now be fomenting polarization and undermining democracy around the globe (Tucker et al. 2017). As the executives of Google, Facebook, Twitter, and other major platform companies have been called to testify before elected officials in North America, Europe, and Asia with policymakers discussing various regulatory measures to reign in the so-called “digital giants” (Moore and Tambini 2018), the public is increasingly demanding greater accountability from the technology companies that operate the services entrusted with their sensitive data, personal communication, and attention. In the past several years, transparency has emerged as one of the leading accountability mechanisms through which platform companies have attempted to regain the trust of the public, politicians, and regulatory authorities. Ranging from Facebook’s efforts to partner with academics and create a reputable mechanism for third party data access and independent research (King and Persily 2019) and the recent release of Facebook’s public-facing “Community Standards” (the rules that govern what the more than 2.2 billion monthly active users of Facebook are allowed to post on the site), to the expanded advertising disclosure tools being built for elections around the world (Leerssen et al. 2019), transparency is playing a major role in current policy debates around free expression, social media, and democracy.

While transparency may seem intuitive as a high-level concept (often conceived narrowly as the disclosure of certain information that may not previously have been visible or publicly available, see Albu and Flyverbom 2016), critical scholarship has long noted that a major reason for the widespread popularity of transparency as a form of accountability in democratic governance is its flexibility and ambiguity. As the governance scholar Christopher Hood has argued, “much of the allure of transparency as a word and a doctrine may lie in its potential to appeal to those with very different, indeed contradictory, attitudes and worldviews” (Hood 2006, 19). Transparency in practice is deeply political, contested, and oftentimes problematic (Etzioni 2010; Ananny and Crawford 2018); and yet, it remains an important — albeit imperfect — tool which, in certain policy domains, has the potential to remedy unjust outcomes, increase the public accountability of powerful actors, and improve governance more generally (Fung, Graham, and Weil 2007; Hood and Heald 2006). As the authors of the Ranking Digital Rights report (an annual effort to quantify the transparency and openness of multiple major technology companies) argue, “Transparency is essential in order for people to even know when users’ freedom of expression or privacy rights are violated either directly by — or indirectly through — companies’ platforms and services, let alone identify who should be held responsible.” (Ranking Digital Rights 2018, 4–5).

In recent years, a substantial literature has developed in what might be called “digital transparency studies,” much of which critiques prevailing discourses of technologically-implemented transparency as a panacea for the digital age (Flyverbom 2015; Hansen and Flyverbom 2015; Albu and Flyverbom 2016; Stohl, Stohl, and Leonardi 2016; Christensen and Cheney 2014). This work has been wide-reaching, touching on the politics of whistleblowing and leaking (Hood 2011), the intricacies of transparency can be politicized or weaponized to fulfill specific policy agendas (Levy and Johns 2016), and the dynamics of transparency as they have been mediated by technological developments and encryption (Heemsbergen 2016). But there has been less work that surveys the scholarship on transparency more broadly (drawing upon organizational studies, political science and governance studies, law, as well as digital media and political communication research; see Flyverbom 2019) and applies it narrowly to the pressing questions around platform companies, social media, and democracy that are the focus of growing public and scholarly discussion, as well as this volume. The goal of this chapter is therefore to contextualize the recent examples of transparency as implemented by platform companies with an overview of the relevant literature on transparency in both theory and practice, consider the potential positive governance impacts of transparency as a form of accountability in the current political moment, and reflect upon the shortfalls of transparency that should be considered by legislators, academics, and funders weighing the relative benefits of policy or research dealing with transparency in this area.

The chapter proceeds in three parts: we first provide a rapid historical overview of transparency in democratic governance, tracing it from its origins in Enlightenment-era liberalism all the way to its widespread adoption in the 20th and 21st centuries. We quickly summarize what transparency seeks to achieve in

theory, and how it is commonly conceptualized as furthering traditional democratic values. In the second section, we provide a necessarily not comprehensive survey of major transparency initiatives as enacted by platform companies in the social media era. Finally, we discuss transparency in practice, with a summary of critical insights from the “digital turn” in transparency studies, which provides ample caution against transparency as simply promoting democratic accountability and good governance.

2 From Bentham to Blockchain: The Historical Evolution of the Transparency Ideal

Depending on how one precisely formulates what transparency is and what its key elements are, the origins of the concept can be traced back to various classical Chinese and Greek ideas about government and governance. Researchers have suggested that there are multiple, inter-related strains of pre-20th century thinking that have substantially inspired the contemporary notions of transparency, including longstanding “notions of rule-governed administration, candid and open social communication, and ways of making organization and society ‘knowable’” (Hood 2006, 5), but it is generally accepted that transparency as it is generally understood today became popularized first in the writings of certain major Enlightenment political theorists, especially Jeremy Bentham, Immanuel Kant, and Jean-Jacques Rousseau. Their notions of transparency — often posited as an antonym to secrecy — reflected their presiding views about human nature, politics, and international relations. For Kant, arguing against government secrecy about treaties in his famous essay on “Perpetual Peace,” a lack of transparency could potentially worsen the effects of international anarchy and contribute to war (O’Neill 2006; Bennington 2011). For Rousseau, transparency was primarily a way to increase the visibility of public servants, making it more difficult for them to defraud the state; he called for measures that would make it impossible for government officials to “move about incognito, so that the marks of a man’s rank or position shall accompany him wherever he goes” (Rousseau 1985: 72, in Hood 2006). Bentham’s even more extreme ideals about publicity and visibility, as best exemplified by his infamous “panopticon,” seem to have been fundamentally rooted in pessimistic expectations about human fallibility and the corrupting influence of power (Gaonkar and McCarthy 1994). Bentham is often identified as the forefather of modern transparency as used in the political sense (Hood 2006; Baume 2018). He famously wrote that that “the more strictly we are watched, the better we behave,” an edict that inspired his approach to open government, arguing that “Secrecy, being an instrument of conspiracy, ought never to be the system of a regular government” (Hood 2006, 9).

Even before Bentham’s writings, however, one of the first apparent initiatives for government transparency in practice was underway in Sweden: the “Ordinance on Freedom of Writing and of the Press” (1766), proposed by the

clergyman and parliamentarian Anders Chydenius (Birchall 2011; Lamble 2002), which provided citizens with statutory access to certain government documents. Chydenius, apparently inspired by the Chinese “scholar officials” of the Tang Dynasty “Imperial Censurate,” who investigated government decisions and corrupt officials (Lamble 2002, 3), helped enact what is widely seen to be the precursor to all modern Freedom of Information Access (FOI or FOIA) legislation. While a handful of detailed historical accounts of the adoption of transparency measures as enacted by governments in specific countries exist, such as in the Netherlands (Meijer 2015), it is generally accepted that modern political transparency emerged in the United States centuries after it did in Scandinavia (Hood and Heald 2006). In the early and mid- 20th century, a number of major American political figures, ranging from Woodrow Wilson and Louis Brandeis to Harry Truman and Lyndon Johnson, began publicly arguing that transparency was a moral good and an essential requirement for a healthy, democratic society (Hood and Heald 2006). Wilson, channeling ideas expressed by Kant more than a century earlier, blamed secret treaties for contributing to the outbreak of the First World War, and made diplomatic transparency a significant feature of his famous “14 points” (Hood 2006). Brandeis, a Supreme Court justice and influential political commentator, advocated for even broader forms of transparency in public affairs, famously claiming that “Sunlight is said to be the best of disinfectants” (Brandeis 1913, 10). Brandeis’ ideas would culminate decades later in what the historian Michael Schudson has called the “transparency imperative,” as cultural changes and technological advances resulted in transparency becoming increasingly institutionalized in the United States across a multitude of public and private domains (Schudson 2015, 12).

