

Introduction

The Perils of Platform Misgovernance

There's no way that Mark Zuckerberg could have imagined, when he created an electronic version of Harvard's traditional collection of student photos – informally called “the facebook” at least since I was in law school there in the 1990s – that a decade and a half later he would have the blood of genocide victims on his hands.

In 2016 and 2017, the Myanmar government carried out a genocide against the country's Rohingya Muslim minority.¹ At the time, Facebook was the overwhelmingly dominant internet platform in Myanmar, with one report commissioned by Facebook itself observing that “[t]here are equal numbers of internet users and Facebook users in Myanmar” and “many people use Facebook as their main source of information.”² This is in part because Facebook and local providers created subsidized forms of internet access such as “Free Basics” and “Facebook Flex” in order to expand their services in the country.³ Instrumental in the atrocities was propaganda which the military distributed over Facebook.⁴ Facebook acknowledged its role in 2018.⁵

The shock and chagrin that Zuckerberg must have felt when he learned of his culpability in the ethnic cleansing of the Rohingya people was probably pretty similar to how the leaders and founders of every major social media company

¹ Domino (2020, 150–1); United Nations Human Rights Council (2018); Beaubien (2018); *BBC News* (2020). On Facebook's culpability, see the independent assessment Facebook commissioned in 2018 of its human rights impact in Myanmar, by Business for Social Responsibility, published at https://about.fb.com/wp-content/uploads/2018/11/bsr-facebook-myanmar-hria_final.pdf in October 2018. The United States Department of State formally and publicly classified the events in question as a genocide in March 2022, only the eighth time since the Holocaust that it has made such a declaration. US Department of State, Genocide, Crimes Against Humanity and Ethnic Cleansing of Rohingya in Burma, www.state.gov/burma-genocide/ (last visited December 4, 2022).

² Warofka (2018, 13).

³ United Nations Human Rights Council (2018, 339–40); for further context, see Stecklow (2018).

⁴ Amnesty International (2022); *BBC News* (2018a); United Nations Human Rights Council (2018, 340–1); Hogan and Safi (2018); Gowen and Bearak (2017); Mozur (2018); Fink (2018). Other media were also implicated, such as the official state media of the country, Lee (2019).

⁵ *BBC News* (2018b).

felt on January 6, 2021. After years of the weaponization of misinformation and polarization by Donald Trump and his allies over social media, his regime finally culminated in an armed mob attack on the US Capitol for the purpose of preventing the peaceful transfer of power. While the major social media companies have knowingly or even intentionally inflicted many terrible things on the world, it strains credulity to suggest that they intended or even expected a coup attempt in the world's richest and most powerful democracy – and not incidentally the country in which all of those companies are headquartered, from which their founders mostly originate, and under whose political order those companies and their leaders have prospered.

This book begins with those two examples, among the many ways in which social media companies have failed their users and the world, because their very extremity highlights a point that will run through this book: *nobody* (except their perpetrators) *wanted those two things to happen*. Mark Zuckerberg might be a bad guy. He may be the face of a business model, which often goes by the name “surveillance capitalism” (e.g., Zuboff 2019) and which causes numerous individual and social harms. But he’s not Hitler. Imagine that a genie had appeared before Mark Zuckerberg and said “you can give up some moderate proportion, say 15%, of your wealth from Facebook and, in exchange, I will magically cancel the genocide that you’d otherwise be culpable for.” I imagine that he would have taken the deal.

This isn’t an apology for Mark Zuckerberg. I don’t really care what you think about him. But this book does assume that he doesn’t want to be responsible for genocide, and that he would be willing to spend a substantial (though perhaps not infinite) amount of Facebook’s money to avoid that fate, if only he knew how to do so. The same goes for the then-leaders of all of the major companies and the coup attempt on January 6.⁶ To avoid being culpable in the next genocide or coup, they need to do a better job at getting control of what happens over the services they run, that is, “platform governance.”

