Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation

*Jack M. Balkin**

TABLE OF CONTENTS

 Knight Professor of Constitutional Law and the First Amendment, Yale Law School. This Essay is based on remarks delivered at the Law in the Information Age Lecture at UC Davis School of Law on March 15, 2017. My thanks to Kate Klonick and Robert Post for their comments on previous drafts.

INTRODUCTION

The problems of free speech in any era are shaped by the communications technology available for people to use and by the ways that people actually use that technology.

Twenty years ago, in 1997, when I began the Information Society Project at Yale, we were just entering the age of the Internet. Most people were still using dial-up modems, there was no Facebook, Google, or YouTube, Instagram or Snapchat; there were no iPhones. Only twenty years later, we have already entered into a new phase — the Algorithmic Society — which features large, multinational social media platforms that sit between traditional nation states and ordinary individuals, and the use of algorithms and artificial intelligence agents to govern populations.[1]

In previous work, I have argued that the digital age makes salient one of the central purposes of freedom of speech. The goal of free speech, I contend, is to protect and foster a democratic culture. A democratic culture is a culture in which individuals have a fair opportunity to participate in the forms of meaning-making and mutual influence that constitute them as individuals.[2]

The early Internet seemed to symbolize the possibilities for such a democratic culture.[3] Now people could become their own broadcasters, speaking to an indefinite number of people, not only in their own countries, but around the world. Armed with digital technologies, ordinary individuals could route around traditional media gatekeepers. They could also participate in culture in ever new ways through new digital forms of cultural production and appropriation. Digital speech, I argued, featured "routing around and glomming on," which were characteristic not only of digital speech, but of free speech generally.[4]

---

[1] Jack M. Balkin, *The Three Laws of Robotics in the Age of Big Data*, 78 OHIO ST. L.J. 1217, 1219 (2017) (defining the Algorithmic Society as "a society organized around social and economic decision-making by algorithms, robots, and AI agents").

[2] Jack M. Balkin, *Cultural Democracy and the First Amendment*, 110 NW. U. L. REV. 1053, 1061 (2016); Jack M. Balkin, *The Future of Free Expression in a Digital Age*, 36 PEPP. L. REV. 427, 438 (2009) [hereinafter Balkin, *Future of Free Expression*]; Jack M. Balkin, Commentary, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1, 3-4 (2004) [hereinafter Balkin, *Digital Speech and Democratic Culture*].

[3] *See, e.g.*, Reno v. ACLU, 521 U.S. 844, 850 (1997) (noting the Internet's ability to "enable tens of millions of people to communicate with one another and to access vast amounts of information from around the world").

[4] *See* Balkin, *Digital Speech and Democratic Culture*, *supra* note 2, at 9-12.

Nevertheless, I warned that the digital age would alter the relative importance of the First Amendment in securing freedom of speech, including the goal of protecting and promoting a democratic culture. The First Amendment, I argued, would prove increasingly inadequate to this task;[5] moreover, if courts interpreted the Constitution in a short-sighted manner, judge-made doctrines of the First Amendment would actually hinder the protection and development of a truly democratic culture.[6]

First, I argued that as the Internet developed, judge-made doctrines of the First Amendment, while necessary to the protection of a democratic culture, would prove increasingly insufficient. Much of the responsibility for securing a democratic culture would fall to private actors, technologists, legislatures, and administrative agencies.[7]

Second, I argued that as business models developed, companies would increasingly employ novel First Amendment theories not only to protect free speech generally, but also to restrict access to the digital infrastructure and forestall regulation in order to promote their business models and protect their profits.[8] To be sure, digital companies would often find themselves on the side of the values of a democratic culture. But just as often they would seek constitutional protection for novel forms of surveillance and control of individuals and groups.[9]

Neither judge-made doctrines of First Amendment law nor private companies will prove reliable stewards of the values of free expression in the twenty-first century.

This means that we must rethink the First Amendment's role in the digital era. On the one hand, the First Amendment retains its central purpose of guarding against state censorship through new devices of control and surveillance. On the other hand, courts should not interpret the First Amendment to prevent the state from regulating infrastructure companies in order to protect the values of a democratic culture and the ability of individuals to participate in the public sphere. Thus, the state, while always remaining a threat to free expression, also needs to serve as a necessary counterweight to developing technologies of private control and surveillance.

---

[5] *See id.* at 48-51.

[6] *See id.* at 19-20.

[7] *Id.* at 2; *see* Balkin, *Future of Free Expression*, *supra* note 2, at 432-33.

[8] Balkin, *Digital Speech and Democratic Culture*, *supra* note 2, at 20-22, 46-47; Balkin, *Future of Free Expression*, *supra* note 2, at 443-44.

[9] Balkin, *Digital Speech and Democratic Culture*, *supra* note 2, at 53; Balkin, *Future of Free Expression*, *supra* note 2, at 437-39.

The transition from the early days of the Internet to our present Algorithmic Society has only enhanced these concerns.

The Algorithmic Society features the collection of vast amounts of data about individuals and facilitates new forms of surveillance, control, discrimination and manipulation, both by governments and by private companies. Call this the problem of Big Data.[10]

The Algorithmic Society also changes the practical conditions of speech as well as the entities that control, limit, and censor speech. First, digital speech flows through an elaborate privately-owned infrastructure of communication. Today our practical ability to speak is subject to the decisions of private infrastructure owners, who govern the digital spaces in which people communicate with each other. This is the problem of private governance of speech.[11]

Nation states, understanding this, have developed new techniques for speech regulation. In addition to targeting speakers directly, they now target the owners of private infrastructure, hoping to coerce or coopt them into regulating speech on the nation state's behalf. This is the problem of "New School" speech regulation.[12]

In the digital age, individuals do not face the familiar dyadic model of speech regulation. In a dyadic model, there are two central actors: the power of the state threatens the individual's right to speak. Instead, the digital age features a pluralist model of speech control. In the pluralist model individuals may be controlled, censored, and surveilled both by the nation state and by the owners of many different kinds of private infrastructure, who operate across national borders in multiple jurisdictions. In fact, the largest owners of private infrastructure are so powerful that we might even regard them as special-purpose sovereigns. They engage in perpetual struggles for control of digital networks with nation states, who, in turn, want to control and coopt these powerful players. The practical ability to speak in the digital age is shaped by the results of these struggles for control and cooptation.[13]

In the Algorithmic Age, in short, the rights of free expression simultaneously face threats in multiple directions. Individuals face threats of control and surveillance by Big Data; and companies may try to use First Amendment arguments (inappropriately) to defend their power to surveil and control populations. Individuals also face threats

---

[10] *See infra* text accompanying notes 14–17.

[11] *See infra* text accompanying notes 90–95.

[12] *See infra* text accompanying notes 63–67.

[13] *See infra* text accompanying notes 103–18.

to freedom of expression by private governance and by new school speech regulation. In this world, the judge-made doctrines of the First Amendment, although still necessary, are inadequate to provide sufficient guarantees of free expression. Each of these threats, in its own way, imperils the values underlying the First Amendment — the promise of a vibrant democratic culture of participation and exchange. The first set of problems expands the First Amendment in the wrong direction; the second set of problems finds the First Amendment inadequate to protect online free expression.

The key issues of free speech theory in the Algorithmic Society flow from these problems. They concern Big Data and the use of algorithms and artificial intelligence to shape people's lives and opportunities. They also concern the new techniques that nation states employ to regulate online speech, and the rise of large private organizations that effectively govern how most people speak today online.

In this Essay, I will introduce some key concepts for understanding our rapidly changing free speech environment. Part I introduces the concepts of *information fiduciaries* and *algorithmic nuisance*. These concepts help us understand when the First Amendment should allow the state to regulate companies that engage in the collection, analysis, and distribution of data. Part II introduces the concepts of *new school speech regulation*, *public/private cooperation (and cooptation)*, and *private governance of speech*. These concepts help us understand how speech is governed in the current era and how we must supplement the First Amendment's guarantees to protect free expression.

I. BIG DATA IS SOYLENT GREEN

There is a saying in Silicon Valley that "Big Data is the new oil."[14] What do people mean by this? Big Data is crucial to the use and development of algorithms and artificial intelligence ("AI"). Algorithms and AI are the machines; Big Data is the fuel that makes the machines run. Just as oil made machines and factories run in the Industrial Age, Big Data makes the relevant machines run in the Algorithmic Society.

---

[14] Jonathan Vanian, *Why Data Is the New Oil*, FORTUNE (July 12, 2016), http://fortune.com/2016/07/11/data-oil-brainstorm-tech; *see* Michael Palmer, *Data Is the New Oil*, ANA MARKETING MAESTROS (Nov. 3, 2006), http://ana.blogs.com/maestros/2006/11/data_is_the_new.html; *The World's Most Valuable Resource Is No Longer Oil, but Data*, ECONOMIST (May 6, 2017), https://www.economist.com/news/leaders/21721656-data-economy-demands-new-approach-antitrust-rules-worlds-most-valuable-resource. Clive Humby is generally believed to have coined the phrase in 2006. *See* Palmer, *supra*.

The Algorithmic Society builds on the previous advances of the early Internet. The Algorithmic Society depends on huge databases that can cheaply and easily be collected, collated, and analyzed. But for this to happen, government and businesses need cheap computing, cheap telecommunications, cheap storage, and ways to describe data in easily computable and interoperable form. Above all, government and businesses need ways to collect huge amounts of data about the world and about people's activities.

The digital age makes all of this possible because digital communication involves creating data, copying it, storing it, and moving the copies from one place to another. In the digital age, more and more things that people say and do leave digital traces that can be collected, copied, collated, and analyzed. In short, the Digital Society gives birth to the Algorithmic Society. It created the technological platform on which algorithms and AI run.

There is a second reason why Big Data is the New Oil. Because oil was crucial to the Industrial Age, it was a source of wealth — and therefore a source of power. New forms of wealth emerge in the Digital Age just as they did in the Industrial Revolution. Four especially important forms of wealth in the Information Age are intellectual property, fame, information security, and Big Data. Big Data includes personal data and data about the world collected from sensors and programs that are everywhere. It also includes metadata — data about the data that is collected.

The goal of the Algorithmic Society is practical omniscience: that is, the ability to know as much as possible about who is doing what, when, and where; and the ability to predict who will do what, when, and where. But the possibility of practical omniscience raises an important question of politics: who will possess and control this omniscience? Behind the machines are people: governments, businesses, and organizations. Through its technologies of information collection, use, analysis, and control, the Algorithmic Society creates a new economy of power.

The early Internet gave people hope that new technology would level political and social hierarchies by spreading political, organizational, and communicative powers widely among the world's population. People hoped that the Internet and its related technologies would be a democratizing force. To some extent, that has been the case. But it is not quite how things have turned out.

We should make a key distinction between *distributed* and *democratic* power. A form of power is *democratic* if many people participate in it and participate in decisionmaking about how to

employ it. A form of power is *distributed* if it operates in many different places and affects many different people and situations. In some ways the Internet and its associated digital technologies have made power more democratic. But in other ways the Internet has made it possible for power to be widely distributed but not democratic.

For example, through its App Store, Apple has the ability to control iPhones and iPads around the world. This control is made possible by the Internet. The Internet connects each person's iPhone or iPad to Apple's servers. This form of power is distributed but not democratic.[15]

In the Algorithmic Society, surveillance and data collection are now widely distributed, but there is no guarantee that they will be democratically controlled. Data about many people are collected in many places, but a relatively small number of people have the resources and the practical ability to collect, analyze, and use this data.

This asymmetry is crucial to understanding the political challenges of the Algorithmic Society. Big Data collects and analyzes information about people — their locations, actions, characteristics, and behaviors. But the people whose information is collected are not necessarily the people who control the information. Quite the contrary: information about the world's populations serves as grist for the mill of computation, analysis, and decisionmaking by governments and large corporations. Big Data enables new ways of classifying people, making decisions about them, and exercising power over them. In this sense, Big Data is not only the New Oil. Big Data is also Soylent Green.

Soylent Green, for those who may not remember, is the name of a 1973 science fiction film starring Charlton Heston.[16] It is set in 2022, only a few years from now. The world is suffering from food shortages, and the Soylent Corporation comes up with a nutritious wafer, soylent green, which can feed the world. The corporation says it is made from "high energy plankton," but we learn that this is a lie. The most famous line in the movie comes at the very end, when Charlton

---

[15] *See, e.g.*, JONATHAN L. ZITTRAIN, THE FUTURE OF THE INTERNET — AND HOW TO STOP IT 106-11 (2008) (noting problems of injustice and inequality of power created by tethered devices and perfect enforcement from a distance); Julie E. Cohen, *Pervasively Distributed Copyright Enforcement*, 95 GEO. L.J. 1, 3 (2006) ("Pervasively distributed copyright enforcement invades, disrupts, and casually rearranges the boundaries of personal spaces and of the intellectual and cultural activities played out within those spaces.").

[16] SOYLENT GREEN (Metro-Goldwyn-Mayer Studios, Inc. 1973).

Heston calls out, as the authorities take him away, "It's people! Soylent Green is made out of people."[17]

This brings me to the first major idea in this Essay. Big Data is not simply a vast new source of wealth, or the fuel that runs the Algorithmic Society. Big Data is Soylent Green. Big Data is people.

A. Technology Mediates (and Constitutes) Relationships of Power Between People

In this case, no one is making people into food. But data about people are being made into fuel. Data about people — their activities, transactions, locations, and preferences — powers the engines of the Algorithmic Society. Data makes possible new, interesting, and increasingly powerful forms of algorithmic computation that yield new insights about human behavior and increase opportunities for prediction, risk management, and control. Equally important, data about people is a central method of governance and control over large populations of people, determining their opportunities and their fates.

When I say that Big Data is Soylent Green, I do not merely mean that data scientists use Big Data to study people and their relationships. I mean that the actual practices of collecting, analyzing, and using Big Data for governance and control involve relationships of power between people. That is true even if all of the data processing and decisionmaking is done by algorithms, AI agents, and robots. Indeed, it is especially true in these situations.

When people think about the Algorithmic Society and the rise of robotics and artificial intelligence, they tend to worry about their relationship to these *technologies*. They worry that these technologies will have power over them or displace them.[18]

But behind the algorithm, the artificial intelligence agent, and the robot is a government, a company, or some group of persons, who are using the technology to affect people's lives. The technology — the collection and analysis of Big Data plus its use in decisionmaking — mediates a relationship of power between the people whose data is

---

[17] BradZ1, *IT'S PEOPLE!*, YOUTUBE (Nov. 19, 2007), https://www.youtube.com/watch?v=8Sp-VFBbjpE.

[18] *See, e.g.*, Sarah Griffiths, *Do You Fear AI Taking Over? A Third of People Believe Computers Will Pose a Threat to Humanity and More Fear They'll Steal Jobs*, DAILY MAIL (Mar. 11, 2016, 2:30 PM), http://www.dailymail.co.uk/sciencetech/article-3487851/Do-fear-AI-taking-people-believe-computers-pose-threat-humanity-fear-ll-steal-jobs.html#ixzz4rkBbKHVo; Patrick Thibodeau, *One in Three Developers Fear A.I. Will Replace Them*, COMPUTERWORLD (Mar. 8, 2016, 12:55 PM), https://www.computerworld.com/article/3041430/it-careers/one-in-three-developers-fear-ai-will-replace-them.html.

being collected, and about whom decisions are being made, and the government, company, or people who operate the technology.[19]

This is a general point about technology. We tend to associate power with the effects of technology itself. But technology is actually a way of exemplifying and constituting relationships of power between one set of human beings and another set of human beings. This was true even of the technology of writing, which, Claude Levi-Strauss famously asserted, was used to organize the labor of slaves.[20] It is true today in the development of decisionmaking by algorithms and AI agents.

Debates about robots and artificial intelligence tend to center on whether robots and AI agents themselves are dangerous, whether they will break free from human control, assert themselves, take over the world, and cast humans aside.[21] This is the trope of the Frankenstein monster, or Skynet in the Terminator movies. This way of thinking deflects our attention from what is actually the most important issue. The question is not the robots; it is the people and companies behind the robots that we should be concerned with.[22]

Technology mediates relationships of power between human beings and other human beings. Behind robots, AI agents, and algorithms are people and companies. They use these technologies to make decisions about and govern populations of human beings. Human beings create the technologies that human beings use to achieve power over and govern other human beings. In this respect, robots and AI agents are no different from many previous technological innovations.

What is especially interesting about the Algorithmic Society, however, is the way that new systems of governance and control arise out of data collection, transmission, and analysis. The Industrial Age fought over freedom of contract and the rights of property; the Algorithmic Age is a struggle over the collection, transmission, use, and analysis of data. For this reason, the central constitutional questions do not concern freedom of contract. They concern freedom of expression.