The move towards today’s much wider embrace of transparency as a facet of contemporary democratic governance began with Truman, who signed the first of a series of important administrative orders, the 1946 Administrative Procedures Act, followed notably by the 1966 Freedom of Information Act, and the 1976 Government in the Sunshine Act (Fung 2013). The 1966 legislation proved to be especially influential, with its main premise (a mechanism for citizens to request the public disclosure of certain government documents) effectively replicated by legislatures in more than 90 countries (Banisar 2006). The goal of these initiatives was to reduce corruption, increase government efficiency by holding officials accountable, and generally promote the public legitimacy of government (Grimmelikhuijsen et al. 2013; Cucciniello, Porumbescu, and Grimmelikhuijsen 2017). Through freedom of information requests — which have continued to expand in their scope since the 1960s, to the extent that “the right to know” has been postulated as a human right (see Schudson 2015) — as well as mandatory or voluntary disclosure programs, auditing regimes (Power 1997), and growing recognition of the importance of whistleblowing (Garton Ash 2016, 337), the concept of transparency became central to 20th century forms of power, visibility, and knowledge (Flyverbom 2016).

Brandeis’ ideas continue to inspire transparency initiatives to this day, primarily among advocates of open government (his famous quote even provided the name for the Sunlight Foundation, a non-governmental organization that

advocates for government transparency). Groups such as Transparency International strive to provide access to information that can combat corruption, potentially unethical behavior, and better document how power is exerted in a democratic society (see for instance, the Sunlight Foundation’s work on campaign finance). But less widely cited is Brandeis’ full quote: “Sunlight is said to be the best of disinfectants; electric light the most efficient policeman. And publicity has already played an important part in the struggle against the Money Trust” (Brandeis 1913, 10). As Kosack and Fung (2014) note, the initial target of Brandeis’ quest for transparency was not corrupt government, but rather opaque corporations: the “Money Trust” of robber barons, bankers, and capitalists who were amassing great fortunes — and potentially defrauding the public — with little public accountability or oversight. In effect, Brandeis was anticipating impending corporate scandals (such as the 1929 Stock Market Crash, which led to an initial measure of corporate transparency and oversight via the formation of the US Securities and Exchange Commission), noting that the market itself provided limited information for consumers:

Then, as now, it was difficult for a grocery shopper to judge the ingredients contained in food products; this difficulty was compounded many times when an investor tried to assess the worth of more complicated products such as financial securities. Government, [Brandeis] thought, should step in to require companies such as food producers and banks to become fully transparent about their products and practices through laws and regulations (Kosack and Fung 2014, 68)

In particular, Brandeis broke with the enlightenment tradition of transparency as primarily a way to hold officials to account, also arguing that it could serve as a check against private power. By doing so, he set the intellectual foundations for the corporate transparency initiatives of the 20th century.

2.1 Transparency and the Corporation

There has been a profound interest in transparency measures since their introduction in the US in the 1960s (Schudson 2015), and today, transparency measures are commonly proposed for a host of private and public actors. Archon Fung has written extensively about the fundamental democratic ideals of transparency underlying these efforts, which he argues are based upon the first principle that “information should be available to citizens so that they can protect their vital interests” (Fung 2013, 185). Crucially, because “democratically important kinds of information may be information about the activities of private and civic organizations rather than governments themselves” (Fung 2013, 188), corporations that play outsized roles in public life should also ideally be as transparent as possible. Nevertheless, corporate transparency and accountability is generally only demanded once a corporation appears to threaten citizen interests — as in the case of the powerful banks that Brandeis was

concerned about, or in the archetypal case of the extractive natural resource company or petrochemical firm (Frynas 2005). As multiple labor and environmental scandals across different industries became public, transparency became increasingly packaged as part of the “Corporate Social Responsibility” (CSR) movement that came to prominence in the 1980s and 90s (Ruggie 2013). For instance, anti-corruption activists noted that good corporate behaviour should go beyond just compliance with legal mechanisms, but also involve active participation in various social responsibility initiatives, and feature as much transparency as possible (Hess 2012).

Transparency was thus advocated as a core element of public accountability for business (Waddock 2004), and enthusiastically announced as a way for firms to demonstrate their social responsibility and integrity (Tapscott and Ticoll, 2003). In practice, however, transparency has many important limitations. Firstly, just like governments, corporations cannot be perfectly transparent, albeit for different reasons than governments — intellectual property is a central concern. Just as with governmental transparency, corporate transparency must achieve a compromise, which lifts “the veil of secrecy just enough to allow for some degree of democratic accountability” (Thompson 1999, 182). Although transparency efforts are often backed by those who oppose regulation, governance scholars have argued that “transparency is merely a form of regulation by other means” (Etzioni 2010, 10) and that transparency alone can be no substitute for regulation. Indeed, meaningful transparency often needs to be backed with regulatory oversight, as scholars critical about corporate transparency in practice emphasize the possible existence of “opaque” forms of transparency which does not actively make the democratically relevant information visible, but rather can be used to obfuscate processes and practices beneath a veneer of respectability (Christensen and Cheney 2014; Albu and Flyverbom 2016).

Overall, the track record of corporate transparency measures for promoting good governance has been mixed. Across multiple domains, from development projects to the private sector, it has been said that “actual evidence on transparency’s impacts on accountability is not as strong as one might expect” (Fox 2007, 664). Corporate actors do not always play along, and may only do the bare minimum without fully implementing voluntary or legislatively mandated transparency measures: as a comprehensive literature review of 25 years of transparency research notes, “the effects of transparency are much less pronounced than conventional wisdom suggests” (Cucciniello, Porumbescu, and Grimmelikhuijsen 2017, 32). Empirical work into the results of transparency initiatives has shown its important limitations: for instance, a survey of mandatory disclosure programs for chemical spills in the United States suggested that the disclosures may have been up to four times lower than they should have (Fox 2007, 665). Measures that run the gamut from audits, inspections, and industry-wide ombudspersons to parliamentary commissions and inquiries are often formulated as mechanisms of “horizontal transparency” that strive to let outsiders (e.g. regulators, the public) see inside the corporate black-box (Hansen and Flyverbom 2015). However, the effects of these efforts can vary significantly, to the extent that “it remains unclear why some transparency initiatives man-

age to influence the behavior of powerful institutions, while others do not” (Fox 2007, 665).

The pessimism about the possibility of successful transparency efforts in both the public and private sectors has been somewhat counterbalanced in the past two decades by increasing optimism about the possibilities of “digital transparency” or “e-transparency” (Bertot, Jaeger, and Grimes 2010). New information and communication technologies have promised not only to make transparency more effective, but also more efficient by increasing the availability of relevant information (Bonsón et al. 2012). Internet utopians like Wired Magazine’s Kevin Kelly began arguing that the use of “digital technology in day-to-day life affords an inevitable and ultimate transparency” (Heemsbergen 2016, 140), and the collaborative, peer-produced “Web 2.0” seemed to provide multiple possibilities for blogs, wikis, online archives, and other tools through which interested stakeholders could access certain forms of democratically important information (Flyverbom 2016). As some hoped that information and communication technologies could start laying the foundations for “a culture of transparency” in countries without a longstanding history of democratic governance (Bertot, Jaeger, and Grimes 2010, 267), Heemsbergen demonstrates how others began proposing even more extreme forms of “radical transparency,” which would deploy “networked digital methods of collecting, processing, and disclosing information” in order to seek maximal social and economic growth.