This book focuses on problems, like genocide and coup attempts, where the interests of companies and their leaders are aligned with the interests of the rest of us – ordinary people across the world and our (decent, liberal-democratic) governments. It rests on the assumption that there is a substantial amount of social harm caused in that territory of interest alignment, such that companies have reasons (whether moral or financial) to work together with people and governments to redesign the services those companies provide in order to address those harms. This book sketches out one way in which we might do so.

I take no position on any of the broader questions surrounding these services and the extent to which their interests might conflict with those of the rest of us. Social media in particular has been subject to a sustained critique rooted in its

⁶ Alas, I can’t confidently say the same about the leaders of rogue minor companies like Parler and Gab, or about the leader of Twitter at the time of this writing, Elon Musk.

revenue model and the way that revenue model encourages companies both to radically undermine individuals' autonomy over the details of their lives and to promote thoughtless, emotion-driven, content in pursuit of the goal of "engagement" (and thereby advertising dollars). I don't purport to evaluate that critique, or to attempt to balance the harms social media generates against the benefits of interpersonal connection and communication which it offers. First, let's solve the genocides and coup attempts, and then we can worry about surveillance capitalism. If I can contribute, even in a small way, to reducing the risk of another Myanmar genocide or another January 6, then I will consider this book a success.

In the pages that follow, I offer a program, built on the insights of political science and allied fields, for radically democratizing services like social media – more broadly, "platforms." We should directly insert ordinary people, including ordinary people from the Global South, from minoritized and indigenous communities in the North, and from other subordinated and excluded groups, directly into the processes of platform rule enforcement, rule development, and ultimately product design. I argue that such innovations would actually be in the long-run interests of platform companies along with the rest of us.

In a nutshell, here's what I will propose (with the details reserved for Chapter 6). Companies, assisted where possible and coerced where necessary by liberal-democratic states, will create a multilevel participatory governance organization, organized along lines of geography as well as identitarian affinity.⁷ Through that organization, randomly selected and well-paid groups of ordinary people will have (a) access to internal company information; (b) privileged channels of communication to companies about their own observations of the local impact of platform policies; (c) some degree of retail-level control over company governance decisions with respect to users (e.g., in the social media context, appellate authority over content moderation decisions); and (d) some degree of wholesale-level control over company policies through the ability to propose (and in some cases veto) rule and product design changes.

This proposal should be backstopped by certain legal interventions, including the judicious use of antitrust and workplace rights law and the subjection of the largest companies to elements of the international human rights regime (mainly the responsibility to protect). However, at its heart, the incentive for companies to participate is that doing so is in their own interests. The tasks that fall under the rubric of "platform governance" as defined in this book – at least within the domain in which company interests and public interests are aligned – are tasks that companies actually need to carry out well in order to protect their own long-term economic viability. But, this book argues, they are persistently hampered in doing so

⁷ That is, I envision local and identity-based first-level groups – also including company employees as individuals – which nominate members to composite second-level groups with more authority, and so forth. Company participation in this system will begin with larger companies and social media companies, with expansion to smaller companies and other types of platforms over time.

by (a) their inability to give people who are socially distant from their personnel the capacity and incentive to supply them with the information they need to frame and implement their rules, and (b) their self-control problems in resisting short-sighted incentives relating to threats (especially from governments) and destructive methods of short-term profit seeking. Giving up some control to ordinary people across the boundaries of geography, hierarchy, and affinity can alleviate both problems at once: their self-control problems by separating those who care about short-term costs and threats from those who control the decisions which implicate them; their information problems by creating channels of communication between the peripheries of a company's domain and its organizational nerve centers as well as making the use of those channels effective at achieving the ends of those at the peripheries, and hence worth doing for them.

A major challenge in writing a book-length work on academic timescales in this domain is that the platform economy and its surrounding social, political, and legal landscape tend to change at astonishing speed. Between the penultimate draft of this manuscript and its final, for example, Elon Musk acquired Twitter, at which point the company went from having some of the most thoughtful and innovative work on platform governance to operating purely arbitrarily.⁸ Alternative platforms, some with radically different organizational structures (such as the Mastodon federated model), suddenly became popular. Around the same time period, the Fifth Circuit issued a decision purporting to uphold a Texas law prohibiting social media platforms from engaging in "censorship" on the basis of viewpoint – obviously motivated by efforts by politicians who are supported by the extreme right to require platforms to host misinformation and hate speech.⁹ By the time this book sees print, the Supreme Court may have reversed that decision – or it might be upheld, and the enterprise of social media content moderation may be effectively dead, at least as it relates to partisan misinformation. That sort of radical change seems to happen every time one looks up from one's keyboard when one tries to write about platforms.