---

[19] Jack M. Balkin, *The Path of Robotics Law*, 6 CALIF. L. REV. 45, 52-54 (2015); Balkin, *supra* note 1, at 1219-20.

[20] CLAUDE LEVI-STRAUSS, TRISTES TROPIQUES 299 (John Weightman & Doreen Weightman trans., Penguin 1973) (1955) ("[T]he primary function of written communication is to facilitate slavery.").

[21] *See, e.g.*, Rory Cellan-Jones, *Stephen Hawking Warns Artificial Intelligence Could End Mankind*, BBC (Dec. 2, 2014), http://www.bbc.com/news/technology-30290540.

[22] Balkin, *supra* note 1, at 1223 ("[T]he problem is not the robots; it is the humans.").

Suppose that the government attempts to regulate the collection, analysis, and use of data by companies that use AI and algorithms. Their natural response will be that such regulation violates the First Amendment.[23] All that companies are doing is collecting, analyzing, using, and distributing information. Information is speech, and speech is protected by the First Amendment. In addition, the manipulation, analysis, and distribution of information are just different forms of expression. Therefore, the argument goes, the First Amendment protects the collection, collation, use and distribution of data in the Algorithmic Society.

When people think about robots, AI, and the First Amendment, they naturally imagine that the central question is whether the speech of robots and AI agents is entitled to First Amendment protection. This is certainly an interesting question — and it has produced a growing literature.[24] But it obscures a deeper issue. Behind the robot there is always somebody who designs, programs, manufactures, and implements the robot, who collects data for and from the robot, and who uses the robot for surveillance, decisionmaking, governance, and control. The most important question is not whether robots have First Amendment rights; it is whether companies will be able to shield themselves from regulation by claiming that their uses of AI agents, robots, and algorithms are First Amendment protected activities.

Behind these technologies are people and organizations who will want to use the First Amendment as a deregulatory tool to protect

---

[23] For thoughtful attempts to explain why this follows from basic First Amendment principles, see Jane Bambauer, *Is Data Speech?*, 66 STAN. L. REV. 57, 84-86 (2014) (explaining why the right to collect and create information suggests a broad right to record). *See also* Eugene Volokh, *Freedom of Speech and Information Privacy: The Troubling Implications of a Right to Stop People from Speaking About You*, 52 STAN. L. REV. 1049, 1050-51 (2000) ("[T]he right to information privacy — my right to control your communication of personally identifiable information about me — is a right to have the government stop you from speaking about me."); *cf.* Jane R. Bambauer & Derek E. Bambauer, *Information Libertarianism*, 105 CALIF. L. REV. 335, 357 (2017) (offering a general theory against regulation of information). Bambauer and Bambauer are careful to note, however, that states can still regulate conduct, so that governments might still regulate certain uses of data. *See id.* at 340.

[24] *See generally* EUGENE VOLOKH & DONALD M. FALK, GOOGLE: FIRST AMENDMENT PROTECTION FOR SEARCH ENGINE SEARCH RESULTS (2012), http://volokh.com/wp-content/uploads/2012/05/SearchEngineFirstAmendment.pdf; Stuart Minor Benjamin, *Algorithms and Speech*, 161 U. PA. L. REV. 1445 (2013); James Grimmelmann, *Speech Engines*, 98 MINN. L. REV. 868 (2014); Toni M. Massaro & Helen Norton, *Siri-ously?: Free Speech Rights and Artificial Intelligence*, 110 NW. U. L. REV. 1169 (2016); Toni M. Massaro, Helen Norton & Margot E. Kaminski, *Siri-ously 2.0: What Artificial Intelligence Reveals About the First Amendment*, 101 MINN. L. REV. 2481 (2017); Tim Wu, *Machine Speech*, 161 U. PA. L. REV. 1495 (2013).

business practices that affect the lives of many other people. In the United States, at last, the constitutional question is whether companies in the Second Gilded Age will be able to use the First Amendment guarantees of speech and press in the same way that industrial organizations used the constitutional idea of freedom of contract in the First Gilded Age.[25] Are the power relationships of the Algorithmic Age safeguarded by the U.S. Constitution and protected by the First Amendment? Which business practices are shielded — and should be shielded — from government regulation by the First Amendment is the truly important question of the Algorithmic Society.

Two key ideas help us understand when the First Amendment permits legal regulation of the people and organizations that use Big Data, algorithms, and artificial intelligence. The first is the concept of information fiduciaries. The second is the concept of algorithmic nuisance.

B. Information Fiduciaries

Some enterprises are *information fiduciaries* toward their end-users. Governments can impose reasonable regulations on how information fiduciaries collect, use, distribute, and sell information derived from their fiduciary relationships with end-users. In general, information collected in the context of a fiduciary relationship is not part of public discourse. Therefore, government can regulate the uses of this information consistent with the First Amendment.[26]

Fiduciary relationships involve asymmetries of power, information, and transparency.[27] The fiduciary collects sensitive information about the client that might be used to the client's disadvantage. The client is relatively transparent to the fiduciary, but the fiduciary is not transparent to the client. By this I mean that the client is not well-equipped to understand and monitor the fiduciary's operations. Moreover, the client relies on the fiduciary to perform valuable services, which the client cannot easily perform for themselves.

Because of these asymmetries, clients have to trust fiduciaries and hope that the latter will not betray them. Fiduciaries, in turn, must act

---

[25] Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 UC DAVIS L. REV. 1183, 1185-86 (2016) [hereinafter Balkin, *Information Fiduciaries*]; *see* Balkin, *Digital Speech and Democratic Culture*, *supra* note 2, at 25-28.

[26] Balkin, *Information Fiduciaries*, *supra* note 25, at 1217-18.

[27] TAMAR FRANKEL, FIDUCIARY LAW xvi, 4, 6, 18, 29 (2011) (noting the role of asymmetries of power and knowledge in fiduciary relationships); Balkin, *Information Fiduciaries*, *supra* note 25, at 1216-17.

in good faith toward their clients, particularly with respect to the information they learn about their clients in the course of the relationship. Where the fiduciary relationship involves the collection and use of significant information about the client, we can speak of information fiduciaries.[28]

The classic examples of information fiduciaries are doctors and lawyers.[29] Both collect lots of personal information about their clients, their operations are not transparent to relatively untrained clients, and clients' ability to monitor professionals is limited by their lack of training. Clients disclose sensitive information because they need particular services, and as a result, they must trust doctors and lawyers. Doctors and lawyers, in turn, have obligations to look out for their clients' and patients' interests, not to create conflicts of interest with them, and not to disclose information about them that might be used to their disadvantage. Above all, professional fiduciaries must act in good faith toward their clients and patients.[30]

The First Amendment does not prevent the state from regulating how professionals interact with their clients and how they use their clients' information. That is because professionals have a fiduciary relationship with their clients.[31] Professionals give information to their clients within a trusted relationship; conversely, clients give information to professionals in confidence. Neither kind of information is part of public discourse; therefore, states may regulate exchanges of information between fiduciaries and clients in ways that they are not usually permitted to regulate ordinary public discourse.[32] For this reason, the concept of information fiduciaries is crucial to understand the appropriate boundaries of regulation that is consistent with the First Amendment.

---

[28] NEIL RICHARDS, INTELLECTUAL PRIVACY: RETHINKING CIVIL LIBERTIES IN THE DIGITAL AGE 168 (2015); Balkin, *Information Fiduciaries*, *supra* note 25, at 1209-10.

[29] Balkin, *Information Fiduciaries*, *supra* note 25, at 1207-09.

[30] *See* RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS §§ 16, 49, 60 (2000) (stating lawyers' fiduciary duties to respect client confidences and to act in the client's interests); MARK A. HALL, MARY ANNE BOBINSKI & DAVID ORENTLICHER, MEDICAL LIABILITY AND TREATMENT RELATIONSHIPS 171 (3d ed. 2013) (stating fiduciary duties of physicians).

[31] Balkin, *Information Fiduciaries*, *supra* note 25, at 1210-11; *cf.* ROBERT C. POST, DEMOCRACY, EXPERTISE, AND ACADEMIC FREEDOM: A FIRST AMENDMENT JURISPRUDENCE FOR THE MODERN STATE 23 (2012) ("Whereas within public discourse the political imperatives of democracy require that persons be regarded as equal and as autonomous, outside public discourse the law commonly regards persons as dependent, vulnerable, and hence unequal.").

[32] Balkin, *Information Fiduciaries*, *supra* note 25, at 1210-11.

Who are the new information fiduciaries in the digital age? They are organizations and enterprises who collect enormous amounts of information about their end-users.[33] End-users are transparent to these organizations, but their operations are not transparent to end-users, and it is difficult if not impossible to monitor their operations. As a result, these organizations enjoy significant asymmetries of knowledge and power over their end-users. These businesses encourage their end-users to trust them and disclose information to them, and end-users must trust them in order to benefit from the services these organizations provide.

This means that many of the digital organizations that people deal with every day — including Internet service providers ("ISPs"), search engines, and social media platforms — should be treated as information fiduciaries with respect to their clients and end-users. Therefore, consistent with the First Amendment, governments can subject the information fiduciary to reasonable restrictions on collection, collation, analysis, use, sale, and distribution of personal information.

I hasten to add that the scope of constitutionally permissible regulation is and should be narrower in the case of digital enterprises than the permissible scope of regulation of doctors and lawyers. The business of a social media platform or Internet service provider is quite different from the business of a doctor or a lawyer, and the degree of reasonable trust that end-users have in digital enterprises is also different.[34] First, although monetizing personal data is not central to the business of doctors and lawyers, it is central to many online service companies, which allows them to subsidize the services they perform for end-users. Second, unlike doctors and lawyers, search engines and social media sites have an interest in getting people to express themselves as much as possible publicly so that they will produce content and data that can be indexed or analyzed, even though people may regret their choices later on. Third, people expect doctors to do more than not harm them; they also expect that doctors will look out for them and warn them about potential risks to their health, their diet, and so on. People do not expect such comprehensive

---

[33] *Id.* at 1221-26 (arguing that the standard justifications for imposing fiduciary obligations on older businesses and professions also apply to certain digital enterprises).

[34] *Id.* at 1226-31 (describing limited nature of fiduciary obligations of digital information fiduciaries).

obligations of care from their ISPs, search engines, and social media sites.[35]

Because of these differences, digital information fiduciaries should have fewer obligations than traditional professional fiduciaries like doctors, lawyers, and accountants. They are special-purpose information fiduciaries, and the kinds of duties that it is reasonable to impose on them should depend on the nature of the services they provide.[36] Their central obligation is that they cannot act like con artists — inducing trust in their end-users to obtain personal information and then betraying end-users or working against their interests.[37]

C. Algorithmic Nuisance

The idea of a fiduciary obligation presupposes a contractual (or quasi-contractual) relationship between a client and some entity or business, whether it is Facebook or a personal physician. The contractual relationship is part of a relationship of trust that gives rise to fiduciary obligations. End-users usually have a contractual relationship with their digital information fiduciaries. That is because digital companies normally require end-users to agree to an end-user license agreement, terms of service agreement, or membership agreement when they create an account and sign up for a digital service or join the digital community that the company provides.[38]

On the other hand, there are a wide range of situations in which people lack a contractual relationship with a digital enterprise or with a business that collects personal information and uses algorithms to make decisions. For example, consider people applying for a job, for credit, or for a mortgage. Applicants seek an opportunity, and the business that decides whether to give them that opportunity relies on

---

[35] *Id.* (describing these three differences between digital information fiduciaries and traditional fiduciaries).

[36] *Id.* at 1229-30 ("What is unexpected or seems like a breach of trust will depend on the kind of service that entities provide and what we would reasonably consider unexpected or abusive for them to do.").

[37] *Id.* at 1225-26 ("Digital information fiduciaries may be held to reasonable ethical standards of trust and confidentiality, even if they do not make specific representations, because of the nature and kind of business they are in.").

[38] *See, e.g.*, *Google Terms of Service*, GOOGLE (Apr. 14, 2014), https://www.google.com/policies/terms; *Statement of Rights and Responsibilities*, FACEBOOK (Jan. 30, 2015), https://www.facebook.com/legal/terms; *Terms of Use*, INSTAGRAM (Jan. 19, 2013), https://help.instagram.com/478745558852511; *Twitter Terms of Service*, TWITTER (Oct. 2, 2017), https://twitter.com/en/tos; *U.S. Terms of Use*, UBER (Mar. 23, 2017), https://www.uber.com/legal/terms/us.

a Big Data company or an algorithm that uses Big Data in order to make the decision.

In such a case, there is no preexisting contractual relationship. It is a transaction between relative strangers. The applicant does not sign up for a digital service, and does not join a community like Facebook or Twitter. In such cases, one cannot employ the concept of information fiduciaries to justify regulation, because there is no fiduciary relationship — at least before a relationship is formed. An information fiduciary may not betray or abuse the trust of its end-users. But not all relationships in the Algorithmic Society involve information fiduciaries because not all relationships are contractual relationships of trust with end-users.[39]

Instead, we must turn to a different idea. This is the idea of algorithmic nuisance. The concept of algorithmic nuisance applies when companies use Big Data and algorithms to make judgments that construct people's identities, traits, and associations that affect people's opportunities and vulnerabilities. Opportunities include things like employment, credit, financial offers, and positions. By vulnerabilities, I mean increased public or private surveillance, discrimination, manipulation, and exclusion.

The concept of algorithmic nuisance stems from the fact that companies collect data about people from multiple sources and use algorithms to make decisions about people. Through this process, companies do more than simply make decisions. They also construct people's digital identities, traits, and associations, which, in turn, construct (and constrict) their future opportunities.

These collections of data, and the resulting digital constructions of traits, associations, and identity, may be employed by still other companies in ever new contexts of judgment. For example, companies might use credit scores in models predicting all sorts of behavior. People's lives become subject to a cascade of algorithmic judgments, each constructing their identity and opportunities over time. You might think of our digital identities as an informational stream into which a collection of new judgments, scores, and risk assessments are constantly being tossed.

This cascade of judgments increasingly shapes people's lives as more and more businesses participate in the collective process of digital identity shaping. This process predictably throws unjustified costs

---

[39] Balkin, *Information Fiduciaries*, *supra* note 25, at 1232-34; Balkin, *supra* note 1, at 1226-27.

onto the people who are the subjects of this system of digital identity in the form of constricted opportunities and increased vulnerabilities.

The concept of algorithmic nuisance tries to capture the idea of companies externalizing the socially unjustified costs of algorithmic decisionmaking onto the populations whose lives are shaped — and opportunities constricted — by algorithmic judgments.[40]

Although these businesses use data and share data, the First Amendment does not prevent regulation of how they make and implement their decisions. That is because permissible regulation aims at the outputs of algorithmic decisionmaking: discrimination and manipulation.[41]

The idea of algorithmic nuisance builds on an analogy to the common law concepts of public and private nuisance. Common law nuisances are non-trespassory invasions that interfere with the quiet use and enjoyment of a person's interests in real property.[42] In economic terms, action for nuisance are justified in cases of market failure when companies can externalize the costs of their operations onto strangers.[43] Pollution is the obvious example. Private nuisances normally involve situations in which the externalization of costs harms a (relatively small) class of people who are harmed in distinctive ways.[44] Public nuisances involve costs spread over a wide range of people; in these cases, the state must decide whether to bring

---

[40] Balkin, *supra* note 1, at 1235-36.

[41] Balkin, *Information Fiduciaries*, *supra* note 25, at 1194, 1212-13 (noting that market behavior regulated in antitrust law, consumer protection law, and antidiscrimination law is not protected by the First Amendment simply because it uses speech); Neil M. Richards, *Reconciling Data Privacy and the First Amendment*, 52 UCLA L. REV. 1149, 1181-82 (2005) (noting that use restrictions are consistent with the First Amendment).

[42] *See* RESTATEMENT (SECOND) OF TORTS § 821A cmt. b (1979) (defining nuisance as "human activity or a physical condition that is harmful or annoying to others"); *id.* § 821D & cmt. a (noting that private nuisance has traditionally been concerned with non-trespassory invasions of interests in the use and enjoyment of land).

[43] *See* Keith N. Hylton, *The Economics of Public Nuisance Law and the New Enforcement Actions*, 18 SUP. CT. ECON. REV. 43, 43-44, 55 (2010) ("Nuisance law induces actors to choose socially optimal activity levels by imposing liability when externalized costs are far in excess of externalized benefits or far in excess of background external costs.").