The latest technological innovations are constantly being harnessed for possible transparency initiatives — the recent spread of projects using distributed ledger systems (e.g., blockchain) to create open registries and databases provides perhaps the best example (Underwood 2016). Bentham would likely have approved of such initiatives: with a blockchain-based registry where each change is documented and permanently encoded in the database itself, theoretical levels of perfect transparency in a specific system can be achieved. From Wikileaks to the encrypted whistleblowing platform SecureDrop, encryption has further helped digital transparency take root, with potentially significant impacts for the redistribution of power between states, citizens, and corporations (Heemsbergen 2016; Owen 2015).

In the past decade, a new generation of technology utopians has seized the ideological foundations set by the enlightenment thinkers, positing “openness” as a organizing principle for contemporary social life. Facebook’s chief executive, Mark Zuckerberg, has preached for many years that his company’s products were creating “radical transparency” at a societal level, fostering more “open and honest communities” (Heemsbergen 2016). Zuckerberg has publically portrayed openness and transparency as key organizing features of the digital age while running a company that effectively made political decisions with global ramifications in secret (Gillespie 2018). However, following the multiple scandals hounding Facebook in the past two years, the mantra is slowly being turned inwards: Zuckerberg has claimed that he will finally bring transparency to some of the companies’ sensitive business dealings, most notably in the realm of political advertising (Feldman 2017). In public discourse, academics, policymakers, and civil society groups are increasingly advocating measures to look into the

corporate black-box of firms like Facebook and Google, positing transparency as a major potential governance mechanism which could rein in platform companies (Brock 2017). But how has transparency historically been enacted by these companies, and what are the recent measures that have been implemented in response to this public outcry?

3 Platforms and Transparency in Practice

A noteworthy feature of the 21st century “platform society” (Van Dijck, Poell, and Waal 2018) is the relationship between, on one hand, the increasingly sophisticated sociopolitical and technical systems that now require transparency due to their political and democratic salience, and the even more technical and complex sociopolitical mechanisms enacted to try and create that transparency on the other. For instance, large-scale algorithmic systems have been in the past few years roundly critiqued for their opacity (Burrell 2016), leading to a growing movement for technical measures that produce “Fairness, Accountability, and Transparency” in machine learning models and data driven systems (Kroll et al. 2016). As research in this area has shown, producing desirable forms of transparency while also preserving privacy and intellectual property in a computationally feasible manner is no easy task (Edwards and Veale 2017; Wachter, Mittelstadt, and Russell 2017). Social media platforms present a similar challenge: companies like Facebook and Google have shown themselves to be highly important venues for political speech and deliberation around the world, and their increasing role as a global channel for news and political information mean that they are now clearly institutions with significant democratic implications (Garton Ash 2016, Gorwa, 2019a; Van Dijck, Poell, and Waal 2018). Their apparent influence, combined with many high-profile scandals, suggest that these firms can pose a threat to the average citizen’s best interests. Today, platform companies clearly meet the threshold articulated in traditional theories of democratic transparency for the types of actors that should require transparency, oversight, regulation, and accountability.

Platform companies, drawing from their ideological roots in historic countercultural and hacker movements (Turner 2009), have come to embody a very distinct flavor of transparency. Internally, companies like Google and Facebook are famous for their founders’ embrace of “openness,” perhaps most clearly manifest in their physical workplace environments. At Facebook’s offices in Menlo Park, cubicles have been eschewed in favor of one of the world’s largest open-plan office spaces, meeting rooms have large glass windows or doors that allow employees to easily observe what goes on inside, and CEO Mark Zuckerberg often works from the center of that open-plan office in a glass ‘fishbowl,’ visible to all (Flyverbom 2016). Google is known for its similarly designed offices and open office culture, including weekly Friday meetings where senior executives share oftentimes sensitive, non-public information with much of the company, engaging in an informal question and answer session where any employee can theoretically pose their questions to higher ups. In a 2009 blog post, a Google

vice president explained that “openness” is the fundamental principle upon which the company operates, ranging from their (admittedly self-serving) goal to make information “open” via their search engine and other products, to their philosophy on open-source code, open protocols, and open corporate culture (Rosenberg 2009).

While platforms have this notable internal culture of transparency, they are considerably less open to the outside world. To visit Facebook or Google as an outsider (perhaps a journalist, researcher, or academic), one must make it through the reception desk and keycard-entry turnstiles, usually after signing strict non-disclosure agreements. Externally, platforms are notable for their corporate secrecy — for example, it took twelve years for Facebook to release detailed information about its “Community Standards,” the rules that govern speech on the site, despite longstanding civil society and academic pressure (Gillespie 2018a). As the management and organization researcher Mikkel Flyverbom has written, platforms are characterized by “strong forms of vertical transparency, where employees and employers can observe each other very easily, but also as very little horizontal transparency because outsiders have very few opportunities to scrutinize the insides of these companies” (Flyverbom 2016, 177). This has made research into platforms difficult, especially as firms shut down the developer APIs and other tools traditionally used by researchers to access data (Hogan 2018). Much like a highly opaque bureaucracy, only a limited transparency has currently been achieved via public statements and interviews granted by their executives, as well as the occasional incident of whistleblowing and leaking. For instance, training documents issued to contractors that engage in commercial content moderation on Facebook were leaked to the Guardian, leading to significant public outcry and a better understanding of how Facebook’s moderation functions (Klonick 2017). As we argue here, transparency initiatives for platforms can be classified into four camps: voluntary transparency around freedom of expression and for content takedowns; legally mandated transparency regimes; self-transparency around advertising and content moderation; and third-party tools, investigations, and audits.

3.1 Voluntary Transparency for Content Takedowns

Virtually since their emergence in the early 2000s, platform companies have had to weigh legal requests for content takedowns from individuals and governments around the world (Goldsmith and Wu 2006). As Daphne Keller explains in this volume, intermediary liability laws inform online intermediaries of their legal obligations for the information posted by their users, placing platform companies in the often uncomfortable position of having to weigh content against their own guidelines, local laws, and normative goals around freedom of expression. In an effort to maintain their legitimacy (and their normative position as promoters of free expression ideologically crafted in the First Amendment tradition), platform companies have become more transparent in the past decade about these processes as they pertain content takedowns for copyright purposes and other legal reasons. To date, the dominant mode for horizontal transparency

implemented by major platform companies has been in the area of these speech and content takedown requests.