This book aims to be timely insofar as it takes account, as best as possible, of the current state of play in the operation of platforms and offers advice that is potentially actionable by existing governments and companies for mitigating their governance problems. However, it also aims to be timeless – and hence unlikely to become obsolete the moment it hits the shelves – insofar as it offers a general account of the sorts of governance problems that platforms face, regardless of which companies

⁸ Incidentally, unless otherwise specified, discussions of specific governance features or policies of Twitter in this book refer to the state of affairs before the Musk acquisition, when the company was making a serious effort to conduct platform governance. At the time of this book's completion, matters on Twitter under Musk are too chaotic to fully take into account.

⁹ *NetChoice v. Paxton* No. 21-51178 (5th Cir., September 16, 2022). That decision was, to put it bluntly, utterly clueless – its analysis of companies' First Amendment interests completely neglected the well-recognized role of content moderation in their core business models.

happen to be in operation on a specific date, and an account of how that arises from the nature of the services they provide. From that perspective, the descriptions of individual companies and incidents given in this book should be taken primarily as evidence for the general structural account of the problem. I hope that in the future this book may even serve as a contribution to the discipline of political science more broadly by providing an illustration of the incidence of state-like problems of governance in nonstate contexts.

The price for this effort to speak to the present as well as the future and to policymakers and company personnel as well as to scholars is that none of its audiences will be fully satisfied. Policymakers and company personnel will have cause for complaint that this book lingers far too long over excursions into the scholarly literature, and its recommendations will not be finely tuned to their most immediate problems but will require adaptation to be useful. Scholars will have cause for complaint that this book sacrifices the extended theoretical development of fewer ideas and in-depth interaction with the literature in favor of a somewhat more brisk development of more ideas that can come together into relatively concrete recommendations to capture that low-hanging governance fruit which I just described. To each of those groups, I acknowledge your complaints and trust and hope that the payoff – real design and policy recommendations backed by decades of scholarship for officials in companies and states, and a novel application of existing tools in political science to new contexts for scholars – warrants your patience.

WHAT IS “PLATFORM GOVERNANCE” ANYWAY?

Before we get going, I should specify the scope of the argument. This is a book about platforms and about governance, and the merger of those two things that has come to be called “platform governance.” Standing on its own, however, that sentence indicates surprisingly little because of the capaciousness of all those ideas.

First, “platform.” I’m going to hazard a more careful definition of a “platform” in Chapter 1, but for present purposes, we can take it to mean internet enterprises that facilitate interactions between individuals, where the economic model of the enterprise is that the company operating the platform captures some of the value produced by that interaction.¹⁰ This description captures two-sided markets like eBay, Uber, Apple’s App Store, and Amazon’s Marketplace, which take commissions on transactions; as well as social media companies like Facebook and Twitter, which

¹⁰ This is generally similar to other extant definitions of platforms in the scholarly literature across several fields. Bonina et al. (2021, 871) helpfully review recent definitions along these lines, as does Jin (2017, 7–10). In the terms of Bonina et al., this book focuses on “transaction” platforms (which encompasses both social media and many-to-many marketplaces like Amazon and eBay) rather than “innovation” platforms, although I shall reserve the term “transactional” for platforms that primarily focus on buying and selling rather than social interaction. There has also long been talk of hardware “platforms,” such as the iPhone. Those are entirely out of scope for this book.

capture some of the value from users producing and viewing one another's content by sticking ads on the top and mining the data they generate. In this book, I'm primarily concerned with the social media kind, with which I'm most familiar and which seem to lately be causing the direst social impacts. But many of the propositions advanced in these pages will be oriented toward the abstract characteristics of platforms and will apply reasonably well to the transactional kind too.¹¹