[44] *See* RESTATEMENT (SECOND) OF TORTS § 821E & cmt. a (1979) ("The liability for private nuisance exists only for the protection of persons having 'property rights and privileges,' that is, legally protected interests, in respect to the particular use or enjoyment that has been affected.").

a judicial action to abate the nuisance or pass new laws or new regulations to respond to the problem.[45]

A nuisance may be socially unjustified because the activity is unjustified as a matter of cost-benefit analysis — a problem of negligence law. But often an activity becomes a nuisance because the defendant engages in too much of the activity. Once again, pollution is the standard example.[46]

How do we apply these ideas to the use of Big Data? Businesses use algorithms and ratings systems derived from algorithms to make decisions about who gets what opportunity — credit, a job, or entrance to and exclusion from any number of different benefits. In order to make these decisions, businesses increasingly rely on Big Data and algorithms, because so many decisions have to be made and it is too costly to engage in individualized decisionmaking.[47]

Some algorithms are negligently constructed. Businesses may use biased or skewed data, the models may be badly designed, or the company's implementation and use of the algorithm may be faulty. In these situations, we have ordinary negligence.

But the more interesting situations are those in which an enterprise can make a plausible argument that the algorithmic model they are using is reasonable for *them* to use, given their particular goals and their need to lower the costs of decisionmaking. For example, suppose a business uses an algorithmic model that generates many false positives (people who would be suitable for an opportunity but are turned away) but very few false negatives (people who are unsuitable for the opportunity but are accepted). This model is reasonable for them to use because they only want to fill a certain number of positions.

---

[45] *See* RESTATEMENT (SECOND) OF TORTS § 821B(1) (1979) (defining public nuisance as an "unreasonable interference with a right common to the general public"); *id.* § 821C (public officials must bring suits to abate a public nuisance unless a private individual suffers a harm different in kind from that suffered by the general public); *id.* § 821B(2)(b) & cmt. c (noting that legislatures and administrative agencies may determine that certain conduct constitutes a public nuisance, thereby obviating the need for an additional showing of unreasonable interference); *id.* §§ 821B(2)(a), 821D, 821D cmt. b (public nuisance is not limited to invasions of interests in real property, but may be concerned with broader matters such as public health, safety, or morals).

[46] *See* Hylton, *supra* note 43, at 48 (noting that nuisance focuses on costs produced by excessive activity levels as well as lack of due care).

[47] *See* Andrew McAfee & Erik Brynjolfsson, *Big Data: The Management Revolution*, HARV. BUS. REV. (Oct. 2012), https://hbr.org/2012/10/big-data-the-management-revolution (explaining that big data increases the volume, velocity, and variety of decisionmaking at lower cost).

Even so, using algorithms to make decisions in areas like policing, employment, housing, and finance may cumulatively disadvantage many people in the long run. That is especially so if organizations further economize by building on algorithmic judgments made by other companies about people's attributes, trustworthiness, and reputation. Using algorithms repeatedly and pervasively over large populations of people may inappropriately treat people as risky or otherwise undesirable, impose unjustified burdens and hardships on populations, and reinforce existing inequalities.[48]

The idea behind algorithmic nuisance is that algorithmic decisionmaking has cumulative side effects on populations as more and more public and private businesses adopt it.[49] Algorithms construct people's identities and reputations by classifying them as risky, associating them with undesirable traits or correlations, or placing them in the same categories as other people who are risky or have undesirable characteristics. In other words, algorithms construct digital portraits of people and lump people into digital categories created by the algorithms. Algorithms work through constructing categories for differentiation and imposing those categories and differences onto populations of people. They bestow characteristics on people and create digital identities and reputations for them.[50]

These categories and differences have social force because businesses rely on them to create and withhold opportunities. Businesses may further economize by purchasing databases and importing algorithmic judgments from other companies, thus spreading people's algorithmic reputations and identities widely throughout society.

What are the possible dangers of creating and proliferating digital reputations? First, algorithmic decisionmaking may propagate discrimination over many different aspects of people's lives. Second, it may create opportunities for businesses to manipulate people. Third, it may produce incentives for people to conform their lives to the requirements of algorithms to avoid their judgment. All of this will happen without transparency, accountability, due process, or the ability to monitor and respond to the decisions.[51]

---

[48] Balkin, *supra* note 1, at 1231-32.

[49] *Id.* at 1232.

[50] *Id.* at 1235-37.

[51] *Id.* at 1239 (explaining that the potential dangers of algorithmic decisionmaking are that "algorithms (a) construct identity and reputation through (b) classification and risk assessment, creating the opportunity for (c) discrimination, normalization, and manipulation, without (d) adequate transparency, accountability, monitoring, or

When companies employ algorithms to satisfy particular organizational needs, they do not always consider how their work shapes digital reputations that may be employed in a wide range of other situations. Because algorithmic judgments may be shared widely in multiple contexts, they can have systemic and cumulative effects. Algorithmic decisionmaking contributes to the digital construction of identities with effects that go well beyond any business's particular decisions. This is like throwing just a little bit of pollutant in a river that everyone else uses. Over time, the consequences are significant.

Hence, there is an analogy to nuisance. Increased activity levels may increase unjustified social costs, even when the activity is conducted non-negligently. Nuisances may emerge when businesses ramp up production as they shift to new technologies.[52] That is what is happening in the Algorithmic Society. Algorithms lower the costs of judgment and therefore increase the amount, rapidity, and spread of judgment, affecting more lives and reputations more quickly, more cheaply, and more pervasively.

All judgments are imperfect and the quality of a judgment is relative to the purposes for which it is being used. The more that businesses engage in judgments that previously would have been prohibitively costly — or even impossible — the more they increase the side effects of judgment and decisionmaking. These side effects of algorithmic decisionmaking are like increased levels of pollution as companies invest more heavily in industrial production. The appropriate remedy is to make companies internalize the costs they shift onto others and onto society as a whole as they employ algorithmic decisionmaking.

D. Personal Robots

Let me give you an example of these ideas in practice. The example is personal robots, by which I also mean to include home robots, smart homes, and digital appliances like Alexa.

---

due process"). On the problems of ensuring due process, accountability and transparency in Big Data operations, see Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 677-92, 718-19 (2016); Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1256 (2008); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 5-6, 18-20 (2014); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 94-101, 121-28 (2014); Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857, 881 (2017); Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 640 (2017).

[52] Hylton, *supra* note 43, at 48 (noting that employing new technologies may increase activity levels, leading to nuisance).

The Algorithmic Society is producing a revolution in the idea of consumer protection. The paradigmatic example of a twentieth century consumer protection problem is a Coke bottle that explodes in your face. The paradigmatic example of a twenty-first century consumer protection problem is a Coke bottle that spies on you.[53]

More generally, the problem is that consumer objects — including appliances, cars, and houses — collect information about you, listen to everything you are doing, and then report back to the corporation that manufactures and services them. The regulatory challenge of the early twenty-first century is the toaster that betrays you.[54]

This shift in focus reveals the value of concepts like information fiduciaries and algorithmic nuisance. We need to describe the emerging relationship between ourselves and a new class of entities that use algorithms and AI not only to provide us with services but also to study us and make decisions about us.

Now consider the personal robot. It need not look like a humanoid artifact. It might also be a smartphone, a smart car, a home appliance like Alexa, a home climate system, or even a toaster. Today people are surrounded by personal robots they may not even recognize as such. As time goes on, an increasing number of the personal devices, appliances, and facilities that people use every day — to travel, to communicate with friends and family, to perform tasks around the house — will be equipped with various degrees of artificial intelligence. Personal robots will use algorithms and AI to make decisions and provide services, all the while collecting and remembering information about the people who use them to perform those services. These devices will often be connected to the Internet and report what they see, hear, and do to businesses and corporations.

What is the relationship between these new devices and the concepts of information fiduciary and algorithmic nuisance? Begin with algorithmic nuisance. Smart devices constantly make decisions about what kinds of services to offer you, and what kind of

---

[53] Jack M. Balkin, *The Difference Between 20th and 21st Century Consumer Protection*, BALKINIZATION (Feb. 8, 2017), https://balkin.blogspot.com/2017/02/the-difference-between-20th-and-21st.html.

[54] *See, e.g.*, M. Ryan Calo, *Robots and Privacy*, *in* ROBOT ETHICS: THE ETHICAL AND SOCIAL IMPLICATIONS OF ROBOTICS 187, 187-89 (Patrick Lin et al. eds., 2012) (arguing that one of the most likely uses of robots will be as surveillance devices); Woodrow Hartzog, *Unfair and Deceptive Robots*, 74 MD. L. REV. 785, 790-96 (2015) (describing how robotic communication may involve fraud, manipulation, and invasions of privacy); Margot E. Kaminski, *Robots in the Home: What Will We Have Agreed To?*, 51 IDAHO L. REV. 661, 661-63 (2015) (noting problems of privacy and consent when robots are introduced into the home).

opportunities you will get; moreover, they constantly construct and revise your digital identity and send this information to other businesses who will, in turn, make judgments and decisions about you that will be folded into your digital dossier. This creates problems of algorithmic nuisance.

Personal or home robots are also a good example of the principle of information fiduciaries. Companies are starting to build intelligent functions into houses, and to place home robots, like Alexa, in people's houses.[55] These home robots are also methods of home surveillance. Alexa knows how to serve because it records and remembers what you ask it. Indeed, in theory, it can remember everything you have ever said to it or around it.

But Alexa is not simply a robot; it is a cloud robot. What you say to it, and what it hears you say, does not stay in your home; it is uploaded and stored in the servers of the business that runs the artificial intelligence system. When you speak to Alexa, you are not simply speaking to an appliance on your kitchen table. You are speaking to a corporation. That corporation, which is privy to the most intimate details of your life, should have a fiduciary duty to deal with you in a trustworthy fashion. It should be considered an information fiduciary.

Many people may recall television shows like Downton Abbey and P.G. Wodehouse's Jeeves and Wooster series of short stories and novels. These stories center on the crucial role of the British butler — or more correctly valet — and their relationships to the aristocrats that employ them.

There is an old saying: no man is a hero to his own valet. Jeeves works for a British upper-class twit named Bertie Wooster. He sees every stupid and venal thing that Bertie does, and as P.G. Wodehouse makes clear, Bertie Wooster is a bit of an idiot. He is always getting into various scrapes and Jeeves has to get him out of them. Fans of Downton Abbey may love the Crawley family, but the Crawleys also are always getting into trouble, and some of them misbehave a great deal. (And however much the butler, Carson, adored Lady Mary, even he had to agree that she was a bit much on occasion.)

---

[55] *See* Jean-Baptiste Coumau, Hiroto Furuhashi & Hugo Sarrazin, *A Smart Home Is Where the Bot Is*, MCKINSEY Q. (Jan. 2017), http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/a-smart-home-is-where-the-bot-is ("Within a decade, our living spaces will be enhanced by a host of new devices and technologies, performing a range of household functions and redefining what it means to feel at home.").

Just as Jeeves was always hovering over Bertie Wooster, the staff of Downton Abbey sees everything that the Crawleys are doing. Over time, like Jeeves, they collect a great deal of personal information that could be used to the detriment of their employers. They were expected to observe, and serve, and not to wash their employer's dirty laundry in public. In other words, they were expected to have a fiduciary relationship to their employers, whether or not they lived up to it.

We do not live in that world today, but increasingly we are living in a sort of twenty-first century version of Downton Abbey. In this world, ordinary people play the roles of the lords and ladies; their servants are Alexa and Siri and the computer in their self-driving cars or home climate systems. The personal robot is our digital valet, and just as no one is a hero to his or her valet, no one is a hero to his or her own personal robot either.

Perhaps even more important, our new digital servants are not self-contained. Behind the robot is a company, a corporation, that designed and implements algorithms and artificial intelligence, and that collects information and analyzes it in order to provide services and opportunities to the people who employ our personal robots. The digital version of Jeeves or Carson is an elaborate technological system — sensors, microphones, cameras, and data collection equipment in constant communication to a series of server farms, algorithms, and artificial intelligence agents, operated by and for business corporations — although its apparent physical manifestation in a consumer's home may be a little box or a seemingly harmless gizmo with a cute little face. Indeed, the physical manifestation may be our homes themselves. When we are most at home, we may be most under surveillance.

To deal with this new organization of consumer products and services, we need the concepts of information fiduciary and algorithmic nuisance. Home robots and smart appliances collect an enormous amount of information about us which, in theory, can be collated with information about many other people that is stored in the cloud. Home robots and smart appliances are always-on, interconnected cloud entities that rely on and contribute to huge databases. Although we may come to trust the home robot and the smart appliance — indeed, we have to — the entity that we really have to trust is not the robot or the appliance. It is the company behind the robot and the appliance that collects the data and makes the decisions. And that company, I argue, should be an information fiduciary.

II. NEW SCHOOL SPEECH REGULATION AND PRIVATE GOVERNANCE

The second set of issues is symbolized by the ideas of "the right to forget" and "fake news." These two issues may seem unrelated. In fact, they are about the same issue: a fundamental change in how freedom of speech is regulated in the digital era. This alteration in governance has two key elements. The first is a change in how governments regulate — or attempt to regulate — speech in the digital era, from "old school" to "new school" speech regulation. The second is that privately owned online platforms engage in private governance of speech.

A. The Right to Be Forgotten and the Problem of Fake News

The "right to be forgotten" is a doctrine of European data protection law.[56] Recently, the doctrine has been applied to search engines on the grounds that they are information processors of personal data. The European Court of Justice has held that the right to forget requires search engines to remove links to webpages containing information about people that is "inadequate, irrelevant [,] . . . no longer relevant, or excessive" to ensure that other people will not have easy access to the information.[57]

The problem of "fake news" arose in the context of the 2016 American presidential election. People were concerned about the distribution of propaganda and false stories spread by individuals, organizations, and armies of bots through social media like Facebook and Twitter.[58] These stories and propaganda were designed to sow confusion and disinformation about the candidates.[59]

---

[56] *See* Robert Post, *Data Privacy and Dignitary Privacy:* Google Spain, *the Right to Be Forgotten, and the Construction of the Public Sphere*, 67 DUKE L.J. (forthcoming 2018) (manuscript at 8-9), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2953468.

[57] Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos, ECLI:EU:C:2014:317 ¶ 94 (May 13, 2014), http://curia.europa.eu/juris/document/document_print.jsf?doclang=EN&docid=152065.

[58] *See, e.g.*, Craig Silverman & Lawrence Alexander, *How Teens in the Balkans Are Duping Trump Supporters with Fake News*, BUZZFEED (Nov. 3, 2016, 4:02 PM), https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.iupnjj946#.dfoAyyb2V ("Over the past year, the Macedonian town of Veles (population 45,000) has experienced a digital gold rush as locals launched at least 140 US politics websites . . . . They almost all publish aggressively pro-Trump content aimed at conservatives and Trump supporters in the US.").

[59] Hunt Allcott & Matthew Gentzkow, *Social Media and Fake News in the 2016 Election*, 31 J. ECON. PERSP. 211, 211-13 (2017). Allcott and Gentzkow report that "fake news was both widely shared and heavily tilted in favor of Donald Trump." *Id.* at 212. They "estimate that the average US adult read and remembered on the order of

A famous example of these fake stories is the claim that Hillary Clinton and other Democratic Party officials were running a child sex ring out of a Washington, D.C. pizza parlor, Comet Ping Pong. One man was so outraged at this news that he actually took a gun and traveled to the pizza place to sort things out and rescue the children.[60]

Following the election, President Trump confused the issue by asserting that news stories that showed him in a bad light were "fake news."[61] In any case, when other people express concerns about "fake news," they are usually calling for either self-regulation or government regulation of social media.[62]

B. New School Speech Regulation

Both the creation of a right to forget and recent calls for a solution to the problem of fake news are examples of a larger phenomenon: the emergence of a new form of government speech regulation. I call it "new school" speech regulation, to distinguish it from earlier, "old school" speech regulation.[63]

---

one or perhaps several fake news articles during the election period, with higher exposure to pro-Trump articles than pro-Clinton articles." *Id.* at 232. Moreover, "people who report that social media were their most important sources of election news were more likely both to correctly believe true headlines and to incorrectly believe false headlines." *Id.* at 228.

[60] Cecilia Kang & Adam Goldman, *In Washington Pizzeria Attack, Fake News Brought Real Guns*, N.Y. TIMES (Dec. 5, 2016), https://www.nytimes.com/2016/12/05/business/media/comet-ping-pong-pizza-shooting-fake-news-consequences.html.