In 2008, as part of an effort to combat censorship and protect human rights online, the Global Network Initiative (GNI) was created, with Microsoft, Yahoo, Google, and a number of civil society organizations and academic institutions as founding members. As part of a commitment to the GNI principles, Google introduced an annual “Transparency Report” in 2010, the first company to publicly release data about content takedown and account information requests filed by governments around the world, along with a “Government Requests” tool, which visualized this data and provided a FAQ of sorts for individuals interested in the government takedown process (Drummond 2010). Grandiosely citing Article 19 of the Universal Declaration on Human Rights, Google positioned itself as a defender of free expression and civil liberties violations as enacted by states (Drummond 2010). In the years to follow, these tools were expanded: in 2011, Google made the raw data underpinning the report public, and in 2012, the company added copyright takedowns under DMCA and other intermediary liability laws to the report. In July 2012, Twitter began publishing a biannual transparency report, which now includes information about account information requests, content removal requests, copyright takedown notices, and requests to check content against Twitter’s own terms of service. In January 2013, as it joined the GNI, Facebook begins publishing its own bimonthly transparency report, which includes numbers about requests for user data, broken down by country, and in July 2013, numbers about how much content was restricted based on local law. In July 2013, documents provided to the press by Edward Snowden documented extensive governmental systems for mass surveillance, with the disclosures suggesting that the US National Security Agency had access to unencrypted information flowing between Google data centres (Gellman and Soltani 2013). Google disputed the extent that information was collected via PRISM and other disclosed NSA programs, but expanded its transparency reporting efforts to include information about National Security Letters and FISA court orders.

3.2 Voluntary Transparency for Content and Advertisements

The significant public pressure on platform companies following the 2016 US election has led to a new series of voluntary horizontal transparency initiatives that go beyond just content takedown requests. The most notable development has been the release of an expanded, public-facing version of Facebook’s “Community Standards” in April 2018. While Facebook transparency reports have since 2013 provided aggregate numbers about content that “is reported to [Facebook] as violating local law, but doesn’t go against Community Standards,” the details of those Community Standards were not public. For instance, the Standards stated that content that featured sexual content, graphic violence, or hate speech was prohibited, but did not define any of those highly contentious categories or provide information into how they had been defined (Gorwa 2018a),

leading civil society and academics to almost universally critique Facebook’s processes as highly opaque and problematic (Myers West 2018).

Facebook’s publication of more detailed “Community Standards” that have information about the policies (eg. how hate speech is defined, how sexual content and nudity is defined) give far more context to users and marks an important step forward. Companies are starting to also provide detail about how these policies are enforced: at the end of April 2018, Google published their first “Community Guidelines Enforcement Report,” a type of transparency report which provides numbers into the amount of violating material taken down in various categories of the Community Standards, and the role of automated systems in detecting content before it is reported. In May 2018, Facebook published a similar “Community Standards Enforcement Report,” which illustrates different types of content takedowns, providing aggregate data into how much content is removed. For six parts of the Community Standards (graphic violence, adult nudity and sexual activity, terrorist content, hate speech, spam, and inauthentic accounts), Facebook provides four data points: how many Community Standards violations were found, the percentage of flagged content upon which action is taken; the amount of violating content found and flagged by automated systems; and the speed at which the company’s moderation infrastructure acts in each case. In May 2019, Facebook added multiple new categories to its Enforcement Report, including interesting data on the amount of content that is removed, appealed, and re-instated. (Since mid-2018, users who have content removed by Facebook’s manual or automated moderation systems can have those decisions appealed). For example, the report states that from January to March 2019, 19.4 million pieces of content were removed from Facebook under their policies prohibiting certain types of nudity and sexual content. Out of that 19.4 million, 2.1 million (approximately 10.8 percent) was appealed by users, with Facebook reversing the decision in 453 000 (approximately 21.5 percent) of those instances (Table 1). Given the sheer quantity of Facebook content being posted by users every day, this may not seem like an enormous number; however, consider a hypothetical situation where each one of these appeals represents a separate user (setting aside for the moment that users could have had multiple pieces of content taken down): before Facebook began allowing appeals in 2018, more than four hundred thousand people around the world would have been frustrated by the wrongful takedown of their images, videos, or writing, which may have had political or artistic merit (see Gillespie 2018, Gorwa 2018b), or may not have been nudity at all, picked up by a faulty automated detection system, for instance.

These reports are a major step, and should continue to be expanded (Garton Ash, Gorwa, and Metaxa 2019). But they still have notable limitations: as the Electronic Frontier Foundation’s Jillian York has written, the Facebook report “deals well with how the company deals with content that violates the rules, but fails to address how the company’s moderators and automated systems can get the rules wrong, taking down content that doesn’t actually violate the Community Standards” (York 2018). The 2019 report focuses heavily on Facebook’s increasing use of automated systems to take down hate speech: as

Policy Area	Amount of Content Appealed	Amount of Content Restored After Appeal
Spam	20.8 M	5.7 M
Adult Nudity and Sexual Activity	2.1 M	453 K
Hate Speech	1.1 M	130 K
Bullying and Harassment	496 K	80.2 K
Violence and Graphic Content	171 K	23.9 K
Regulated Goods - Drugs	87.4 K	18.1 K
Regulated Goods - Firearms	80.8 K	6.2 K
Terrorist Propaganda (ISIS, Al-Qaeda and affiliates)	40.1 K	20.3 K
Child Nudity and Sexual Exploitation of Children	20.6 K	0.7 K

Figure 1: Appeals data provided in Facebook's May 2019 Community Standards Enforcement Report (Rosen 2019)

of March 2019, 65 percent of content Facebook removes under its hate speech policies globally are automatically flagged, up from 24 percent at the end of 2017. This is presented as an unqualified good thing, but the increasing use of automated systems in content moderation will have a number of underexplored consequences, and may make content moderation more opaque by adding a layer of computational complexity at a time where almost everyone is seeking greater transparency (Gorwa, Binns, and Katzenbach, forthcoming).

In 2017 and 2018, Facebook also tested and rolled out a number of features with the stated aim to improve transparency around political advertising (especially in electoral periods). Following the discovery that Russian operatives had purchased Facebook advertisements to target American voters in the lead up to the 2016 US election, Mark Zuckerberg issued an extended public statement in which he announced Facebook’s goal to “create a new standard for transparency in online political ads.” These efforts began in October 2017, when Facebook announced advertisers wishing to run election related advertisements in the United States would have to register with Facebook and verify their identity, and that these ads would be accompanied by a clickable “Paid for by...” disclosure. In April 2018, Facebook announced that this verification process would be expanded to all political “issue ads.” Beginning in November 2017, a feature called “view ads” was tested in Canada, where users could navigate the drop-down menu on a page to see all of the ads that it was currently running. This was expanded to the United States and Brazil in the Spring of 2018, and in June 2018, allowed users to view ads run by a page across platforms (Facebook, Instagram, and Messenger). In May 2018, Facebook launched a public political ad archive of all political ads being run, and in August 2018, Facebook provided a political ad API that would allow researchers and journalists to access this archive.

Other platform companies simultaneously rolled out similar initiatives. In October 2017, Twitter announced a number of future advertising transparency efforts, including a “Transparency Center,” which launched in June 2018 and provided users with the ability to view ads targeted to them (along with the personalized information being used to target that ad, as well as a view of all ads being run on the platform. As well, Twitter has mandated that political “electioneering ads” are clearly distinct from other ads, and has opted to provide information about the advertiser, the amount spent on the ad campaign, targeting data used, and other relevant information. These efforts remain a work in progress: in May 2018, Twitter launched announced their policy on political campaigning, to explain how they defined political advertising and the steps that political advertisers would have to take to register with the company, and in August, they outlined their specific policy on issue ads. Google announced in May 2018 that it would require identification from any party seeking to run an election related advertisement on a Google platform for a US audience, and that all such ads would have a “paid for by” notice. In August 2018, Google added a “political advertising” section to its transparency reports, providing a public facing tool for users to search for certain advertisers, a district level breakdown of political ad spending in the United States, and searchable ad database fea-

turing all video, text, and image advertisements about “federal candidates or current elected federal officeholders.”