The difference between transactional and social media platforms has the potential to vex any academic work attempting to analyze them together. One challenge which might fairly be raised against the entire project is to question whether they even bear enough in common to be understood together – do commentators and scholars just lazily use the word “platform” to describe both kinds of enterprises? Are not the problems posed by each – such as evasion of ordinary commercial regulation for the transactional kind and political polarization and disruption for the social kind – distinct?¹²

I will attempt to answer that worry through the back door with the additional definitional work in Chapter 1, but here we might just notice that many of our major companies have demonstrated a striking tendency to leverage their assets to operate across these business models. For example, Meta's major properties, Facebook and Instagram, both operate transactional marketplaces – Facebook's focused on eBay-like individual-to-individual transactions, and Instagram's on business-to-consumer transactions.¹³ I submit that the reason that this seems like a good idea is because the core properties of a successful social media company, in terms of the scale of its userbase and the incentives it offers for activity, are also the core properties of a successful transactional platform. Moreover, the basic criteria for successful design are similar across these platforms. For example, Gillespie (2018b, 211–12) identifies that these two types of platforms offer something like algorithmically curated and moderated user-generated content within the core of their business.

As a first pass, we can say that “platform” is an abstract description of an economic model focused on facilitating a variety of kinds of third-party interactions, while “social

¹¹ The big Chinese platforms, such as Sina Weibo and WeChat, are outside the scope of this book. The goals and challenges of and tools available to a largely single-state platform operating under the thumb of a world-historically sophisticated and effective autocracy are wholly distinct from those of a platform with a global userbase operating out of a liberal democracy.

¹² For example, Lobel (2016, 94–95) suggests that social media is merely a “prelude” to the real platform companies, like Uber.

¹³ Incidentally, it is quite inconvenient that two of the major platform companies, Facebook and Google, changed their names while maintaining the original name for a subset of their original businesses, during the course of the events described in this volume. Generally, I will use “Facebook” and “Meta” interchangeably, and the same goes for “Google” and “Alphabet.” However, I will endeavor (with only moderate consistency) to primarily use “Meta” for references to elements of the company formally known as Facebook in contexts of continuing operation, for example, the “Meta Oversight Board” rather than the “Facebook Oversight Board.” It should be clear from context when I mean to refer to Facebook the service in contrast to other Meta services like Instagram, and to Google the service in contrast to other Alphabet services like YouTube.

media” describes a particular kind of interaction facilitated by some platforms. To the extent the problems of governance in social media are rooted in the abstract economic model, they will be shared with other kinds of platforms like Amazon and eBay; however, some problems of social media governance arise in particular from the *communicative* character of the interactions social media platforms facilitate.

I also think it is a mistake to be too strict about one’s definition of “platform.” In reality, “platform” is probably a Wittgensteinian family resemblance concept rather than susceptible to definitive criteria of inclusion and exclusion.¹⁴ There are characteristic properties of platforms, but not everything which we might want to call a platform will have all those characteristic properties. For example, the outer limits of the concept of a social media platform on the definition I favor are probably at Google Search, which I consider to fit within the definition insofar as users are both creators of and consumers of search results. (They are creators both by creating the underlying web pages and other content and by creating the links and other activity which Google relies upon for search rankings.) Moreover, Google Search experiences the characteristic problems associated with social media, such as user gaming of recommendation algorithms to promote low-quality content. Finally, Google Search monetizes activity in exactly the same way that Facebook or YouTube does, that is, by using behavioral data to create recommendation algorithms that give users an incentive to make more use of their service, and by using predictions from that behavioral data to target advertisements.¹⁵ In order to distinguish platforms from more longstanding business models such as brokers of various sorts, it will also help to focus our attention on the kind of novelty that platforms generate – platforms tend to create new kinds of interactions between people or transform existing kinds of interactions in fairly dramatic ways.

Why and How Do Platforms Govern?

This leads to the notion of “governance.” There are at least two distinct problems of “platform governance,” though, as I shall argue, they are closely related. First is the governance *of* platforms, that is, of platform companies, by governments.

¹⁴ I mean to invoke the weaker version of the notion of a family concept as described by Wennerberg (1967, 109–10).