[61] Danielle Kurtzleben, *With 'Fake News,' Trump Moves from Alternative Facts to Alternative Language*, NPR (Feb. 17, 2017, 8:27 PM), http://www.npr.org/2017/02/17/515630467/with-fake-news-trump-moves-from-alternative-facts-to-alternative-language ("Now, Trump casts all unfavorable news coverage as fake news. In one tweet, he even went so far as to say that 'any negative polls are fake news.'"); Kristen Mitchell, *President Trump Changed Meaning of 'Fake News,'* GW TODAY (Apr. 30, 2017), https://gwtoday.gwu.edu/president-trump-changed-meaning-%E2%80%98fake-news%E2%80%99 ("Mr. Trump has turned 'fake news' into a phrase that undercuts news reports his administration simply does not like."); Robert Schlesinger, *The Maestros of Fake News*, U.S. NEWS & WORLD REP. (July 7, 2017, 11:20 AM), https://www.usnews.com/opinion/thomas-jefferson-street/articles/2017-07-07/donald-trump-doesnt-know-what-fake-news-is ("Trump denounces as fake news pretty much any news story or organization which displeases him.").

[62] Anthony L. Fisher, *Fake News Is Bad. Attempts to Ban It Are Worse*, VOX (July 5, 2017, 10:24 AM), https://www.vox.com/the-big-idea/2017/7/5/15906382/fake-news-free-speech-facebook-google (describing a series of calls for government and self-regulation of fake news).

[63] Jack M. Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2298 (2014) [hereinafter Balkin, *Old-School/New-School*].

Old school speech regulation primarily aims at speakers and publishers of content. It uses traditional methods of enforcement, including civil and criminal fines, injunctions, imprisonment, and in some countries, violence or the threat of violence to deter and censor speakers and publishers.[64]

New school speech regulation, by contrast, is not aimed at speakers or publishers; it is aimed at digital infrastructure.[65] What is the digital infrastructure? It includes the Internet backbone, cloud services, the international domain name system ("DNS"), Internet service providers, web hosting services, social media platforms, and search engines. It also includes payment systems — credit card companies such as Master Card and Visa and new financial intermediaries such as PayPal — who make it possible to fund a whole host of online enterprises.[66] Together this infrastructure makes possible our current system of digital communication.

---

[64] *See id.* at 2340 ("Old-school regulation tries to control bodies, spaces, and predigital technologies of mass distribution.").

[65] For surveys of different techniques of new school speech regulation, see *id.* at 2308-29; Joseph Hall et al., A Survey of Worldwide Censorship Techniques (July 8, 2016) (working paper), https://www.ietf.org/archive/id/draft-hall-censorship-tech-04.txt.

[66] JAMES GRIMMELMANN, INTERNET LAW: CASES & PROBLEMS 33-34 (2016) (describing elements of the Internet "stack"); Balkin, *Old-School/New-School*, *supra* note 63, at 2303-04 (listing elements of the digital infrastructure of free expression); *Free Speech: Only as Strong as the Weakest Link*, ELECTRONIC FRONTIER FOUND., https://www.eff.org/free-speech-weak-link (last visited Nov. 11, 2017) (describing elements of digital infrastructure). Matthew Prince has listed thirteen different types of organizations that can regulate content on the Internet:

> [1] Platforms (e.g., Facebook, Wordpress, etc.), where the content is published.
>
> [2] Hosts (e.g., Amazon Web Services, Dreamhost, etc.), that provide infrastructure on which the platforms live.
>
> [3] Transit Providers (e.g., Level(3), NTT, etc.), that connect the hosts to the rest of the Internet.
>
> [4] Reverse Proxies/CDNs (e.g., Akamai, Cloudflare, etc.), that provide networks to ensure content loads fast and is protected from attack.
>
> [5] Authoritative DNS Providers (e.g., Dyn, Cloudflare, etc.), that resolve the domains of sites.
>
> [6] Registrars (e.g., GoDaddy, Tucows, etc.), that register the domains of sites.
>
> [7] Registries (e.g., Verisign, Afilias, etc.), that run the top level domains like .com, .org, etc.

Nation states have not abandoned old school speech regulation. But they have increasingly moved to new school speech regulation because online speech is hard to govern. Speakers may be judgment proof, anonymous, and located outside the country, and they may not be human at all, but an army of bots. By contrast, owners of infrastructure are usually large for-profit enterprises, they are readily identifiable, and they have assets and do business within nation states.

An additional feature makes new school speech regulation especially attractive to nation states: the large private enterprises that constitute the digital infrastructure have the technical and bureaucratic capacity to regulate and govern speech, through blocking, filtering, and removing content, through otherwise controlling access to their facilities, and through digital surveillance.[67] New school speech regulation depends on these capacities for governance by infrastructure owners, and, to a certain extent, even encourages them.

New school speech regulation has three features that are worthy of note. First, it involves *collateral censorship*, which often has many of the same problems as administrative prior restraints. Second, it involves *public/private cooperation or cooptation*. Third, as noted above, it involves *private governance* by infrastructure owners, and especially

---

> [8] Internet Service Providers (ISPs) (e.g., Comcast, AT&T, etc.), that connect content consumers to the Internet.
>
> [9] Recursive DNS Providers (e.g., OpenDNS, Google, etc.), that resolve content consumers' DNS queries.
>
> [10] Browsers (e.g., Firefox, Chrome, etc.), that parse and organize Internet content into a consumable form. . . .
>
> [11] Search engines (e.g., Google, Bing, etc.), that help you discover content.
>
> [12] ICANN, the organization that sets the rules for the Registrars and Registries.
>
> [13] RIRs (e.g., ARIN, RIPE, APNIC, etc.), which provide the IP addresses used by Internet infrastructure.

Matthew Prince, *Why We Terminated Daily Stormer*, CLOUDFARE (Aug. 16, 2017), https://blog.cloudflare.com/why-we-terminated-daily-stormer.

[67] *See* Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. (forthcoming 2018) (manuscript at 42-49), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2937985 (describing bureaucracies at Facebook, YouTube, and Twitter); Catherine Buni & Soraya Chemaly, *The Secret Rules of the Internet*, THE VERGE (Apr. 13, 2016), https://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech ("During a panel at this year's South by Southwest, Monika Bickert, Facebook's head of global product policy, shared that Facebook users flag more than one million items of content for review every day.").

by search engines (such as Google) and social media platforms (such as Facebook, YouTube, and Twitter) that operate across many countries.

1. Collateral Censorship

The first key feature of new school speech regulation is collateral censorship. Collateral censorship occurs when the state aims at *A* in order to control *B*'s speech.[68] If *A* and *B* are the same enterprise or the same publication, there is not a significant free speech problem. For example, we hold newspapers liable for defamatory speech published by their reporters, and we hold publishers liable for defamatory content by the authors they publish.[69]

On the other hand, if *A* is an infrastructure provider or conduit like an ISP or a social media site, and *B* is an independent speaker, then *A* will tend to over-block and over-censor to avoid liability or government sanction. That is because it is not *A*'s speech that is at stake, but that of a stranger, *B*.

Problems of collateral censorship occur whenever governments adopt intermediary liability rules.[70] In fact, collateral censorship is just the flip side of a rule of intermediary liability. The whole point of holding Internet intermediaries liable for content that they host or that flows through them is to encourage these intermediaries to engage in various forms of collateral censorship, whether it is denial of access, blocking, filtering, or other forms of digital control.

---

[68] J.M. Balkin, *Free Speech and Hostile Environments*, 99 COLUM. L. REV. 2295, 2298 (1999); Balkin, *Old-School/New-School*, *supra* note 63, at 2309-11; *see also* Christina Mulligan, *Technological Intermediaries and Freedom of the Press*, 66 S.M.U. L. REV. 157, 160 (2013) (arguing that collateral censorship threatens freedom of the press). Professor Michael Meyerson coined the term. *See* Michael I. Meyerson, *Authors, Editors, and Uncommon Carriers: Identifying the "Speaker" Within the New Media*, 71 NOTRE DAME L. REV. 79, 118 (1995) (defining collateral censorship as "the silencing by a private party of the communication of others"); *see also* Seth F. Kreimer, *Censorship by Proxy: The First Amendment, Internet Intermediaries, and the Problem of the Weakest Link*, 155 U. PA. L. REV. 11, 16 (2006) (coining the terms "proxy censorship" and "censorship by proxy").

[69] *See* RESTATEMENT (SECOND) OF TORTS § 578 (1977) ("Except as to those who only deliver or transmit defamation published by a third person, one who repeats or otherwise republishes defamatory matter is subject to liability as if he had originally published it.").

[70] Balkin, *Old-School/New-School*, *supra* note 63, at 2309 ("[Exerting] control over privately held intermediaries . . . leads directly to practices of collateral censorship, a characteristic technique of speech regulation in the digital age.").

Collateral censorship in the digital era involves nation states putting pressure on infrastructure providers to censor, silence, block, hinder, delay, or delink the speech of people who use the digital infrastructure to speak. Nation states have a range of different strategies to exert pressure. They can impose fines or criminal penalties. They can threaten prosecution. Or they can engage in jawboning — urging digital infrastructure operators to do the right thing and block, hinder, or take down content.[71] Ex ante methods of speech regulation include filtering and blocking. Ex post methods involve takedown, with or without notice to the speaker.

2. Digital Prior Restraint

New school speech regulation is troubling not only because of collateral censorship. It is also troubling because it has aspects of an administrative prior restraint.

The principle against prior restraint is not limited to judicial injunctions; in fact, administrative prior restraints are a much older form. A key problem of administrative prior restraint is that it involves informal or bureaucratic censorship.[72] Decisions about who gets to speak are made by bureaucrats — or in the digital age, by programmers or algorithms — often without notice or an opportunity to be heard and without the civil liberties and other procedural protections that apply in judicial proceedings. By contrast, in a criminal prosecution, there has to be a judicial determination that speech is unprotected, and the defendant enjoys a full panoply of procedural protections before the defendant is sanctioned and the speech is blocked or censored. New school speech regulation is analogous to prior restraint; when speech is filtered, blocked, or taken down by a filter, an algorithm, or a corporate employee, speakers get none of these procedural benefits.[73]

A second problem of administrative prior restraints is that they flip the burden of action (and the corresponding effects of inertia).[74] In a world without prior restraints, speakers decide for themselves whether to speak and risk prosecution; they do not have to obtain prior

---

[71] Balkin, *Old-School/New-School*, *supra* note 63, at 2327-29; Derek E. Bambauer, *Against Jawboning*, 100 MINN. L. REV. 51, 84-88 (2015).

[72] Balkin, *Old-School/New-School*, *supra* note 63, at 2316-17; Thomas I. Emerson, *The Doctrine of Prior Restraint*, 20 LAW & CONTEMP. PROBS. 648, 656-60 (1955).

[73] *See* Balkin, *Old-School/New-School*, *supra* note 63, at 2315, 2318-24; Emerson, *supra* note 72, at 657-58.

[74] Balkin, *Old-School/New-School*, *supra* note 63, at 2316-17; Emerson, *supra* note 72, at 657.

permission from the government. If they choose to speak, government officials must decide whether to use resources to respond. This arrangement tends to be speech protective — officials may choose to exercise prosecutorial discretion, they may choose to avoid expending resources on a large number of cases, and they may not be aware of every communication within their jurisdiction.[75] These are examples of the burden of action — it is more costly to act than to do nothing and leave matters as they are. The burden of action in this case benefits the speaker. In a system of prior restraints, by contrast, the effects of the burden of action are flipped. The speaker may not speak unless he or she gets prior permission; until the bureaucrat or employee gets around to giving permission, the speech is forbidden. Here again, new school speech regulation — which encourages blocking and filtering — is analogous to prior restraint: until the social media company programmer or algorithm gives permission, the author's speech is blocked, usually without explanation.[76]

For these two reasons, although new school speech regulation may not always be technically an administrative prior restraint, it has many of the same features and functions as prior restraint, and that is what makes it so troublesome from the standpoint of free expression.

Because of the dangers of collateral censorship, some governments, like the United States, provide for varying degrees of intermediary immunity.[77] Intermediary immunity rules relieve collateral censorship by holding the infrastructure owner harmless for content that is stored on their sites, or moves through their channels, when certain conditions are met. These policies are codified in the United States in section 230 of the 1996 Telecommunications Act.[78] A somewhat different set of rules applies to copyright lawsuits through section 512 of the Digital Millennium Copyright Act.[79]

Intermediary immunity provisions work in different ways, they have different effects, and they do not protect all parts of the digital infrastructure — such as the domain name system — but they offer

---

[75] Balkin, *Old-School/New-School*, *supra* note 63, at 2316; Emerson, *supra* note 72, at 657.

[76] *See* Balkin, *Old-School/New-School*, *supra* note 63, at 2318-19, 2326, 2332.

[77] *Id.* at 2313.

[78] *See* 47 U.S.C. § 230(c)(1) (2018) ("No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.").

[79] *See* 17 U.S.C. § 512 (2018); *id.* § 512(a) (providing immunity for "[t]ransitory [d]igital [n]etwork [c]ommunications"); *id.* § 512(b) (providing immunity for temporary caching); *id.* § 512(g) (describing notice-and-takedown procedure for service providers).

protection against some forms of collateral censorship. Other countries have a variety of different rules. Some have very limited intermediary immunity, resulting in much greater degrees of collateral censorship.[80]

C. Public Private Cooperation/Cooptation

A second key feature of new school speech regulation is public/private cooperation and cooptation.[81] Governments aim at infrastructure providers in order to get them to censor or regulate the speech of people that governments cannot easily otherwise control. New school speech regulation seeks to coax the infrastructure provider into helping the state in various ways. These methods range from blocking and filtering content ex ante, to removing content (and access) ex post, to surveilling end-users and providing information about them and their online activities to government officials. Whether willingly or unwillingly, infrastructure providers help the nation state police the digital infrastructure and speech flowing through the infrastructure.

Often it is not even necessary for nation states to threaten infrastructure providers directly. Jawboning sends the message that infrastructure providers should be patriotic and cooperate with the government, rather than getting on the bad side of government officials. Public officials may also appeal to the public to put pressure on infrastructure providers. In general, infrastructure providers prefer a stable, predictable environment in which they are free to do business and make money; therefore, they will often seek to obtain a

---

[80] *See, e.g.*, Case 64569/09, Delfi AS v. Estonia, ¶ 4 (June 16, 2015), http://www.klgates.com/files/Upload/CASE_OF_DELFI%20AS_v._ESTONIA.pdf (holding that Article 10 of the European Convention on Human Rights was not violated by holding a news site liable for anonymous defamatory comments posted by readers, even when the comments are removed on request); Noah C.N. Hampson, Comment, *The Internet is Not a Lawless Prairie: Data Protection and Privacy in Italy*, 34 B.C. INT'L & COMP. L. REV. 477 (2011) (describing Italy's prosecution of three Google executives for a YouTube video that violated the privacy rights of an autistic student who was shown being bullied by classmates); Miquel Peguera, *The DMCA Safe Harbors and Their European Counterparts: A Comparative Analysis of Some Common Problems*, 32 COLUM. J.L. & ARTS 481 (2009) (describing differences between American DMCA safe harbors and the European Union Council Directive on Electronic Commerce).

[81] *See* Balkin, *Old-School/New-School*, *supra* note 63, at 2324-29; Yochai Benkler, *WikiLeaks and the PROTECT-IP Act: A New Public-Private Threat to the Internet Commons*, 140 DAEDALUS 154, 155 (2011); Michael D. Birnhack & Niva Elkin-Koren, *The Invisible Handshake: The Reemergence of the State in the Digital Environment*, 8 VA. J.L. & TECH. 1, 14-17, 57 (2003).

relationship of peaceful coexistence and cooperation with government officials.

For example, when WikiLeaks began disclosing the contents of diplomatic cables in 2010, various members of the American government let it be known that it was outrageous that Amazon web services hosted WikiLeaks; they also let it be known that it was outrageous that Master Card and Visa allowed people to make donations to WikiLeaks.[82] As a result, these services decided to no longer do business with WikiLeaks. WikiLeaks was forced to quickly find substitute infrastructure providers.[83]

The relationship between nation states and infrastructure providers varies along a spectrum. It ranges from direct regulation, to threats, to suggestions that things will go better for infrastructure operators if they cooperate, to negotiations over the terms of cooperation. Sometimes companies will willingly participate with states in order to remain in their good graces. Sometimes, as in the case of digital surveillance and threats to cybersecurity, states and infrastructure operators have common concerns. But equally often, companies are pushed, cajoled, and coerced into cooperation — then they are effectively coopted into assisting states in governance. As we will see shortly, in the case of the right to be forgotten, the European Union has not only ordered Google to comply with European law; it has essentially handed off enforcement of the right in the first instance to Google.[84]

New school speech regulation is not just a feature of states; it capitalizes on the growing power of governance by private owners of infrastructure. Companies like YouTube and Facebook, for example, have created algorithms and policies that decide what is posted or taken down.[85] They have also created private bureaucracies to govern their end-user communities in the interests of the community (and the company's profits).[86] As these technical abilities and bureaucracies

---

[82] *See* Derek E. Bambauer, *Orwell's Armchair*, 79 U. CHI. L. REV. 863, 891-93 (2012) (describing the multipronged campaign to apply pressure to WikiLeaks); Yochai Benkler, *A Free Irresponsible Press: Wikileaks and the Battle over the Soul of the Networked Fourth Estate*, 46 HARV. C.R.-C.L. L. REV. 311, 330-51 (2011) (describing the various efforts the United States government undertook to impede WikiLeaks).