3.3 Mandated Transparency Regimes

Broader forms of transparency have been mandated by a combination of local regulation and by membership in certain voluntary organizations. Along with a commitment to public transparency via transparency reports, the GNI requires independent-third party assessments undertaken every two years to ensure compliance with the GNI principles. These assessments are performed by a series of trusted auditors, who look into processes (the policies, and procedures employed by the member companies) and review specific cases, with their findings compiled into a report released by the GNI. Under the 2011 and 2012 Federal Trade Commission consent decrees about Google and Facebook’s deceptive privacy practices, the two companies were required to subject themselves to privacy audits for 20 years (Hoofnagle 2016). As part of this regime, “compliance reports” are publicly available online at the FTC website, but the mandated third-party audits, which initially caused excitement as a possible transparency mechanism, are only published in heavily redacted and jargon-heavy form. These audits are thin on substance due to the watered-down language in the final FTC decree, making them, as one legal scholar argued, “so vague or duplicative as to be meaningless” (Gray 2018, 4). Recently, German legislation seeking the removal of “manifestly illegal” content (the *Netzwerkdurchsetzungsgesetz* or *NetzDG*), has included mandatory public transparency reporting for large platform companies operating Germany. This has resulted the publication of reporting that includes: details about operational procedures, staffing and training (with information about the number of employees reviewing *NetzDG* reports, the contractors hired to fulfill those reports, the support and wellness services available to those contractors and employees, and more), detailed breakdowns of content removals per each relevant section of the German Criminal Code; detailed breakdowns of the time taken for those content removals; and the way in which users were notified about content take downs (Schulz 2018). While the legislation is controversial and has been widely criticized on freedom of expression grounds, many of *NetzDG*’s harshest civil society critics stated their support for the transparency provisions it enacted (see Tworek and Leerssen 2018). However, researchers have noted the limitations of the data provided in the first transparency reports, released in late 2018, which do not have sufficient granularity to be empirically useful (Heldt 2019).

In the past two years, a growing number of governments pushed for platform companies to create publicly accessibly archives of political advertisements, with legislation requiring ad archiving being passed in France and Canada and being proposed and debated in the United Kingdom and United States (Leerssen et al. 2019; McFaul 2019). The European Commission’s Code of Practice on Disinformation includes a voluntary commitment from of firms to develop systems for disclosing issue ads, and recommends transparency without articulating specific mechanisms. Ad archiving appears to be an important new frontier for

transparency and disclosure, but major questions around scope and details of implementation remain (see Leerssen et al. 2019 for a comprehensive discussion of advertising archives as a form of transparency).

3.4 Third Party “Audits” and Investigations

A final, indirect source of transparency and information about the dealings of platform companies is created by research and investigative work by third-parties, such as journalists. The investigative journalism non-profit ProPublica has conducted multiple key investigations into Facebook’s advertising interfaces, successfully showing how, for example, those interfaces can be used to target anti-Semitic users, or exclude certain minorities from seeing ads for housing or jobs. This work also can rely on crowdsourcing: before Facebook made their ad API available, ProPublica built a browser extension that Facebook users could install to pull the ads that they saw on their Facebook, and a similar strategy was employed by the British group WhoTargetsMe, as well as researchers at the University of Wisconsin who found that the majority of issue ads they studied in the lead up to the 2016 US election did not originate from organizations registered with the Federal Election Commission (Kim et al. 2018). Although Facebook itself seems to conceptualize third party research as a meaningful accountability mechanism, it has not made it easy for this work to be undertaken, which generally violates Facebook’s terms of service and puts researchers on precarious legal footing. Recently, following recommendations from researchers and civil society, Facebook has commissioned a few of third-party assessments into its impact in certain domains, e.g. in fragile states like Myanmar (Warofka 2018; Garton Ash, Gorwa, and Metaxa 2019).

4 The Future of Platform Transparency

Facebook, Google, and Twitter have made efforts to bring more transparency to political advertising and content policy. These efforts certainly demonstrate a willingness to be more transparent about processes than a decade ago, but according to civil society, still do not go far enough. Efforts to measure the transparency measures of various companies discussed, such as an annual report conducted by the EFF, gave Facebook one out of five stars (citing a lack of meaningful notice to users who will have content taken down by a government or copyright request, as well as a lack of appeal on these decisions; see Cardozo et al. 2018). Ranking Digital Rights, a project based at the New America Foundation which produces an annual “Corporate Accountability Report” with scorecards for 22 platform companies, internet service providers, and other online intermediaries, gave Facebook a score of 55 out of 100 in 2018. The researchers noted that Facebook “disclos[es] less about policies affecting freedom of expression and privacy than most of its U.S. peers,” and that the company “provided users with limited options to control what information the company collects, retains, and uses, including for targeted advertising, which appears to

be on by default” (Ranking Digital Rights 2018, 87). Google ranked first out of the companies assessed (with a score of 63 of 100), with the report noting that they now allow users to opt out of targeted advertising and make more disclosures about freedom of expression, content takedowns, and privacy than other platforms, but still lack robust grievance and appeal mechanisms (Ranking Digital Rights 2018, 89). The Report’s 2019 iteration applauded Facebook for providing more transparency and user autonomy around free expression issues (for example, by publishing its first Content Standards Enforcement report, and beginning appeals for content takedown decisions), but noted that it still lagged behind on privacy, offering “less choice for users to control the collection, retention, and use of their information than all of its peers other than Baidu and Mail.Ru” (Ranking Digital Rights 2019, 20).

Transparency is important for legislators and the public to have realistic expectations about what the processes and capabilities of platform companies to make decisions and take down content are. As Daphne Keller writes in this volume, without transparency, “policymakers misjudge what platforms can do... They draft and support poorly tailored laws as a result.” But the existing forms of horizontal transparency as currently enacted by these companies have major limitations. First, they are predominantly voluntary measures that have little enforcement mechanisms: the public must hope that the data provided is accurate, but has little mechanisms for verifying that it is. A broader question can also be raised about the overall utility of transparency reports: while some features have clear democratic, normative importance, such as notice and appeals processes, the highly technical, aggregate statistics that are the core of most transparency reports are not necessarily useful in reducing the overall opacity of the system, where key processes, protocols, and procedures remain secret. Finally, a major limitation of the existing voluntary transparency measures, especially the advertising transparency efforts which have clearly been crafted to supplant possible regulatory regimes, is that these primarily public-facing tools may provide a rich and informative resource for students, researchers, and journalists, but are less useful for regulators (e.g. electoral officers). Certainly, while Google’s new advertising library may be useful for students trying to understand attack ads, and may yield interesting tidbits worthy of investigation for reporters, electoral commissions will require their own machine-readable database which permanently archives advertisements. At the time of writing, roughly a third of the advertisements on the front page of the Google Ad library were no longer viewable — perhaps due to their violation of Google’s guidelines, or due to requests for removal filed on behalf of the advertiser.