¹⁵ However, some important clusters of network affordances, some of which are even controlled by platform companies, probably don’t meet any formal definition of “platform” which we might want to adopt. We might call them “quasi-platforms.” For example, WhatsApp seems to have many of the characteristic problems of social media platforms, such as viral misinformation, but it lacks many of the standard features – because data flowing over it are encrypted, Meta has limited opportunities at best to monetize those data; as far as I know, it doesn’t feature recommendation algorithms in any significant sense; the fact that it does seem susceptible to things like viral misinformation seems, as far as I can tell (and this is with extremely low confidence), to be a product of the combination of its dominant position in certain communicative markets as well as more user-interface style affordances like the capacity to forward messages to many people at once and, perhaps, a certain degree of immersiveness not shared by, for example, email. At any rate, we can pretend that WhatsApp is a platform to the extent the governance techniques described in this book might be useful for it, but ignore it otherwise.

There are numerous debates about the extent to which we should regulate those companies, and the manner in which we should do so; for example, how much should we require them to protect the data of their users, or forbid them from combining that data in certain ways? Should they be required to offer users some degree of interoperability in terms of being able to move data back and forth between them?¹⁶

However, there is also the problem of governance *by* platforms – that is, the regulation of the behavior of users of platforms, such as buyers and sellers in Amazon marketplace or people posting and reading tweets.¹⁷ It is now commonplace to recognize that platforms are engaged in acts analogous to public governance when they regulate user behavior – thus, for example, many scholars have identified the quasi-governmental character of social media speech regulation, and have further identified that this is intrinsic to the products companies offer (Klonick 2018, 1638–30; Gillespie 2018b). One scholar (Eichensehr 2019) has gone even further, comparing companies to “Digital Switzerlands” that compete for authority with physical-world governments; a claim that may seem overheated but for the fact that the author leads her article with a citation to the president of Microsoft arguing for just such a thing (albeit in the limited domain of protecting their customers from cybersecurity threats).¹⁸

In the process of exercising governing power, platform companies routinely adopt both the methods and the personnel of government regulators, hiring lawyers – many with prosecutorial or other government experience – to write lawlike rules which they enforce with formal processes. The apotheosis of this trend is perhaps Meta’s Oversight Board, colloquially called its “Supreme Court” since before it was even created.

This degree of government-like organization and government-like behavior is an outlier in contemporary capitalism. The point should not be overstated: As Rory Van Loo (2016) has shown, corporations frequently offer dispute resolution services (consider credit card chargebacks as the canonical example). But the degree of lawlike formalization in the platform economy seems unique.¹⁹ When my bank decides whether or not to offer me a line of credit, I don’t have the benefit of a

¹⁶ I use “governance” and “regulation” interchangeably.

¹⁷ See Gorwa (2019a, 855); Gillespie (2018c) for the governance *of*/governance by distinction.

¹⁸ See also Srivastava (2021, 7–8), who articulates a similar idea from an international relations perspective; and Cohen (2019, 129–31) giving an example of a tug of war over surveillance which positions companies as both defenders of the public against state surveillance and themselves agents of both private and state surveillance.

¹⁹ Another important potential counterexample is supply chain regulation, in which companies control the behavior of their suppliers for social goals (albeit driven by their business interests, such as consumer demand), in fields such as labor rights and environmental protection. Green (2014, 1–2) cites Walmart’s sustainability rules as a major source of regulation for numerous global producers. Van Loo (2020) gives a variety of other examples in that vein. Still, regulating the other business firms with whom one deals is a different ballgame from regulating a mass public.

published set of rules to which I may appeal in litigating their decision before a formal process within the company. Shopping malls don't have their own codes of laws to justify the decisions of their security personnel to kick out rowdy shoppers. Even credit card chargebacks don't feature anything like a system of appeals or a published set of rules meant to draw a balance between the interests of customers in getting what they purchased and the interest of retailers in avoiding fraud and manipulation.