[83] Benkler, *supra* note 81, at 154, 157-58 (describing WikiLeaks' responses).

[84] *See Factsheet on the "Right to Be Forgotten" Ruling (C-131/12)*, EUROPEAN COMM'N, http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_en.pdf (last visited Oct. 16, 2017) (describing Google's obligations under European law to make initial determinations about the right to be forgotten).

[85] *See* Klonick, *supra* note 67.

[86] *See id.* (manuscript at 47-49); Nick Hopkins, *Facebook Moderators: A Quick Guide to*

develop, they are subject to cooptation by states; indeed, these bureaucracies develop in part in response to pressure and complaints by states.

Quite apart from government pressure, however, infrastructure owners have multiple incentives to create their own governance regimes. First, companies want to enforce their terms of service and end-user license agreements to the extent that these rules are important to their profitability. Second, companies want to keep their existing customers happy and attract new customers by preventing bad behavior by strangers and other end-users. Third, companies want to be able to make credible commitments to their current and potential business partners that they can effectively locate, block, filter, tag, or remove content.[87] Fourth, to the extent that social media companies — like Facebook or YouTube — create and maintain communities, they have interests in maintaining and enforcing community norms.[88] They need to arbitrate disputes both with and between end-users. Enforcing community norms is designed to keep the vast majority of end-users happy, deter bad behavior by insiders and outsiders, and help attract new end-users.

The result of these incentives is that companies that began as technology companies soon discover not only that they are actually media companies, but that they are also governance structures.[89] To be

---

Their Job and Its Challenges, GUARDIAN (May 21, 2017), https://www.theguardian.com/news/2017/may/21/facebook-moderators-quick-guide-job-challenges ("Facebook has 4,500 'content moderators' — and recently announced plans to hire another 3,000.").

[87] *See* Lyor Cohen, *Five Observations from My Time at YouTube*, YOUTUBE OFFICIAL BLOG (Aug. 17, 2017), https://youtube.googleblog.com/2017/08/five-observations-from-my-time-at.html ("YouTube's team has built a system in Content ID that helps rightsholders earn money no matter who uploads their music. As of 2016, 99.5 percent of music claims on YouTube are matched automatically by Content ID and are either removed or monetized.").

[88] Klonick, *supra* note 67 (manuscript at 31) (noting that the "economic viability [of online platforms] depends on meeting user's speech and community norms"); *see* Mark Zuckerberg, *Building Global Community*, FACEBOOK (Feb. 16, 2017), https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634 ("[O]ur next focus will be developing the social infrastructure for community — for supporting us, for keeping us safe, for informing us, for civic engagement, and for inclusion of all.").

[89] Klonick, *supra* note 67 (manuscript at 70-72) ("[A]nalysis of online speech is best considered from the perspective of private governance and self-regulation."); Kerry Flynn, *After Charlottesville, Tech Companies Are Forced to Take Action Against Hate Speech*, MASHABLE (Aug. 16, 2017), http://mashable.com/2017/08/16/after-charlottesville-tech-companies-action-nazis/#kxrJzxU9pOqP (statement of Brittan Heller of the Anti-Defamation League) ("Previously tech companies felt like their job was to work behind the scenes . . . since Charlottesville . . . companies [now feel] free to create online communities that reflect the type of communities they want to see in

sure, not all infrastructure owners face the same sets of pressures. Master Card and Visa, for example, face very different pressures than DNS registries and registrars, and both face different pressures than Facebook and Google. But over time, many of the largest infrastructure owners — especially social media companies — find that they have to devote significant resources to private governance.

Perhaps equally important, the largest infrastructure companies, like Google and social media platforms, do business in many different parts of the world and face pressure from many different countries. Their private governance crosscuts national boundaries, and they become intermediate institutions for the regulation of speech around the world. Facing the state, they are businesses subject to regulation — and occasional threats, jawboning, and cooptation. Facing their end-users, they are a new system of governors, special-purpose sovereigns ruling over the members of their communities.

III. PRIVATE GOVERNANCE

This brings us to the third feature of new school speech regulation: private governance. Private governance means that the infrastructure provider governs the flow of information through the infrastructure that it owns, and it governs the behavior of the end-users and customers who employ the digital infrastructure. The ability to govern comes from telecommunications law, from infrastructure owners' property rights, and from their contractual agreements with end-users.

Capacities for private governance are the flip side of new school speech regulation and collateral censorship. New school speech regulation would be ineffective if private infrastructure operators lacked ways to block, filter, surveil, and censor. Thus, new school speech regulation relies on the fact that infrastructure owners have the ability to govern speech either directly (in the case of ISPs' web hosting services and social media platforms) or indirectly (in the case of credit card companies and domain name registrars). New school speech regulation also assumes that if infrastructure owners do not currently have the ability to govern, they will develop such abilities over time — either in the natural evolution of their business models or because territorial governments demand it. In short, new school speech regulation is made possible by the development of privately-owned and privately-employed digital technologies of control and surveillance.

---

real life . . . .").

As Kate Klonick has recounted, over time, companies like YouTube, Facebook, and Twitter have developed elaborate bureaucracies to decide what is consistent with their terms of service, end-user licensing agreements, or other internal company policies, and to adjudicate whether something should be allowed to be posted, or whether it should be taken down.[90]

Often end-users do not want to be exposed to certain kinds of things on the platform — for example, hate speech, pornography, or abusive speech. As a result, end-users, their relatives, government officials, and social activists complain.

Originally, digital infrastructure companies thought of themselves primarily as technology companies.[91] Over time, however, infrastructure owners have faced pressure from two directions. On the one hand, nation states expect them to control and govern their end-users. On the other, the end-users themselves — and other people affected by end-users' speech — expect companies to enforce norms of appropriate behavior.

Social media companies and search engines in particular understood that their end-users required increasing amounts of care and regulation. They realized that a substantial aspect of their product was creating a hospitable environment for end-users, and that meant governing communities of people who used their services.[92]

As a result, many digital infrastructure providers gradually recognized that they were media companies, and that in many cases they were governing communities of end-users, whether they liked it or not. They had to make policies about what could flow through their channels, what could be posted on their facilities, and what would be taken down, and how they would address consumer complaints.[93] Each element of the digital infrastructure has different policies and governance structures — due in part to differences in the nature of their businesses, their history, and their corporate culture. Social media and search engine companies in particular have developed and implemented increasingly elaborate systems of private governance,

---

[90] *See* Klonick, *supra* note 67 (manuscript at 22-26) (describing institutional histories at Facebook, YouTube, and Twitter).

[91] *Id.* (manuscript at 23) (describing Nicole Wong's account of the early history of online platforms).

[92] *See id.* (manuscript at 36-42) (describing the gradual evolution of moderation policies as companies expanded around the world).

[93] *See id.* (describing the evolution of enforcement policies at various social media sites).

enforced by internal bureaucracy, social norms, and code or technology.[94]

Why do we call these systems of control private governance? First, these actors are private corporations, not nation states. Second, these actors promulgate rules, implement them, and enforce them against the people who use their services and platforms. In some cases, such as social media platforms, infrastructure owners create and maintain digital communities of their end-users; they govern the behavior of the people who are part of their community and sanction misbehavior. These sanctions include deciding whether to expel people from the community if they misbehave sufficiently.[95]

Today, if you post or watch videos on YouTube, you are part of the YouTube community. If you have a Facebook account and make posts or read the posts of others, you are part of the Facebook community. These communities are governed by the rules of the infrastructure owner. These companies are the governors of these digital communities, and if you have an account and use the service, you are part of the governed.

A. From Game Gods to Social Media

The idea of private governance is one of the oldest ideas in cyberlaw, and it inspired some of the earliest debates about what made the Internet distinctive and attractive.

At the beginning of the digital age, there were text-based adventure games, called multiuser domains ("MUDs").[96] They had administrators, sometimes, called systems operators or sysops, and sometimes jokingly referred to as wizards or game gods.[97] These

---

[94] *Id.* (manuscript at 42) ("Content moderation at YouTube and Facebook developed from an early system of standards to an intricate system of rules due to (1) the rapid increase in both users and volume of content; (2) the globalization and diversity of the online community; and (3) increased reliance on teams of human moderators with diverse backgrounds.").

[95] *Id.* (manuscript at 56) (describing Facebook and YouTube remedies); *see also How to Appeal*, ONLINECENSORSHIP.ORG, https://onlinecensorship.org/resources/how-to-appeal (last visited Oct. 4, 2017) (describing appeal process at major social media sites).

[96] For an early history of these spaces and how they developed law-like features and governance structures, see Jennifer Mnookin, *Virtual(ly) Law: The Emergence of Law in LambdaMOO*, 2 J. COMPUTER-MEDIATED COMM. 1 (1996), http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1996.tb00185.x/full.

[97] *Id.* ("The oligarchs — MOO-founder Pavel Curtis as well as several other players who had participated in LambdaMOO since its infancy [were] known as 'wizards'; they were responsible for both technical integrity and social control on the

administrators kept their games running. Occasionally, participants would devise various exploits; they would use the software to stretch the rules, do unexpected things — sometimes very obnoxious things — and create all sorts of problems and havoc.[98] Systems operators quickly understood that they had to respond to these exploits and the complaints they engendered. They would have to alter the code or expel people from the game to keep people from misbehaving and acting like trolls with respect to the other players. Thus, the wizards or game gods had to decide what was a permissible move and an impermissible exploit. They were, in effect, digital governors of their game worlds; they set down the rules for operating within the metaphorical space created by the game; they governed the community with code.[99]

In the early days of cyberlaw, David Johnson and David Post imagined that the Internet would foster the creation of multiple digital spaces; each of these spaces would provide different rule sets for how to behave in the space.[100] People would choose which rule sets they wanted to be governed by. Johnson and Post's vision emphasized the freedom of the Internet and the ability to escape the control of territorial governments. People would be able to choose their digital governors, and their ability to choose would enhance their freedom.[101]

MOO."); Beth Simone Noveck, *The State of Play*, 49 N.Y.L. SCH. L. REV. 1, 15-16 (2005) ("Game Gods versus the Law").

[98] Julian Dibbel, *A Rape in Cyberspace*, VILLAGE VOICE (Dec. 23, 1993), http://www.juliandibbell.com/articles/a-rape-in-cyberspace. This article, which remains an excellent introduction to the central problems of governance of cyber communities, describes the exploits of an early troll named Mr. Bungle, who committed a "cyberrape." *Id.*

[99] Mnookin, *supra* note 96; *cf.* Richard A. Bartle, *Virtual Worldliness: What the Imaginary Asks of the Real*, 49 N.Y.L. SCH. L. REV. 19 (2004) (arguing that law should defer to systems operators to govern their own spaces). *See generally* THE STATE OF PLAY: LAW, GAMES, AND VIRTUAL WORLDS (Jack M. Balkin & Beth Simone Noveck eds., 2006) [hereinafter THE STATE OF PLAY] (discussing the governance of virtual worlds and game spaces).

[100] David R. Johnson & David Post, *Law and Borders — The Rise of Law in Cyberspace*, 48 STAN. L. REV. 1367, 1393-97 (1996) [hereinafter Johnson & Post, *Law and Borders*] (arguing for the self-governance of spaces with distinctive rule sets); *see also* David R. Johnson & David G. Post, *And How Shall the Net Be Governed? A Meditation on the Relative Virtues of Decentralized, Emergent Law*, *in* COORDINATING THE INTERNET 62, 65 (Brian Kahin & James H. Keller eds., 1997); Joel R. Reidenberg, *Governing Networks and Rule-Making in Cyberspace*, 45 EMORY L.J. 911, 917-21 (1996) (arguing that network borders would displace territorial borders and that networked communities would create their own norms and rules).

[101] Johnson & Post, *Law and Borders*, *supra* note 100, at 1398-99 ("In Cyberspace . . . any given user has a more accessible exit option, in terms of moving

Critics of Johnson and Post, on the other hand, emphasized that the metaphor of cyberspace as a separate place was misleading. People on the Internet were still subject to national jurisdiction, and nation states had a variety of different techniques to control digital enterprises, even those that styled themselves as separate cyberspaces.[102]

Both sides of this early controversy were partly correct. Digital culture did give rise to online communities that were similar to what Johnson and Post had imagined, but not in all respects. Massive multiplayer online games are one example.[103] More important for our purposes, social media companies also produced digital communities; they established norms that they enforce through a combination of law and code. But Johnson and Post's critics were also right: these companies are constantly being threatened, regulated, and hemmed in by nation states. That is the point of new school speech regulation.

B. From the Dyadic to the Pluralist Model of Speech Governance

The early debates over the governance of game spaces and cyberspace in the early years of the digital age help us understand the problems of private governance by social media sites like Facebook, YouTube, and Twitter. The corporations who operate these platforms are the equivalent of the game gods. They organize a community, they

---

from one virtual environment's rule set to another's, thus providing a more legitimate 'selection mechanism' by which differing rule sets will evolve over time."); *see also* David G. Post, *Governing Cyberspace*, 43 WAYNE L. REV. 155, 166-67 (1996). As David Post explained in a prescient statement:

> [I]ndividual network access providers, rather than territorially-based states, become the essential units of governance; users in effect delegate the task of rule-making to them — confer sovereignty on them — and choose among them according to their own individual views of the constituent elements of an ordered society. The "law of the Internet" thus emerges, not from the decision of some higher authority, but as the aggregate of the choices made by individual system operators about what rules to impose, and by individual users about which online communities to join.

Id.

[102] JACK GOLDSMITH & TIM WU, WHO CONTROLS THE INTERNET? 65-86 (2006) (arguing that nation states had multiple devices for regulating Internet content, including regulating or pressuring infrastructure owners); Lawrence Lessig, *The Zones of Cyberspace*, 48 STAN. L. REV. 1403, 1406 (1996) (arguing that cyberspace "will be regulated by real space regulation to the extent that it affects real space life, and it will quite dramatically affect real space life").

[103] *See generally* THE STATE OF PLAY, *supra* note 99, at 100-02.

impose a set of values on the community, and they enforce those values through a combination of norms, contract law, and code.

In addition to private governance, however, we must add the phenomenon of new school speech regulation. Above or beyond the digital governors, there are territorial governments, who constantly put pressure on digital governors to control their populations in certain ways.

The result is a new system of free expression and speech governance. The scope of speech governance delimits the space for free expression. In the twentieth century model in which modern First Amendment doctrine arose, there was the state on the one hand, and there were speakers and publishers on the other. The state governed both speakers and publishers, producing a law of free speech (and a law of government censorship). The twentieth century model is a *dyadic* model: the state is on one side, speakers and publishers are on the other.

During the early age of the Internet, people imagined that territorial governments would lose much of their power to control speech. John Gilmore famously remarked that "[t]he Net interprets censorship as damage and routes around it."[104] It did not turn out precisely that way, in part because nation states developed the techniques of new school speech regulation. Instead of simply trying to control the speakers, the nation states aimed at the digital infrastructure.

What has emerged is a new model of free expression. This model is pluralist rather than dyadic. For convenience, we can imagine it involving a struggle among at least three different groups of people and organizations. On one side of the triangle we have the state and supra-national entities like the European Union. Although states have not abandoned old school speech regulation, they now rely heavily on new school speech regulation to coerce, coax, and coopt the owners of digital infrastructure.