In an important article, Mike Ananny and Kate Crawford outline ten challenges facing transparency efforts that strive to govern complex socio-technical systems such as forms of algorithmic decision making (Ananny and Crawford 2018). The article is an important contribution to the recent literature on transparency as it pertains to technology policy issues and the digital age — a literature, one should note, that is roundly critical of the productive possibilities of transparency as a form of governing and knowing (Flyverbom 2016).

Many of Ananny and Crawford’s insights (Table 1) can be applied to the case

Transparency...
can be disconnected from Power
can be harmful
can intentionally occlude
can create false binaries
can invoke neoliberal models of agency
can privilege seeing over understanding
does not necessarily build trust
entails professional boundary work
has technical limitations
has temporal limitations

Table 1: Common Digital Transparency Pitfalls (Ananny & Crawford, 2018)

of platform transparency. For example, they warn that if “If transparency has no meaningful effects, then the idea of transparency can lose its purpose” and effectively backfire, as it becomes disconnected from overarching power relationships (Ananny and Crawford 2018, 979). If platform transparency efforts around advertising serves to make people cynical about the inevitability of poisonous attack ads, or, on the other hand, ameliorate some narrow concerns about the generally problematic system of privacy-invasive advertising that funds contemporary platforms without actually making these systems more just, the efforts will fail to achieve a democratically desirable outcome. Furthermore, certain forms of transparency can “intentionally occlude” (Ananny and Crawford 2018, 980), burying the important insights in mounds of superfluous data or providing a distraction from the fundamental harms caused by a system. How useful are existing transparency reports, and do they distract from the end reality that some online intermediaries are making tremendously important decisions about the boundaries of acceptable political speech around the world in a fundamentally undemocratic and unaccountable manner? Do these measures provide the illusion of seeing inside a system, without providing meaningful understanding of how it really functions and operates (Ananny and Crawford 2018, 982)? Do existing transparency measures lend a modicum of legitimacy and due process to a host of American private companies with global reach and impact, without actually providing good governance? These are crucial questions for scholars interested in the future of social media and its relationship with democracy.

4.1 The Facebook Oversight Board

As governments grapple with a host of complex policy options, transparency mechanisms seem to provide a logical path forward. Platform companies seem willing to become more transparent, and are even increasingly doing so voluntarily. When compared with broad and potentially messy legislation, transparency seems as a fruitful, if limited, form of accountability that has the additional benefit of yielding information that better inform the future policy climate.

(Platform companies themselves even come to support some limited legislation which is squarely in the realm of transparency, such as the US Honest Ads Act, which legally enshrines political ad transparency and is now supported by Twitter and Facebook). Legislated transparency, which comes with meaningful third-party audits and verification measures, has the potential to improve current systems, although the track-record of the admittedly small sample of such efforts has been mixed. There is a great deal that firms could do to make the data they release as part of their transparency reporting more meaningful and reproducible (Bradford et al. 2019).

One major developing initiative is the ‘Oversight Body’ for content policy that Facebook has been workshopping in 2019. The proposal, which built on Mark Zuckerberg’s comments in 2018 about a ‘Supreme Court’ style mechanism for external appeals (Garton Ash, Gorwa, and Metaxa 2019), seeks to formalize a system of third-party input into Facebook’s Community Standards. The initial conception of the Oversight Body focused primarily on procedural accountability, but during a consultation process with academics and civil society groups, including the authors of this chapter, it appears to have become more broadly conceived as a mechanism “intended to provide a layer of additional transparency and fairness to Facebook’s current system” (Darne and Miller 2019). In September 2019, Facebook published a nine-page charter for the Oversight Board, which outlines how the Board will be set up (via an independent trust), how members will be selected (Facebook will select an initial group, which will then select the remaining 40 part-time ‘jurors’), and commits the company to acting on the board’s binding decisions.

Much yet remains to be seen about the Board’s implementation, and whether it truly results in a more democratically legitimate system of speech governance (Kadri & Klonick 2019) for the average Facebook user — someone who is far more likely to be a non-English speaker, mobile-first, and located in Bombay or Bangkok than Boston — or if it merely becomes a new type of informal governance arrangement, with more in common with the many certification, advisory, and oversight bodies established in the natural resource extraction or manufacturing industries (Gorwa 2019b). While a group of trusted third-parties publishing in-depth policy discussions of the cases they explore, if adequately publicized and implemented in a transparent manner, certainly has the potential to increase the familiarity of social media users with the difficult politics of contemporary content moderation, it may not significantly increase the transparency of Facebook’s actual practices. The Oversight Board could even become a ‘transparency proxy’ of sorts, attracting public attention but remaining little more than a translucent layer that floats on top of a still-opaque institution.

5 Conclusion

Transparency should help consumers make informed choices. As the traditional argument for democratic transparency articulates, “People need information to assess whether such organizations protect their interests or predate upon them,

to choose which organizations (and so which products and services) to rely upon, to decide whether to oppose or support various organizations, and to develop and execute strategies to affect and interact with them” (Fung 2013, 184). Yet, the contemporary “attention economy” that characterizes most digital services is profoundly opaque, and users are intentionally kept from seeing much of how their behavior is shaped and monetized, whether by the multitude of hidden trackers and behavioral advertising pixels that follow them around the Web, or by the secret rules and processes that determine the bounds of acceptable speech and action. As citizens weigh the tradeoffs inherent in their usage of social media or other services provided by companies like Facebook, Google, or Twitter, they deserve to better understand how those systems function (Gillespie 2018b).

The companies have displayed a commendable effort in the past decade to provide some data about the way they interact with governments when it comes to freedom of expression; however, these tend to offer more insight into government behavior rather than their own. Extracting the policies, practices, and systems through which platform companies affect free expression or other democratically essential areas such as privacy is a slog, as journalists, academics, and activists painstakingly try to obtain what they can, as if pulling teeth from an unwilling patient. Measures enacted willingly thus far by Facebook, Google, and Twitter (or even those enacted in consent decrees with the Federal Trade Commission and other agencies) remain unlikely to result in meaningful long-term reform. Digital transparency — whether it is enacted through technological solutions or more classical administrative and organizational means — will not on its own provide an easy solution to the challenges posed by the growing role of platforms in political and public life. More research will be required to examine how platform companies enact and perform transparency, and how this transparency functions in an increasingly contested governance landscape.

As Ananny and Crawford (2018, 983) argue, “A system needs to be understood to be governed.” Transparency in key areas, such as content moderation, political advertising, and automated decisionmaking, should be an important first step reinforced by legislators. In the United States, the Brandeisian role of transparency in combatting corporate power has often been neglected by the new cadre of advocates picking up the anti-trust banner. While it is important to think critically about the pitfalls and shortcomings of transparency initiatives, a significant amount of robust transparency, legislatively mandated where necessary, would certainly contribute to the public understanding of the digital challenges facing democracies around the world.

6 References

Albu, Oana Brindusa, and Mikkel Flyverbom. 2016. “Organizational Transparency: Conceptualizations, Conditions, and Consequences.” *Business & Society*: 0007650316659851.

Ananny, Mike, and Kate Crawford. 2018. “Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Account-

ability.” *New Media & Society* 20(3): 973–89.