Notably, these other businesses offer their customers less of the protection of formal legalism even though the individual stakes for those customers are typically much higher than at least the social platforms. If I have one of my Facebook posts removed, I typically don't experience a serious injury.²⁰ By contrast, if some merchant defrauds me and Chase refuses to reverse the charge on my credit card, I could lose lots of money. This suggests that the increasing formality of platform user-governance decisions is not simply a response to the stakes involved for their users.

Rather, I claim that lawlike, formal, governance methods respond to the sorts of decisions that platform companies have to make. Lawlike forms of governance aren't just chosen at random; rather, human societies have developed social technologies in the most foundational sense, like the independent[-ish] judge and the written code of laws, because those are effective at solving certain kinds of governance problems (Gowder 2016, 40, 59–62; 2018b, 91).

The existence of platform law and platform law enforcement is a kind of return to an (alleged, albeit highly controversial) earlier day of weak states in which private law was required to fill in the gaps for the purposes of facilitating things like trade.²¹ With that precedent in mind, it's easy to start by observing that part of the explanation for the lawlike form of platform governance proceeds from the global nature of the problems posed by the largest of such platforms and the difficulty domestic governments have in controlling them (consider our experience in the United States with counterfeit products from China and election interference from Russia).

But global scale isn't the only reason these platforms occupy a kind of governance role, and there are lots of global companies that don't. Rather, I think it is the combination of scale, vast diversity (explored further in Chapter 1), and the fact that platforms inherently (indeed, as part of the definition I've lightly sketched so far) create surfaces for interactions between third parties which by their very nature

²⁰ While I might experience a serious injury if I ask a company to remove someone else's post and the company says no (e.g., if that other person's post includes my private information), most of the lawlike protections social media companies offer are directed at protecting the interests of the poster, not the complainer. For example, when the Meta Oversight Board was first created, there was no way to appeal the company's refusal to take down content someone else produced, only the company's decision to take down content one has produced. (This appears, however, to have changed.) Similarly, there's an appeals process for YouTube creators to seek review of platform “strikes,” but as far as I can discern there is not one for people who report policy-violating videos.

²¹ For example, Milgrom, North, and Weingast (1990); for a skeptical take, see Kadens (2015).

enable and promote some interactions and disable and deter others that drives the phenomenon. The need to regulate (i.e., govern) the behavior of the parties using their platforms, at least in their own interests, is built into the economic model itself: For platform companies to make money, there must be activity to monetize; for that activity to be sustainable in the long term, users must on the whole understand themselves to be experiencing positive outcomes from their usage of a platform.²² But human sociality has the unfortunate habit of turning vicious on a regular basis, and a platform too-plagued by viciousness (as its users understand it) will not be able to keep its users. For the simplest example: if Amazon lets buyers and sellers rip one another off with abandon, before long there won't be any transactions on their platform for the company to take a cut of.

That is the core dynamic driving the distinctive enterprise of “platform governance” and it's the reason that things like credit card chargebacks don't count: Amex doesn't have to worry that people will stop using its credit cards if merchants rip them off, because the credit cards don't (except in the special case of things like skimmers) have any particular connection to the rip-offs – a person sold a piece of garbage by their local store would have gotten just as ripped off if they'd paid in cash. Accordingly, Amex's business model doesn't directly depend on governing transactional honesty (though it might be able to obtain a competitive advantage by doing so, and doing so effectively, or government regulators might impose it) in the same way that Facebook's and eBay's business models depend on governing what shows up on their platforms.

“Platform governance” in the sense of governance-by-platforms is helpfully divided into three subcategories, which we can, at a first pass, call “organizational governance,” “architectural governance,” and “regulatory governance.”

In the first category falls choices made about the internal structure and processes of an organization itself, such as the organization of decision-making responsibility among employees for the exercise of authority over behavior using the platform. Traditional “corporate governance” falls within this category, but so do novel innovations such as Facebook's creation of a content moderation oversight board (Douek 2019; Klonick 2019).