On the second side of this triangle, we have the companies that operate the digital infrastructure, especially search engines and social media platforms. Many, perhaps most, people now use this

---

[104] Philip Elmer-DeWitt, David S. Jackson & Wendy King, *First Nation in Cyberspace*, TIME (Dec. 6, 1993), http://content.time.com/time/magazine/article/0,9171,979768,00.html (quoting John Gilmore). There are many versions of this famous quote. *See, e.g.*, John Perry Barlow, *Censorship 2000*, ON THE INTERNET, http://www.isoc.org/oti/articles/1000/barlow.html (last visited May 10, 2014) (statement of John Gilmore at the Second Conference on Computers, Privacy, and Freedom that "[t]he Internet treats censorship as though it were a malfunction and routes around it").

infrastructure to communicate with each other. These companies — Facebook, YouTube, Twitter, Google, and their competitors and successors — are the new governors of digital expression. They develop and enforce their terms of service, end-user license agreements, and internal company policies, applying them to end-users and community members. They operate in many different countries but are privately owned. Some of them are richer than some of the countries in which they do business.[105]

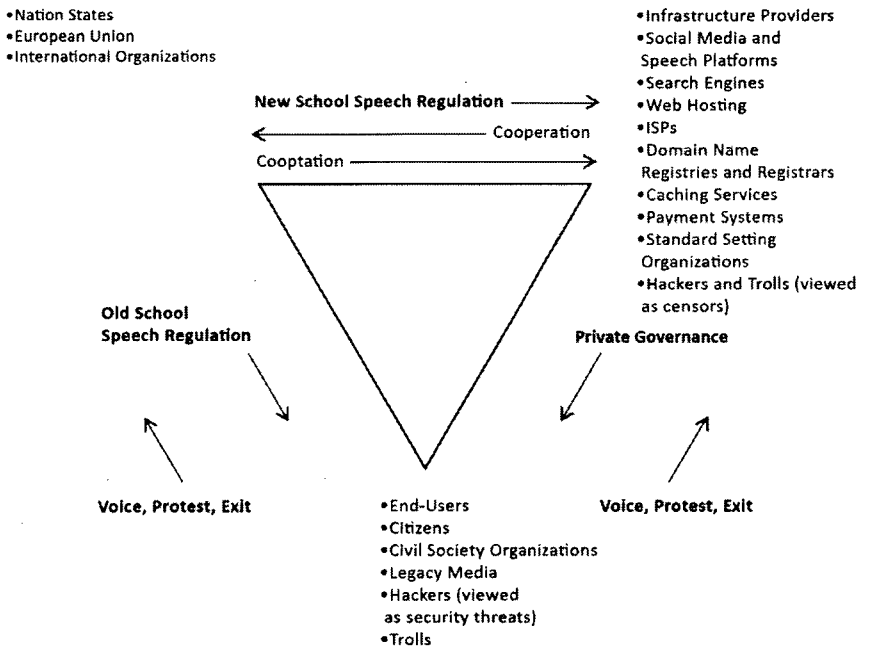
On the third side of the triangle we have the speakers who use the digital infrastructure to communicate. They include people who are well behaved and who are not well behaved: parents and children; consumers and activists; trolls and norm enforcers; people who use their own names and people who want to remain anonymous; people who hide behind bots and people who take on multiple identities. The two other sides of the triangle subject them to various forms of surveillance and control. But speakers have a variety of techniques to push back, protest, evade, resist, troll, and exploit the system of governance. Moreover, they can use the digital public sphere to place social pressure on these digital platforms to modify their policies.

The substantive problem of free expression in the early twenty-first century is that the practical ability to speak online is affected by power struggles among these various groups. Our practical ability to speak is shaped by (1) our relationship to the state; (2) our relationship to the owners of the digital infrastructure that we use to speak; and (3) the relationships of cooperation, cooptation, and coercion between states and digital infrastructure owners.

---

[105] *See, e.g.*, Anupam Chander, *Facebookistan*, 90 N.C. L. REV. 1807 (2012) (arguing that Facebook has become so large, rich, and powerful that in many ways it behaves like a nation state).

Diagram 1. The Pluralist Model of Speech Regulation



As a first approximation, we might call this system triadic, because of the three sides of the triangle. In fact, the new system of speech regulation is actually far more complicated, because there are more than three sets of players.

First, international organizations and stakeholders like the International Corporation for Assigned Names and Numbers ("ICANN") and international standard setting organizations also create rules and impose governing regimes.[106]

Second, end-users can be governed by more than one infrastructure provider. Even if one digital infrastructure provider grants access, another may block or hinder. Imagine an unpopular speaker who operates a website or uses social media. The digital infrastructure has many different potential points of control, staffed by different private companies. A search engine company can delink the site or demote its page ranking. An ISP can block or filter the site. A webhosting service

---

[106] *See generally* LAURA DENARDIS, PROTOCOL POLITICS: THE GLOBALIZATION OF INTERNET GOVERNANCE (2009) (explaining the role of standard setting in controlling Internet communication and traffic).

can refuse service. A social media site can filter, block, or revoke posting privileges altogether. A DNS registrar can refuse to host a site's domain name. Payment systems companies can refuse to allow payments to the site or its operators.[107]

Third, legacy media organizations like newspapers, broadcasters, cable networks, and movie studios may strike deals with digital infrastructure providers to regulate content; they may also pressure them directly — or indirectly, by lobbying nation states — to regulate digital content. Fights over intellectual property management in the early twenty-first century often pitted mass media organizations against digital infrastructure owners.[108] More recently, legacy media organizations, as organs of public opinion, may pressure digital infrastructure owners to regulate hate speech or other content.

Fourth, civil society organizations as well as end-users may put pressure on digital infrastructure providers. They may object to speech that they find obnoxious, bigoted, racist, or abusive.[109] They may demand that social media platforms block or take down content and expel trolls or misbehaving end-users. There will also be pushback in the opposite direction by end-users and civil society organizations that want to relax content controls and expand end-user rights.

Fifth, cyberattacks by hackers and other rogue elements on the Internet infrastructure — and on particular websites and speakers — place additional pressure on digital infrastructure companies. Hackers may seek to censor or dissuade speakers through distributed denial of service or other kinds of attacks. Some infrastructure companies, like Cloudflare, exist primarily to protect websites and speakers from attacks by hackers.[110]

---

[107] *See* Balkin, *Old-School/New-School*, *supra* note 63, at 2297, 2303-04, 2308-10, 2322-24, 2326, 2328 (describing techniques of speech control involving digital infrastructure); Hall et al., *supra* note 65 (providing an overview of the mechanisms used internationally to block Internet content); Prince, *supra* note 66 (describing targets of speech control in digital infrastructure).

[108] *See generally* WILLIAM PATRY, MORAL PANICS AND THE COPYRIGHT WARS (2009) (offering a history of legal struggles between the content industries and Internet companies over digital copyright).

[109] *See, e.g.*, Klonick, *supra* note 67 (manuscript at 64-67) (discussing the role of "third-party influencers" in shaping content moderation policies); Russell Brandom, *Charlottesville Is Reshaping the Fight Against Online Hate*, VERGE (Aug. 15, 2017, 1:54 PM), https://www.theverge.com/2017/8/15/16151740/charlottesville-daily-stormer-ban-neo-nazi-facebook-censorship ("This morning, the neo-Nazi website *Daily Stormer* dropped off the internet, the result of sustained campaigning by the Southern Poverty Law Center and other groups.").

[110] *See* Prince, *supra* note 66. Matthew Prince of Cloudflare has described the role of hackers in censorship:

Sixth, quite apart from attacks on infrastructure, groups can use a variety of techniques to discipline, harass, confuse, discourage, and silence other speakers, and to confuse, distract, and mislead Internet audiences.[111] These techniques discourage and control speech informally rather than through direct state prohibitions or mandates. For example, speakers may engage in *online harassment and threats*, often through the use of coordinated armies of trolls consisting of combinations of humans and bots posing as humans.[112] These troll armies may be loosely affiliated with nation states and political movements or encouraged by them.[113] Speakers may also *flood the Internet with propaganda* or other material designed to discredit reliable sources of information, or to distract or confuse people who get their news from social media.[114] The goals of flooding and propaganda are twofold. The first goal is to make it difficult or impossible for people to figure out what is true and false. The second goal is to deflect and monopolize people's limited attention spans so that they pay little attention to or lose interest in what other speakers are saying.[115] Again, nation states and social movements may secretly support or encourage this propaganda.[116]

Because the emerging system of speech regulation has so many possible players, who can regulate speech in so many different ways, we call it *pluralist*.

> The size and scale of the attacks that can now easily be launched online make it such that if you don't have a network like Cloudflare in front of your content, and you upset anyone, you will be knocked offline. In fact, in the case of the Daily Stormer, the initial requests we received to terminate their service came from hackers who literally said: "Get out of the way so we can DDoS this site off the Internet." *Id.*

111 TIM WU, EMERGING THREATS: IS THE FIRST AMENDMENT OBSOLETE? 9-11 (Sept. 2017), https://knightcolumbia.org/sites/default/files/content/Emerging%20Threats%20Tim%20Wu%20Is%20the%20First%20Amendment%20Obsolete.pdf.

112 *Id.* at 11-17 (discussing the use of troll armies to harass and silence critics).

113 *Id.* at 11-13 (describing the implementation of Russian tactics and their adoption by the alt-right to attack journalists who criticize Donald Trump).

114 *Id.* at 15-17 (describing various techniques that employ reverse censorship, flooding, and propaganda robots).

115 ZEYNEP TUFEKCI, TWITTER AND TEAR GAS: THE POWER AND FRAGILITY OF NETWORKED PROTEST 241 (2017) (describing the goals of overwhelming people with confusing and distracting propaganda); WU, *supra* note 111, at 15 (explaining how these techniques leverage limited human attention spans).

116 WU, *supra* note 111, at 11-13, 15 (describing Russian interference in the 2016 American Presidential election and the use of bots).

One might object that a pluralist — or at least, a triadic — system already existed in the twentieth century, because mass media companies already played the same role as digital infrastructure providers. After all, during the twentieth century, mass media companies — newspapers, publishing houses, books stores, movie production studios, television, and radio broadcasters — acted as gatekeepers for the speech of others.

However, the role of digital infrastructure companies in the twenty-first century is not the same as the role of mass media companies in the twentieth century. These mass media companies were not conduits for the speech for the vast majority of the people who constituted the audience for their products. Rather, these companies (1) produced their own content, (2) published the content of a small number of creative artists, or (3) delivered content made by other organizations to a mass public. In the twentieth century model, the vast majority of people were members of an audience for mass media products, but very few actively used mass media as speakers or broadcasters. Twenty-first century governors of digital speech, by contrast, make their money by facilitating and encouraging the production of content by ordinary people and governing the communities of speakers that result.

New media companies like Facebook, Google, YouTube, and Twitter do not produce most of the content they serve. Rather, their business model requires them to induce as many people as possible around the world to post, speak, and broadcast to each other. Constant production of content by end-users, in turn, captures audience attention. This allows digital media companies to sell advertising, collect data about end-users, and use this information to sell even more advertising.

The twentieth century model required large audiences. The twenty-first century model requires both large audiences and large numbers of speakers who actively participate in the production of the content. If people stopped posting on Facebook, the company would wither and die on the vine. Ideally, Facebook would like each and every person on Earth to talk incessantly and check their Facebook feeds constantly. Unlike a twentieth century movie studio or television station, it is not enough for Facebook to have a large audience. Facebook needs people perpetually to speak to each other and post new content. That is why, whether they like it or not, companies like Facebook and Google find themselves in a position of governance.

Twentieth century mass media companies that spoke to mass audiences had to be the editors and publishers of content they produced. But digital platforms that make money by encouraging

people to speak to others inevitably become the governors of the communities they produce. It is this difference that creates the triadic — or, more correctly, pluralist — structure of contemporary freedom of speech.

The best analogy to the pluralist system of speech regulation is not twentieth century mass media companies. It is to systems of speech regulation by professional organizations and universities.[117] Professional organizations and universities govern the speech of members. At the same time, territorial governments put regulatory pressure on both professionals and professional organizations; they attempt to regulate universities, students, and teachers. The result is a complicated set of power relationships. It is therefore not surprising that the rules governing academic freedom and professional regulation do not correspond to standard First Amendment doctrine.[118]

C. Problems of the New System of Public/Private Governance

New school speech regulation and private governance are interconnected. Digital communication leads states to new school speech regulation. New school speech regulation creates incentives for private governance. Platform operators become special-purpose sovereigns who govern populations of end-users.

The emerging system of private governance raises three important issues for freedom of expression. The first issue, which I have already touched on, is that free speech protection involves a power struggle involving multiple players. The second is that the practical ability to speak does not correspond to standard First Amendment norms or doctrine. The third is that in this new system, due process becomes an increasingly important value.

---

[117] *See generally* PAUL HORWITZ, FIRST AMENDMENT INSTITUTIONS (2013) (emphasizing the central infrastructural role played by First Amendment institutions in public discourse); POST, *supra* note 31 (arguing for a distinctive approach to speech regulation outside of public discourse and within professional and knowledge producing institutions).

[118] Claudia E. Haupt, *Professional Speech*, 125 YALE L.J. 1238, 1241-42 (2016) (arguing that professional speech is treated specially because of the role of professional communities in producing knowledge). *See generally* POST, *supra* note 31 (arguing that First Amendment doctrine for academic freedom and professional regulation is distinctive because it involves speech that lies outside of general public discourse).

1. Private Governance and Private Norms

Most forums in which most people speak and interact these days are governed by the norms of digital infrastructure companies. Those norms do not conform to the requirements of the First Amendment.[119]

One might object that this fact is irrelevant to the question of freedom of expression, because, by hypothesis, private governors are not the state. But this neglects the importance of the emerging pluralist system of speech regulation.

First, most speech travels through digital infrastructure and is subject to private governance. Second, because of new school speech regulation, private governance is not wholly private; it results from constant pressure by states and constant interactions and cooperation between private companies and states.

Third, although, online speech platforms are not public forums in a First Amendment sense, they are public in a different sense — they are sites for public discourse directed not only to specific individuals but to an undifferentiated public.[120] And although platform owners are private businesses, they are also governors — they govern the communities of people who use the applications, create and apply norms, and settle disputes among their end-users.

For these reasons, it is unhelpful to impose a rigid distinction between public and private power to understand digital speech today. This is not a claim that infrastructure owners should be treated as state actors — unless, as in some countries, they happen to be owned by states themselves. Arguing that infrastructure owners are state actors simply replicates the outmoded assumptions of a pre-digital world. Rather, the present world features at least two sources of governing authority: new school speech regulation by states and speech governance by different kinds of Internet infrastructure owners. We cannot understand the situation by collapsing these two groups into one in order to preserve a specious duality between public and private power. Private governance is now essential to digital speech regulation and hence to digital freedom of expression.

Online communities enforce speech norms that protect far less expression than the corresponding obligations of government under

---

[119] *E.g.*, Klonick, *supra* note 67 (manuscript at 55) (noting that Facebook flags content that involves "promotion or encouragement of bestiality, . . . bullying, self-harm content, poaching of endangered animals, Holocaust denial, all attacks on Ataturk, maps of Kurdistan and Burning Turkish Flags").

[120] Balkin, *Digital Speech and Democratic Culture*, *supra* note 2, at 23 ("[D]igital communications networks are 'public' in the sense that the public uses them as a space for general interaction.").

the American First Amendment. Online communities police abusive speech, sexual expression, and hate speech that the American First Amendment normally shields from government regulation.[121] In addition, although the First Amendment generally protects the right to engage in anonymous speech in the public sphere, some online service providers, like Facebook, may require people to use their real names (or the names by which they are known to their community) to discourage trolling and abusive behavior.[122]

Platform owners impose these rules out of mixed motives. But one reason they do so is to make people feel safe and respected within online communities. If end-users feel safe, they will continue to participate, post content, and make the platform part of their daily lives. This allows platform owners to be profitable. But it also helps foster a constant, vibrant flow of ideas and opinions.[123]

Because platform owners are private actors, constitutional law permits them to engage in content-based regulation that would be prohibited under the First Amendment if they were treated as state actors.[124] Moreover, platform owners do business in many different countries, and as a result, their moderation policies may differ significantly from American free speech norms.[125]

---

[121] Klonick, *supra* note 67 (manuscript at 52-54, 57-59) (describing similarities and differences between Facebook's "Abuse Standards" and First Amendment law).

[122] *What Names Are Allowed on Facebook?*, FACEBOOK, https://www.facebook.com/help/112146705538576?helpref=faq_content (last visited Oct. 5, 2017) ("Facebook is a community where everyone uses the name they go by in everyday life. This makes it so that you always know who you're connecting with and helps keep our community safe.").

[123] Danielle Keats Citron & Helen Norton, *Intermediaries and Hate Speech: Fostering Digital Citizenship for our Information Age*, 91 B.U. L. REV. 1435, 1454-55 (2011) (arguing that intermediaries regulate speech as a matter of corporate responsibility and to protect profits); Klonick, *supra* note 67 (manuscript at 19-20).

[124] *See, e.g.*, Lloyd Corp. v. Tanner, 407 U.S. 551, 564-65 (1972) (holding that the First Amendment does not prevent a private shopping center owner from prohibiting the distribution on center premises of handbills unrelated to the center's operations). On the other hand, online platforms could be required to conform to First Amendment doctrine if courts analogized them to company towns that provided identical functions to a state-run municipality. *See* Marsh v. Alabama, 326 U.S. 501, 507-10 (1946) (holding that when a company town assumed the functions of a state municipality, it could not exclude the distribution of religious literature within the city).