Banisar, David. 2006. “Freedom of Information around the World 2006: A Global Survey of Access to Government Information Laws.” London: Privacy International.

Bannister, Frank, and Regina Connolly. 2011. “The Trouble with Transparency: A Critical Review of Openness in e-Government.” *Policy & Internet* 3(1): 1–30.

Baume, Sandrine. 2018. “Publicity and Transparency: The Itinerary of a Subtle Distinction.” In *Transparency, Society and Subjectivity*, eds. Emmanuel Alloa and Dieter Thomä. Cham, Switzerland: Springer, 203–24.

Bennington, Geoffrey. 2011. “Kant’s Open Secret.” *Theory, Culture & Society* 28(7–8): 26–40.

Bertot, John C., Paul T. Jaeger, and Justin M. Grimes. 2010. “Using ICTs to Create a Culture of Transparency: E-Government and Social Media as Openness and Anti-Corruption Tools for Societies.” *Government Information Quarterly* 27(3): 264–271.

Birchall, Clare. 2011. “Introduction to ‘Secrecy and Transparency’: The Politics of Opacity and Openness.” *Theory, Culture & Society* 28(7–8): 7–25.

———. 2014. “Radical Transparency?” *Cultural Studies? Critical Methodologies* 14(1): 77–88.

Bok, Sissela. 1983. *Secrets: On the Ethics of Concealment and Revelation*. 1st ed. New York: Pantheon Books.

Bonsón, Enrique, Lourdes Torres, Sonia Royo, and Francisco Flores. 2012. “Local E-Government 2.0: Social Media and Corporate Transparency in Municipalities.” *Government Information Quarterly* 29(2): 123–132.

Bradford, Ben et al. 2019. *Report Of The Facebook Data Transparency Advisory Group*. The Justice Collaboratory: Yale Law School.

Brandeis, Louis D. 1913. “What Publicity Can Do.” *Harper’s Weekly*.

Brock, George. 2017. “How to Regulate Facebook and the Online Giants in One Word: Transparency.” *The Conversation*. (September 28, 2018).

Burrell, Jenna. 2016. “How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms.” *Big Data & Society* 3(1): 2053951715622512.

Bushman, Robert M., Joseph D. Piotroski, and Abbie J. Smith. 2004. “What Determines Corporate Transparency?” *Journal of Accounting Research* 42(2): 207–252.

Cardozo, Nate et al. 2018. *Who Has Your Back? Censorship Edition* 2018. Electronic Frontier Foundation. (September 29, 2018).

Christensen, Lars Thøger, and George Cheney. 2014. “Peering into Transparency: Challenging Ideals, Proxies, and Organizational Practices.” *Communication Theory* 25(1): 70–90.

Cucciniello, Maria, Gregory A. Porumbescu, and Stephan Grimmlikhuijsen. 2017. “25 Years of Transparency Research: Evidence and Future Directions.” *Public Administration Review* 77(1): 32–44.

Drummond, David. 2010. “Greater Transparency around Government Requests.” Google BlogGoogle Blog. (September 23, 2018).

- Edwards, Lilian, and Michael Veale. 2017. "Slave to the Algorithm: Why a Right to an Explanation Is Probably Not the Remedy You Are Looking For." *Duke Law & Technology Review* 16: 18.
- Etzioni, Amitai. 2010. "Is Transparency the Best Disinfectant?" *Journal of Political Philosophy* 18(4): 389–404.
- Feldman, Brian. 2017. "Facebook Hands Over Data to Congress and Promises Transparency on Political Ads." *New York Magazine*. (September 28, 2018).
- Fink, Katherine. 2018. "Opening the Government's Black Boxes: Freedom of Information and Algorithmic Accountability." *Information, Communication & Society* 21(10): 1453–71.
- Flyverbom, Mikkel. 2015. "Sunlight in Cyberspace? On Transparency as a Form of Ordering." *European Journal of Social Theory* 18(2): 168–184.
- . 2016. "Transparency: Mediation and the Management of Visibilities." *International Journal of Communication* 10: 13.
- . 2019. *The Digital Prism: Transparency and Managed Visibilities in a Datafied World*. Cambridge, UK: Cambridge University Press.
- Fox, Jonathan. 2007. "The Uncertain Relationship between Transparency and Accountability." *Development in Practice* 17(4–5): 663–671.
- Frynas, Jędrzej George. 2005. "The False Developmental Promise of Corporate Social Responsibility: Evidence from Multinational Oil Companies." *International Affairs* 81(3): 581–598.
- Fung, Archon. 2013. "Infotopia: Unleashing the Democratic Power of Transparency." *Politics & Society* 41(2): 183–212.
- Fung, Archon, Mary Graham, and David Weil. 2007. *Full Disclosure: The Perils and Promise of Transparency*. Cambridge, UK: Cambridge University Press.
- Gaonkar, Dilip Parameshwar, and Robert J. McCarthy Jr. 1994. "Panopticism and Publicity: Bentham's Quest for Transparency." *Public Culture* 6(3): 547–575.
- Garton Ash, Timothy. 2016. *Free Speech: Ten Principles for a Connected World*. New Haven, CT: Yale University Press.
- Garton Ash, Timothy, Robert Gorwa, and Danaë Metaxa. 2019. "Glasnost! Nine Ways Facebook Can Make Itself a Better Forum for Free Speech and Democracy." Oxford, UK: Reuters Institute for the Study of Journalism.
- Gaventa, John, and Rosemary McGee. 2013. "The Impact of Transparency and Accountability Initiatives." *Development Policy Review* 31: s3–s28.
- Gellman, Barton, and Ashkan Soltani. 2013. "NSA Infiltrates Links to Yahoo, Google Data Centers Worldwide, Snowden Documents Say." *Washington Post*. (September 23, 2018).
- Gillespie, Tarleton. 2018a. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven: Yale University Press.
- . 2018b. "Regulation of and by Platforms." In *The SAGE Handbook of Social Media*, eds. Burgess, Jean, Alice Marwick, and Thomas Poell. London: SAGE, 254–78.

- Goldsmith, Jack L, and Tim Wu. 2006. *Who Controls the Internet?: Illusions of a Borderless World*. New York: Oxford University Press.
- Gorwa, Robert. 2018a. "Platform Moderation and Its Discontents." *Los Angeles Review of Books*. (September 29, 2018).
- . 2018b. "Facebook's Decency Conundrum. *Wired Magazine* (UK). November/December.
- . 2019a. "What Is Platform Governance?" *Information, Communication & Society* 22(6): 854–71.
- . 2019b. "The Platform Governance Triangle: Conceptualising the Informal Regulation of Online Content." *Internet Policy Review* 8(2). doi:10.14763/2019.2.1407
- Gray, Megan. 2018. "Understanding and Improving Privacy 'Audits' under FTC Orders." Stanford CIS White Paper.
- Grimmelikhuijsen, Stephan, Gregory Porumbescu, Boram Hong, and Tobin Im. 2013. "The Effect of Transparency on Trust in Government: A Cross-National Comparative Experiment." *Public Administration Review* 73(4): 575–86.
- Hansen, Hans Krause, and Mikkel Flyverbom. 2015. "The Politics of Transparency and the Calibration of Knowledge in the Digital Age." *Organization* 22(6): 872–889.
- Heemsbergen, Luke. 2016. "From Radical Transparency to Radical Disclosure: Reconfiguring (in) Voluntary Transparency through the Management of Visibilities." *International Journal of Communication* 10: 14.
- Heldt, Amélie. 2019. "Reading between the Lines and the Numbers: An Analysis of the First NetzDG Reports." *Internet Policy Review* 8(2). doi:10.14763/2019.2.1398
- Hess, David. 2012. "Combating Corruption through Corporate Transparency: Using Enforcement Discretion to Improve Disclosure." *Minnesota Journal of International Law* 21: 42–74.
- Hogan, Bernie. 2018. "Social Media Giveth, Social Media Taketh Away: Facebook, Friendships, and APIs." *International Journal of Communication* 12: 20.
- Hood, Christopher. 2006. "Transparency in Historical Perspective." In *Transparency: The Key to Better Governance?*, eds. Christopher Hood and David Heald. Oxford, UK: Oxford University Press.
- . 2010. "Accountability and Transparency: Siamese Twins, Matching Parts, Awkward Couple?" *West European Politics* 33(5): 989–1009.
- . 2011. "From FOI World to WikiLeaks World: A New Chapter in the Transparency Story?" *Governance* 24(4): 635–38.
- Hood, Christopher, and David Heald, eds. 2006. *Transparency: The Key to Better Governance?* Oxford, UK: Oxford University Press.
- Hoofnagle, Chris Jay. 2016. "Assessing the Federal Trade Commission's Privacy Assessments." *IEEE Security & Privacy* 14(2): 58–64.
- Hume, L. J. 1981. *Bentham and Bureaucracy*. New York: Cambridge University Press.
- Kim, Young Mie et al. 2018. "The Stealth Media? Groups and Targets behind Divisive Issue Campaigns on Facebook." *Political Communication*: 1–29.

- King, Gary, and Nathaniel Persily. 2019. "A New Model for Industry-Academic Partnerships." *PS: Political Science and Politics*.
- Klonick, Kate. 2017. "The New Governors: The People, Rules, and Processes Governing Online Speech." *Harvard Law Review* 131: 1598.
- Kosack, Stephen, and Archon Fung. 2014. "Does Transparency Improve Governance?" *Annual Review of Political Science* 17: 65–87.
- Kroll, Joshua A. et al. 2016. "Accountable Algorithms." *University of Pennsylvania Law Review* 165: 633.
- Lamble, Stephen G. 2002. "Freedom of Information, a Finnish Clergyman's Gift to Democracy." *Freedom of Information Review* (97): 2–8.
- Leerssen, Paddy et al. 2019. "Platform Ad Archives: Promises and Pitfalls." SSRN Scholarly Paper. (July 17, 2019).
- Levy, Karen EC, and David Merritt Johns. 2016. "When Open Data Is a Trojan Horse: The Weaponization of Transparency in Science and Governance." *Big Data & Society* 3(1): 2053951715621568.
- Lodge, Juliet. 1994. "Transparency and Democratic Legitimacy." *JCMS: Journal of Common Market Studies* 32(3): 343–68.
- McFaul, Michael, ed. 2019. "Securing American Elections: Prescriptions for Enhancing the Integrity and Independence of the 2020 U.S. Presidential Election and Beyond". Palo Alto, CA: Stanford Cyber Policy Center.
- Meijer, Albert. 2015. "Government Transparency in Historical Perspective: From the Ancient Regime to Open Data in The Netherlands." *International Journal of Public Administration* 38(3): 189–99.
- Moore, Martin, and Damian Tambini, eds. 2018. *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*. Oxford: Oxford University Press.
- Myers West, Sarah. 2018. "Censored, Suspended, Shadowbanned: User Interpretations of Content Moderation on Social Media Platforms." *New Media & Society*: 1461444818773059.
- O'Neill, Onora. 2006. "Transparency and the Ethics of Communication." In *Transparency: The Key to Better Governance?*, eds. Christopher Hood and David Heald. Oxford, UK: Oxford University Press, 75–90.
- Owen, Taylor. 2015. *Disruptive Power: The Crisis of the State in the Digital Age*. Oxford: Oxford University Press.
- Power, Michael. 1997. *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.
- Ranking Digital Rights. 2018. 2018 Corporate Accountability Index. Washington, DC: New America Foundation.
- Ranking Digital Rights. 2019. 2019 Corporate Accountability Index. Washington, DC: New America Foundation.
- Roberts, John. 2009. "No One Is Perfect: The Limits of Transparency and an Ethic for 'Intelligent' Accountability." *Accounting, Organizations and Society* 34(8): 957–970.
- Rosen, Guy. 2019. "An Update on How We Are Doing At Enforcing Our Community Standards." Facebook Newsroom. (January 1, 2019).

- Rosenberg, Jonathan. 2009. "The Meaning of 'Open.'" Google Blog. (September 23, 2018).
- Ruggie, John Gerard. 2013. *Just Business: Multinational Corporations and Human Rights*. New York: W. W. Norton & Company.
- Schudson, Michael. 2015. *The Rise of the Right to Know: Politics and the Culture of Transparency, 1945-1975*. Cambridge, MA: Harvard University Press.
- Stohl, Cynthia, Michael Stohl, and Paul M. Leonardi. 2016. "Managing Opacity: Information Visibility and the Paradox of Transparency in the Digital Age." *International Journal of Communication* 10: 15.
- Tan, Yeling. 2014. "Transparency without Democracy: The Unexpected Effects of China's Environmental Disclosure Policy." *Governance* 27(1): 37–62.
- Thompson, Dennis F. 1999. "Democratic Secrecy." *Political Science Quarterly* 114(2): 181–193.
- Tucker, Joshua A, Yannis Theocharis, Margaret E Roberts, and Pablo Barberá. 2017. "From Liberation to Turmoil: Social Media And Democracy." *Journal of Democracy* 28(4): 46–59.
- Turner, Fred. 2009. "Burning Man at Google: A Cultural Infrastructure for New Media Production." *New Media & Society* 11(1–2): 73–94.
- . 2010. *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*. University of Chicago Press.
- Underwood, Sarah. 2016. "Blockchain beyond Bitcoin." *Communications of the ACM* 59(11): 15–17.
- Vaccaro, Antonino, and Peter Madsen. 2009. "Corporate Dynamic Transparency: The New ICT-Driven Ethics?" *Ethics and Information Technology* 11(2): 113–122.
- Vaidhyanathan, Siva. 2018. *Antisocial Media: How Facebook Disconnects Us and Undermines Democracy*. Oxford: Oxford University Press.
- Van Dijck, José, Thomas Poell, and Martijn de Waal. 2018. *The Platform Society: Public Values in a Connective World*. New York: Oxford University Press.
- Wachter, Sandra, Brent Mittelstadt, and Chris Russell. 2017. "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR." *Harvard Journal of Law and Technology* 31(2).
- Waddock, Sandra. 2004. "Creating Corporate Accountability: Foundational Principles to Make Corporate Citizenship Real." *Journal of Business Ethics* 50(4): 313–27.
- Warofka, Alex. 2018. "An Independent Assessment of the Human Rights Impact of Facebook in Myanmar." Facebook Newsroom.
- York, Jillian C. 2018. "Facebook Releases First-Ever Community Standards Enforcement Report." Electronic Frontier Foundation. (September 29, 2018).