[125] *See, e.g.*, Alex Hern, *Facebook, YouTube, Twitter and Microsoft Sign EU Hate Speech Code*, GUARDIAN (May 31, 2016, 8:16 AM), https://www.theguardian.com/technology/2016/may/31/facebook-youtube-twitter-microsoft-eu-hate-speech-code (describing an agreement by technology companies "to review the 'majority of valid notifications for removal of illegal hate speech' in less than 24 hours, and to make it easier for law enforcement to notify the firms directly").

In her study of online speech platforms, Kate Klonick has pointed out that currently the speech policies of the major online platforms — Facebook, YouTube, and Twitter — tend to be fairly free-speech protective, in large part because the people who created these companies were either American or were heavily influenced by American free speech law.[126] But there is no guarantee that social media platforms — or infrastructure owners generally — will continue to be so free-speech friendly in the future.

First, there is no guarantee that the dominant players will continue to be American. Second, platform policies are the result of a tug of war between the demands of company owners and shareholders, end-users and nation states. The direction of this tug of war is unpredictable.

For example, in response to white supremacist and neo-Nazi demonstrations at Charlottesville, Virginia, in August 2017, a number of digital infrastructure companies — domain name registrars, web hosting services, Internet proxy services, payment systems, and even a music streaming service, Spotify — refused service to white supremacist and neo-Nazi organizations.[127] It is likely that we will see more of the same governance by private infrastructure owners in the future — including American-owned businesses who assert a firm belief in the values of the First Amendment.[128]

2. Private Governance and Due Process

In addition to substantive problems of private speech regulation, there are also procedural problems. Currently, speech platforms do

---

[126] Klonick, *supra* note 67 (manuscript at 30) ("The early history and personnel of these companies demonstrates how American free speech norms and concerns over censorship became instilled in the speech policies of these companies.").

[127] Elizabeth Flock, *Spotify Has Removed White Power Music from Its Platform. But It's Still Available on Dozens of Other Sites*, PBS (Aug. 18, 2017, 6:18 PM), http://www.pbs.org/newshour/art/spotify-removed-white-power-music-platform-still-available-dozens-sites; Flynn, *supra* note 89 ("Facebook, Google, Spotify, Uber, Squarespace, and a variety of other tech companies are taking action to curb the use of their platforms and services by far-right organizations.").

[128] *See* Violet Blue, *Options for Neo-Nazis on the Internet Are Starting to Shrink*, ENGADGET (Aug. 18, 2017), https://www.engadget.com/2017/08/18/options-for-neo-nazis-on-the-internet-are-shrinking (listing a wide range of digital infrastructure and platform companies who have denied service to far right groups); Franz Paasche, *PayPal's AUP — Remaining Vigilant on Hate, Violence & Intolerance*, PAYPAL (Aug. 15, 2017), https://www.paypal.com/stories/us/paypals-aup-remaining-vigilant-on-hate-violence-intolerance (explaining PayPal's Acceptable Use Policy and defending "the limiting and closing of sites that accept payments or raise funds to promote hate, violence and intolerance").

not govern in the same way that liberal democratic states do. Enforcement of community norms often lacks notice, due process, and transparency.[129] Platform operators may behave like absolutist monarchs, who claim to exercise power benevolently, but who make arbitrary exceptions and judgments in governing online speech.

For many participants, procedural values may be as important if not more important than substantive values. Procedural values become important precisely because online companies act like governors. They police — and are expected to police — the spaces they operate, they promulgate rules and policies, and they enforce them in the name of the online community. The more that end-users view businesses as governors, or as special-purpose sovereigns, the more end-users will expect — and demand — that these companies should conform to the basic obligations of governors towards those they govern. These obligations include procedural fairness in handling complaints and applying sanctions, notice, transparency, reasoned explanations, consistency, and conformity to rule of law values — the "law" in this case being the publicly stated norms and policies of the company.

Indeed, end-users may accept, to some degree, that companies will take down materials that violate a company's internal policies and expel end-users who are abusive or violate the company's terms of service. But what end-users may especially resent is that the criteria are kept hidden, that the takedowns are done summarily, that the rules are applied arbitrarily, that powerful people and organizations are given exceptions to the rules, and that end-users are booted off the platform without notice and an opportunity to defend themselves.

The shift from a binary model of state versus speaker to a pluralist model of speakers, private governors, and states has the curious effect of making procedural norms especially salient and therefore especially valuable to end-users.[130] End-users cannot expect that private

---

[129] Buni & Chemaly, *supra* note 67 ("The details of moderation practices are routinely hidden from public view, siloed within companies and treated as trade secrets when it comes to users and the public."); Klonick, *supra* note 67 (manuscript at 73); Hopkins, *supra* note 86 (describing leak of secret Facebook manuals for content moderators).

[130] *See* Jeremy Malcolm, Cindy Cohn & Danny O'Brien, *Fighting Neo-Nazis and the Future of Free Expression*, ELECTRONIC FRONTIER FOUND. (Aug. 17, 2017), https://www.eff.org/deeplinks/2017/08/fighting-neo-nazis-future-free-expression (arguing for content neutrality at the level of certain levels of the digital infrastructure, and for due process protections at other levels); Prince, *supra* note 66 ("I, personally, believe in strong Freedom of Speech protections, but . . . it is a very American idea that is not shared globally. On the other hand, the concept of Due Process is close to universal."); *Manila Principles on Intermediary Liability*, MANILA PRINCIPLES, https://www.manilaprinciples.org

governors will conform in every respect to the speech norms that should apply to states. They may accept — and some may even insist upon — private platforms' efforts to police content and abusive speech. In such a world, procedural norms become far more valuable and important.

Put another way, as online speech platforms govern, and increasingly resemble governments, it is hardly surprising that end-users expect them to abide by the basic obligations of those who govern populations in democratic societies. These expectations include (1) obligations of transparency, notice, and fair procedures; (2) the offer of reasoned explanations for decisions or changes of policy; (3) the ability of end-users to complain about the conduct of the institution and demand reforms; and (4) the ability of end-users to participate, even in the most limited ways, in the governance of the institution.

Online speech platforms may often frustrate these expectations of procedural and participatory fairness. But online speech platforms will find that they cannot ignore them entirely. These expectations point to the ways that online speech platforms will have to accommodate their end-users in the future. Moreover, they point to the kinds of reforms that democratic governments may eventually try to require of online speech platforms as they become indispensable features both of commerce and democracy within democratic nation states.

An anecdote may help underscore this point. I was recently on a panel on online speech with a very well-known journalist who regards herself as a great defender of freedom of speech and press. She was quite upset that YouTube had taken down certain videos that she and her colleagues had posted. She strongly resisted the idea that governments should intervene on her behalf to regulate YouTube's speech. As a champion of First Amendment values, she retained the binary model of state power in opposition to the rights of individual speakers. Nevertheless, she wanted something from YouTube — she wanted prior notice and an explanation of the takedown, as well as an opportunity to defend her actions. What she wanted from private governors, in short, was due process.[131]

(last visited Oct. 5, 2017) (arguing for transparency and accountability in digital intermediaries).

[131] Future.Today Summit, *The Future of the First Amendment*, LIVESTREAM (Dec. 6, 2016), https://livestream.com/92Y/FutureToday/videos/143684069 (remarks of Judith Miller).

3. Exit, Voice, and Loyalty in Private Governance

We can restate some of the features of private speech governance in terms of Albert O. Hirschman's theory of how people respond to organizations that they perceive as failing them.[132] Participants can exit from the platform, or they can attempt to voice their objections by complaining to the platform operators.[133] They exercise loyalty by sticking with the organization and hoping that it improves. Hirschman noted that participants' choice between exit and voice is shaped by the nature and design of the institution.[134] He also pointed out that easy exit would, all other things being equal, undermine resort to voice.[135] But the choice to stay and complain rather than exit is affected by loyalty to the institution. Loyalty to the institution, in turn is shaped by what participants believe about how likely it is that that things will improve, how much they feel respected by the organization, how difficult or costly it is to change communities, and how much they feel invested in the community and their ability to participate in it.

Exit from online platforms is normally not the dominant strategy. First, exit from a platform may be costly because of network effects. Second, Klonick points out that many end-users already belong to more than one platform and may view platforms as complementary, rather than as substitute goods.[136] Thus, leaving Facebook is not like leaving Russia for the United States. One can inhabit multiple platforms simultaneously. Exit may not be the favored option because participation in one platform may make participation in others more valuable.

Social media companies, in turn, encourage the idea that end-users are part of a larger community and encourage participants to continue to check in — and post — as much as possible. Companies work hard to design their sites so as to attract — and capture — their end-users' time and attention.[137] In addition, companies may have compiled

---

[132] ALBERT O. HIRSCHMAN, EXIT, VOICE, AND LOYALTY: RESPONSES TO DECLINE IN FIRMS, ORGANIZATIONS, AND STATES 1-3 (1972).

[133] Charles Tiebout had emphasized the exit option in the context of local governments. *See* Charles M. Tiebout, *A Pure Theory of Local Expenditures*, 64 J. POL. ECON. 416, 423-24 (1956). Hirschman added the option of voice and argued that it was as important if not more important than exit. *See* Albert O. Hirschman, *"Exit, Voice, and Loyalty": Further Reflections and a Survey of Recent Contributions*, 13 SOC. SCI. INFO. 7, 8-9 (1974).

[134] HIRSCHMAN, *supra* note 132, at 86.

[135] *Id.* at 76.

[136] Klonick, *supra* note 67 (manuscript at 35).

[137] TIM WU, THE ATTENTION MERCHANTS: THE EPIC SCRAMBLE TO GET INSIDE OUR

significant databases of end-users' preferences, interests, and behavior. Although one can exit in the sense of no longer visiting the site, this by itself does not erase data about end-users or prevent later use of the data by the company — or by others to whom the data is later sold. Unless the company has promised to erase all data collected — or has been required to erase it by law — personal data remains with the company for its later use. In this specific sense, exit is not possible.

Hence, when end-users feel badly treated on online platforms, the dominant strategies are likely to be loyalty — bearing with the situation and hoping for improvement — or voice — complaining about bad treatment and demanding accountability and/or reform. Note that when, as often happens, people are kicked off a platform, they are not choosing exit — they would like to stay and complain.

The most significant exceptions — when exit becomes the dominant strategy — are cases in which the platform becomes unreliable or no longer offers significant network effects. Then participants will simply stop using the platform rather than officially quit it.[138] And, as previously noted, doing so will not remove personal data previously compiled by the company.

In addition to calls for reform, perhaps the most important form of voice in online platforms occurs when end-users tag or identify content that they consider inappropriate. Not only is this kind of voice ubiquitous, it is constitutive of the online community and its norms. The moderation systems of online companies actually depend on this kind of voice.[139] Moderation systems, like many open-source projects, employ end-users as a kind of unpaid workforce that polices the platform, drawing attention to potential violations of community norms, and, in the process, shaping their evolution.[140]

Because voice tends to dominate exit in online speech platforms, system operators face continual pressure from end-users (as well as

---

HEADS 255-66 (2016).

138 Something like the latter happened to MySpace, although it is still not exactly clear what caused the decline. *See* Sean P. Aune, *Why Did Everyone Leave MySpace for Facebook?*, TECHNOBUFFALO (July 18, 2010), https://www.technobuffalo.com/2010/07/18/why-did-everyone-leave-myspace-for-facebook. It is possible that generational change will lead people, while retaining accounts on old social media platforms, to move most of their attention to new ones. *Id.*

139 Buni & Chemaly, *supra* note 67 ("[U]sers play a critical role in moderation, since almost every content moderation system depends on users flagging content and filing complaints, shaping the norms that support a platform's brand.")

140 *Id.* ("[U]sers are not so much customers as uncompensated digital laborers who play dynamic and indispensable functions (despite being largely uninformed about the ways in which their labor is being used and capitalized).").

from non-members who regard the platform as socially important). The exercise of voice leads naturally to a demand for due process.

We should not confuse the influence of end-users on social media companies with genuine democracy. Private governors are much like nineteenth century enlightened despots. They champion a set of enlightened values that they believe that their end-users want — or should want — but they implement these values through bureaucracy and code without taking any sort of vote. And, like enlightened despots, they reserve the right to act arbitrarily on occasion.[141]

IV. SPEECH GOVERNANCE, THE RIGHT TO BE FORGOTTEN, AND THE PROBLEM OF FAKE NEWS

We can apply this analysis to the right to be forgotten and the problem of fake news.

A. The Right to Be Forgotten

The right to be forgotten, a feature of European data privacy law, was extended to online search engines in 2014 in the *Google Spain* case by the Court of Justice of the European Union ("CJEU").[142] The case arose out of a 2010 complaint to the Spanish Data Protection Agency by a Spanish lawyer, Mario Costeja González. Costeja González complained that people searching for his name on the

---

[141] For example, Facebook's Mark Zuckerberg, acting in his capacity as liberal autocrat, decided not to take down candidate Donald Trump's posts even though they technically violated the company's hate speech policy. Deepa Seetharaman, *Facebook Employees Pushed to Remove Trump's Posts as Hate Speech*, WALL ST. J. (Oct. 21, 2016, 7:43 PM), http://www.wsj.com/articles/facebook-employees-pushed-to-remove-trump-posts-as-hate-speech-1477075392.

Kate Klonick offers the example of Facebook's removal of a famous picture of a Vietnamese girl running naked following a Napalm attack. Klonick, *supra* note 67 (manuscript at 63-64). The photo likely violated Facebook's terms of service banning nudity, but it was restored and an exception to the policy made after complaints by a Norwegian newspaper that had also tried to publish the picture. *Id.* Klonick notes that Facebook's response likely reflected the power of the people complaining, since the photo had likely been posted and removed "thousands of times" before. *Id.* (manuscript at 64). "To the best of our knowledge, however," Klonick explains, "all prior instances had failed to happen to a famous author, political world leader, or the editor-in-chief of a newspaper — and thus, the content had never been reinstated." *Id.*

[142] *See* Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos, ¶ 99 (May 13, 2014), http://curia.europa.eu/juris/document/document_print.jsf?doclang=EN&docid=152065. The right to be forgotten in European law is much older. *See* Post, *supra* note 56 (manuscript at 1-3) (discussing the French Privacy Law of 1978 and the 1995 European Data Privacy Directive).

Internet would discover two brief newspaper accounts in January and February 1998 available on the site of the La Vanguardia newspaper.[143] These stories were public announcements "mentioning Mr Costeja González's name" in connection "with attachment proceedings for the recovery of social security debts."[144] Costeja González argued that the ability of the public to access these stories violated his rights under the European Data Privacy Directive,[145] and he asked for the newspaper to delete his name and Google to remove links to the newspaper accounts.

The remedy that the CJEU eventually ordered, however, was not directed at newspapers but at search engines, which are part of the digital infrastructure. It held that a data subject — such as Mr. Costeja González — could complain if information is "inadequate, irrelevant or no longer relevant, or excessive in relation to [the] purposes [of the processing] and in the light of the time that has elapsed."[146] In such cases, the operator of a search engine must remove links to the relevant webpages.[147]

To require newspapers to take down stories would appear to be a serious intrusion into the freedom of the press.[148] Instead, at the outset, the CJEU has targeted search engines, on the ground that under the terms of the Directive, search engine companies are data controllers that process personal information.[149]

Usually nation states target infrastructure in new school speech regulation because it is difficult to locate and sanction the actual speaker. That is not the case here. The articles complained of appear in the archives of European newspapers; the newspapers are not anonymous, do business within the jurisdiction of European courts, and are easily located. Rather, the right to be forgotten aims at search

---

[143] *Google Spain SL*, Case C-131/12, ¶ 14.

[144] *Id.*

[145] *See* Council Directive 95/46, 1995 O.J. (L 281) 31 (EC).

[146] *Google Spain SL*, Case C-131/12, ¶ 93.

[147] *Id.* ¶ 94.

[148] Unlike the European Data Protection Directive, Italy does not recognize a journalism exception to its national data privacy law. In 2006 an Italian court applied the right to be forgotten to newspapers and ordered them to take down old stories. *See* Post, *supra* note 56 (manuscript at 27 n.97) (citing Athalie Matthews, *How Italian Courts Used the Right to Be Forgotten to Be an Expiry Date on the News*, GUARDIAN (Sept. 20, 2016, 4:12 AM), https://www.theguardian.com/media/2016/sep/20/how-italian-courts-used-the-right-to-be-forgotten-to-put-an-expiry-date-on-news?CMP=share_btn_tw).

[149] *Google Spain SL*, Case C-131/12, ¶ 28.

engines because the goal is to produce *practical obscurity* rather than complete censorship.[150]

In a pre-digital era, old newspaper articles that contained embarrassing information were quite literally yesterday's news. People threw away the old copies and one had to go to library or some other location where the archives were stored. In the digital age, organizations publish but do not delete. Instead, old articles are freely searchable in newspaper archives, which remain online. This fact changes the nature of a newspaper as an institution of the public sphere. The newspaper is no longer simply a report of the day's events, to be cast aside tomorrow and stored, if at all, in a relatively small number of libraries and other archival locations that are not quickly and easily accessible to the public. Instead, the newspaper becomes an increasingly important and valuable online archive. It becomes an institution of memory that is widely and easily accessible through search engines. Newspapers become important records of history experienced in real time that remain present for people to search and read days, months, and years later.

When embarrassing material was published in newspapers in the past, the subjects eventually enjoyed practical obscurity when the newspapers were discarded (so that access to older newspaper stories was limited to those who visited libraries and archives). This practical obscurity has vanished because of search engines. Embarrassing articles may show up in search engine results and continue to appear indefinitely. By targeting search engine providers, the right to be forgotten attempts to restore the practical obscurity (and thus privacy protection) of the pre-digital era.

The right to be forgotten raises three issues.

1. Collateral Censorship

First, the right to be forgotten is a classic example of collateral censorship. Instead of going after the speaker, the state targets the infrastructure provider, and it threatens to hold the search engine company liable if it does not delink embarrassing articles from newspapers. The government puts pressure on the infrastructure owner to muffle (but not completely silence) the voice of the original speaker. The speaker is not completely silenced because if one knows

---

[150] *See* David Hoffman, Paul Bruening & Sophia Carter, *The Right to Obscurity: How We Can Implement the* Google Spain *Decision*, 17 N.C. J.L. & TECH. 437, 458 (2016) ("[T]he result is . . . much more about obscurity than it is about a right to be forgotten.").

the URL of the offending article, one can still access it; but of course, the point of the delisting is that without a search engine link most people will not be able to find it.[151]

2. Threats to the Global Public Good of the Internet

Second, the right to be forgotten threatens the global Internet because the concern is that courts will eventually require global delinking as the appropriate remedy.[152] Currently, if Google violates the right to be forgotten in France (for example), Google delinks the offending article on Google.fr, the French language search engine directed at French users, as well as all European Google sites (such as Google.de, Google.es, etc.).[153] It uses geographical filtering to identify people whose IP address suggests that they are making queries from France (or from another EU country).[154] In 2016 Google announced an additional remedy: it would use geographical filtering to prevent people located in France (or in another EU member state) from accessing search engine listings on *any* Google affiliated site, including Google.com — the general English language Google search engine that serves the United States.[155] People outside the EU consulting Google.com, however, would not be affected.

---

[151] In other respects, however, the remedy does not even produce practical obscurity. The CJEU order only requires Google to delink search results that appear when someone types Mr. Costeja González's name, not when they make any other query. *See Google Spain SL*, Case C-131/12, ¶ 94 (prohibiting "the inclusion in the list of results displayed following a search *made on the basis of his name* of the links to web pages published lawfully by third parties and containing true information relating to him personally" (emphasis added)).

Finally, the many stories and articles about the right to be forgotten and the *Google Spain* litigation that mention Mr. Costeja González are unaffected by the decision in *Google Spain*, including this law review Essay.

[152] The European Court of Justice is currently considering this question. *See* Alex Hern, *ECJ to Rule on Whether 'Right to Be Forgotten' Can Stretch Beyond EU*, GUARDIAN (July 20, 2017, 5:19 PM), https://www.theguardian.com/technology/2017/jul/20/ecj-ruling-google-right-to-be-forgotten-beyond-eu-france-data-removed.

[153] *Transparency Report: Search Removals Under European Privacy Law*, GOOGLE, https://transparencyreport.google.com/eu-privacy/overview (last visited Oct. 6, 2017) ("We delist URLs from all European Google Search domains (google.fr, google.de, google.es, etc.) and use geolocation signals to restrict access to the URL from the country of the person requesting the removal on all domains.").

[154] *See id.*; *Portée du déréférencement de M.Plaignant appliqué par Google*, CNIL (Mar. 24, 2016), https://www.cnil.fr/fr/infographie-portee-du-dereferencement-de-mplaignant-applique-par-google (showing how right to be forgotten operates on different versions of Google accessed in different countries).

[155] Samuel Gibbs, *Google to Extend 'Right to Be Forgotten' to All Its Domains Accessed*

However, courts may decide that even geographical filtering is not enough. That is because French citizens could access Google.com when they are outside of France or use a virtual private network or proxy that makes it appear that they are outside of France. To prevent these people from accessing offending websites, Google must delink the websites worldwide.[156]

Such a remedy would undermine the global public good of the Internet. Nothing would then prevent other countries — pursuing their own speech regulation policies — from requiring global filtering, blocking, or delinking of speech that these countries wish to regulate

---

in EU, GUARDIAN (Feb. 11, 2016), https://www.theguardian.com/technology/2016/feb/11/google-extend-right-to-be-forgotten-googlecom ("If a German resident successfully requests Google remove a search result under queries for their name, the link will not be visible on any version of Google's website, including Google.com, when the search engine is accessed from Germany. Google will use the browser's IP address to determine their location.").

[156] *See, e.g.*, Google Inc. v. Equustek Solutions Inc., [2017] S.C.R. 34 (Can.). In *Equustek*, the Supreme Court of Canada ruled that Google must delink search results not only in Canada, but everywhere else in the world. *Id.* The case involved a dispute over intellectual property involving the sale of pirated products, and the trial court had issued a preliminary injunction requiring delinking pending the resolution on the merits, which in practical effect was a permanent injunction against Google. *Compare id.* ¶ 17 (granting an interlocutory injunction against Google to remove all of a company's websites from its worldwide search engine), *with id.* ¶ 62-64 (Côté and Rowe, JJ., dissenting) (arguing that the remedy against Google was inappropriate because the injunction was effectively a permanent injunction and thus subject to a different test with a higher burden than an interlocutory injunction).

Justice Abella, writing for the majority, argued that the case did not involve free speech questions: "This is not an order to remove speech that, on its face, engages freedom of expression values, it is an order to de-index websites that are in violation of several court orders. We have not, to date, accepted that freedom of expression requires the facilitation of the unlawful sale of goods." *Id.* ¶ 48. Justice Abella noted that "If Google has evidence that complying with such an injunction would require it to violate the laws of another jurisdiction, including interfering with freedom of expression, it is always free to apply to the British Columbia courts to vary the interlocutory order accordingly. To date, Google has made no such application." *Id.* ¶ 46.

The question left open in *Equustek* is what happens if a court regards an issue as not concerning freedom of expression, but other countries disagree. *See* Michael Geist, *Global Internet Takedown Orders Come to Canada: Supreme Court Upholds International Removal of Google Search Results*, MICHAEL GEIST (June 28, 2017), http://www.michaelgeist.ca/2017/06/global-internet-takedown-orders-come-canada-supreme-court-upholds-international-removal-google-search-results ("[T]he court has effectively concluded that those seeking global takedown orders do not need to canvass the laws in other countries to consider the potential for conflicts with their request. In doing so, it places the obligation on intermediaries such as Google and increases the likelihood that those companies will pick and choose among the orders they are willing to follow.").

or censor. By promoting its parochial interests, each country will restrict access for end-users around the world.

The result will be a race to the bottom (or to the top, depending on how you look at it). Currently the Internet is mostly governed by the values of the least censorious regime — that of the United States. If nation states can enforce global filtering, blocking, and delinking, the Internet will eventually be governed by the most censorious regime. This will undermine the global public good of a free Internet.

3. Coopting Private Governance

Third, the right to be forgotten is an example of how nation states (and in this case, the European Union) have tried to coopt private infrastructure owners and their capacities for private governance. It is an example of how new school speech regulation trades on and seeks to coopt the emergence of private governance.

It is not enough to create a right to be forgotten. One must also enforce it in practice. It is impractical for the individual states of the European Union to administer a system of individual hearing requests for delinking. Too many people would make requests. The European Union needed a bureaucracy with sufficient technical capabilities to perform this function for them. The bureaucracy that the European courts have chosen to administer the right to be forgotten was none other than Google itself. The European authorities have required Google to set up an internal system for receiving and reviewing requests for delinking, which, if denied, are appealable to the national data privacy supervisory authority or to the courts.[157] In other words, the European Union has deputized Google to create a bureaucracy within Google that will administer the right to be forgotten in the first instance. The assumption is that Google will settle most of the claims at this level, thus conserving government resources.

---

[157] *Factsheet on the "Right to Be Forgotten" Ruling (C-131/12)*, *supra* note 84. A Frequently Asked Questions factsheet on the Right to Be Forgotten issued by the European Commission explains that complaining individuals should make initial requests to Google. *Id.* At that point, "Google will have to assess deletion requests on a case-by-case basis and to apply the criteria mentioned in EU law and the European Court's judgment." *Id.* The search engine company can turn down the request if it concludes that the right to be forgotten does not apply and/or the public interest requires that the links should continue to appear. *Id.* At that point, the complainant can "complain to national data protection supervisory authorities or to national courts. Public authorities will be the ultimate arbiters of the application of the Right to be Forgotten." *Id.*

These developments represent new school speech regulation and public/private cooptation taken to their logical conclusion. By creating a right to be forgotten, the European Union has, in effect, deputized a private organization to become its governing agent. The state (or in this case, the EU) coopts Google's growing system of private governance to turn it to the state's ends.

This possibility, however, was already implicit in a system of collateral censorship. Ordering a private infrastructure provider to block illegal or harmful content was effectively an order to develop the technical capabilities — through algorithmic decisionmaking, through bureaucracy, or through some combination of the two — to carry out the state's orders. New school speech regulation tends towards commandeering or at the very least coopting the new governors of online speech.

Private companies like Google and Facebook have an ambivalent relationship to these developments. On the one hand, they may think it better that they, and not government bureaucrats, create and apply the new system. This gives them a greater say about how the right will operate in practice. Moreover, in the long run, rather than merely limiting these companies, it also empowers them. They become indispensable to nation states and to the system of speech regulation. On the other hand, the easier it becomes for nation states to coopt private speech platforms to do their work for them, the more likely they are to make additional demands in the future.

The promise, and the danger, of the evolution of new school speech regulation is a world in which large, global, privately-owned platforms become the regulatory agents of nation states. The more these businesses regulate, the more indispensable and powerful they become to the nation states that purport to regulate them.

B. The Problem of Fake News

Now consider the problem of fake news. Fake news travels through social media. Suppose one believes that fake news is a genuine problem for democracies. How should one regulate it? The least efficient way is for states to pursue the people and organizations that produce the fake news — however that term is defined. Most of these people are anonymous, or outside the country, and many of them employ armies of bots.

Instead, the government might put pressure on parts of the digital infrastructure — most likely social media companies — to solve the problem. In other words, a governmental remedy for the problem of fake news would likely be some form of new school speech regulation

that encourages social media companies to develop new forms of private governance. Social media companies would be directed to identify and surveil fake news stories and producers, block links to fake news stories and fake news sites, or else supplement them with clarifying and counteracting material.[158] Such blocking, filtering, and surveillance, of course, raises all of the problems of collateral censorship and digital prior restraint that I mentioned before.

To be sure, governments might engage in jawboning and social pressure rather than employ direct mandates. Moreover, just as in the case of hate speech or abusive speech, pressure for private regulation will not only come from governments. It may also come from end-users. This second form of pressure is especially important because it is a feature of community governance.

Social media companies may not even need to take the hint. They may decide as a public relations matter, or as a routine part of the governance of their online communities, they will take various steps to counteract fake news. These solutions might include curating news feeds, making purchases of advertisements more transparent, marking suspected links, or supplementing suspected links with suggested alternatives. Facebook has already announced policies along these lines.[159]

Pressure from governments, citizens, and end-users to respond to the problem of fake news exemplifies the evolution of private digital governance of speech in the twenty-first century. To solve a perceived problem of speech regulation, a wide variety of public and private

---

[158] *See, e.g.*, Emma Llansó, *German Proposal Threatens Censorship on Wide Array of Online Services*, CTR. FOR DEMOCRACY & TECH. (Apr. 7, 2017), https://cdt.org/blog/german-proposal-threatens-censorship-on-wide-array-of-online-services (arguing that a proposed German bill to combat fake news "would create massive incentives for companies to censor a broad range of speech").

[159] Rob Goldman, *Update on Our Advertising Transparency and Authenticity Efforts*, FACEBOOK NEWSROOM (Oct. 27, 2017), https://newsroom.fb.com/news/2017/10/update-on-our-advertising-transparency-and-authenticity-efforts (announcing new policies on advertisements); Will Oremus, *Facebook Is Finally Fixing the Ad System That Let Russia Secretly Influence Elections*, SLATE (Oct. 27, 2017, 7:55 PM) http://www.slate.com/blogs/future_tense/2017/10/27/facebook_is_finally_fixing_the_ad_system_that_let_russia_secretly_influence.html ("Facebook announced . . . that it's beginning to implement a slew of new policies aimed at making its ads more transparent."); Sara Su, *News Feed FYI: New Test with Related Articles*, FACEBOOK NEWSROOM (Apr. 25, 2017), https://newsroom.fb.com/news/2017/04/news-feed-fyi-new-test-with-related-articles/ (announcing test of new system); Kaya Yurieff, *Facebook Steps Up Fake News Fight with 'Related Articles,'* CNN TECH (Aug. 3, 2017, 2:27 PM), http://money.cnn.com/2017/08/03/technology/facebook-related-articles/index.html (reporting that Facebook "will also include articles fact-checked by third-party sites such as Snopes and PolitiFact, which employ editors . . . intend[ed] to help users think twice about whether a story is true").

actors urge infrastructure owners — in this case, social media companies — to develop their own programs, algorithms, and bureaucracies, and to help end-users make decisions about what kinds of news stories they should read and trust. In other words, our new pluralist system of speech regulation encourages platform owners to develop ever more extensive and elaborate systems of private governance.

CONCLUSION: NEW SOCIAL OBLIGATIONS FOR DIGITAL MEDIA COMPANIES

This is the direction of the future, and it is by no means guaranteed to be free speech friendly. From the standpoint of free speech values, the best solution would be for large international infrastructure owners and social media platforms to change their self-conception. Ideally, they would come to understand themselves as a new kind of media company, with obligations to protect the global public good of a free Internet, and to preserve and extend the emerging global system of freedom of expression. Defenders of democratic values should work hard to emphasize the social responsibilities of digital infrastructure companies and help them both to understand and to accept their constitutive role in the emerging global public sphere.

This is not the first time this has happened. In the twentieth century, the norms of American journalism changed. In the 1890s newspapers were still rabidly partisan.[160] In the early twentieth century, influenced by Progressive era reforms, newspaper publishers and reporters gradually recognized that they (and their competitors) had social responsibilities to the public as a whole rather than to political parties, and over time they developed the professional norms of objectivity that we now think of as the goals of properly trained professional journalists.[161] Modern notions of journalistic objectivity and fairness, however, did not emerge fully until the 1920s.[162] They arose in response to pressures both from within and external to the profession; journalists perceived threats both to journalism and to democracy from the rise of political propaganda and public relations campaigns in the 1910s and 1920s.[163] Journalists "grew self-conscious

---

[160] *See* Michael Schudson, *The Objectivity Norm in American Journalism*, 2 JOURNALISM 149, 159-60 (2001).

[161] *Id.* at 160-61.

[162] *Id.* at 161.

[163] *Id.* at 162 ("Journalists had rejected parties only to find their new-found independence besieged by a squadron of information mercenaries available for hire by

about the manipulability of information in the propaganda age"[164] and the deliberate use of disinformation by propagandists and publicity agents. They developed new norms and social responsibilities as a result.

A new set of social responsibilities confront new media companies in the twenty-first century. There is reason to hope that the story of the early twentieth century will repeat itself in the twenty-first.

In this Essay, I have identified some of the characteristic free speech problems of the early twenty-first century, and a set of new ideas to help people understand them: information fiduciaries, algorithmic nuisance, old school versus new school speech regulation, public/private cooperation and cooptation, and finally the idea of private governance of speech and of speech communities. These are the central concepts for thinking about the free speech problems of the twenty-first century.

The question is not whether digital speech should have these features. They are already with us. The question is what kind of private governance we will have, how states will employ new school techniques, how they will attempt to coopt infrastructure owners, and how infrastructure owners will respond. Digital infrastructure owners are not in precisely the same situation as twentieth century mass media enterprises. But like them, they must take up a new set of social obligations to preserve the global public good of a free Internet and a healthy and vibrant global public sphere.

---

government, business, politicians, and others.").

[164] *Id.* at 162.