In the second category would be what Lessig (1999) described as the regulatory capacity of code or architecture. While code may (and often does) unintentionally regulate, we should limit terms like “architectural governance” to the intentional modification of the affordances made by networked internet platforms in order to control behavior. For example, Facebook has experimented with restricting the ability of non-posters to see the “like” counts on posts, apparently in order to

²² I leave aside here the problem of platform addiction, which could be understood as users fearing negative experiences from leaving the platforms to which they've been habituated. However, the possibility of addiction or of less psychological and more economic analogues (such as lock-in due to high switching costs for platforms that provide important services) imposes some limit on the scope of this claim.

reduce the behavioral incentives supplied by the visibility of “likes” (Constine 2019). One of the persistent features of platforms due to their artifactual character is the fairly blurry boundary between product design and governance (further discussed below).

Finally, the third category describes the most explicit forms of platform governance, by which platform operators make (stated or unstated) rules that divide conduct into permissible and impermissible behavior. Such rules may be enforced by human beings making decisions or by automated processes, and can be backed by sanctions, such as the removal of offending content, offending products or listings on transactional platforms, or offending users altogether (“bans”) as well as a variety of other “remedies,” such as the downranking of content in algorithmic feeds or the removal of a user’s capacity to earn money on their content (Goldman 2021).

Of course, the boundaries between these categories are inevitably fuzzy and unstable. Sometimes internal governance structures may be compelled by external actors, or may be adopted in order to stave off regulation by those actors. Facebook’s content moderation oversight board could be described as a form of organizational governance or a form of regulatory governance. Even architectural governance is not completely bounded, as, for example, the affordances available to users may be modified in the course of a platform’s exercise of regulatory governance, as when a platform disables certain features in order to control user behavior; moreover, architectural governance is backstopped by law itself.²³ Nonetheless, keeping the three broad categories of platform governance in mind will help in clarifying our thoughts about the options for platform operators and states.

Another area of unavoidable ambiguity is the boundary between governance and ordinary operation of a platform. Platform operators may make the same categories of choices in order to control user behavior that is perceived to be harmful and in order to optimize for other desirable qualities. For example, Google may choose to rearrange the ranking of websites in its search results, or Facebook of posts in its News Feed, in order to display results beneficial to their revenue models for non-regulatory reasons – that is, to display more relevant search results or more engaging posts in order to drive more usage (and hence advertising revenue). But they may also reorganize their rankings in order to prioritize behavior considered harmless over behavior considered harmful, like the distribution of viral hoax content (e.g., Constine 2018; Hearn 2017). Moreover, those motivations may merge: It might be the case that some harmful content is also detrimental to engagement or to the relevance of search results. However, they may also conflict,

²³ Law forbids the use of technical means to evade architectural restraints, for example, as in the American Computer Fraud and Abuse Act, 18 U.S.C. 1030, which forbids at a minimum traditional “hacking”-type evasion of platform architecture, and potentially could be used (abused) to forbid a much broader class of activity, such as using adversarial machine learning examples to fool artificial intelligence systems (Calo et al. 2018).

as if it turns out that the viral hoax content is also the most engaging, and, hence, the most profitable. Governance is at least partly a matter of product design, and different governance decisions can vary depending on whether those making the decisions take, for example, a long-term or a short-term perspective on the health of a product.²⁴

Political Governments Can Help Platforms Govern

The point of intersection between the problem of governance of platforms and the problem of governance by platforms is that, as I shall argue in somewhat more detail in Chapter 2, one helpful way that governments and the democratic (hopefully) peoples behind them might govern platforms is by intervening on how they govern their users. Existing efforts to regulate users by regulating the platforms over which they interact are already familiar, particularly in the “intermediary liability” context most famous in the form of the US Digital Millennium Copyright Act’s notice and takedown process, which does not merely dictate that companies are not to host copyrighted content but also imposes specific processes that they must offer to their users for raising and disputing claims of copyright infringement in order to benefit from a “safe harbor” provision against the companies themselves being held liable for such infringement. In effect, then, internet companies have been pressed into service (on pain of their own copyright liability) as enforcers of copyright law against their users.²⁵

The example of the Digital Millennium Copyright Act also indicates some of the dangers of the government recruiting companies into a governance role. It’s fairly clear that the DMCA has led to overenforcement of copyright online, at least in some respects – relying on platform enforcement is a cheap and easy method for copyright holders to get relief – so cheap and easy that it tends to be overused, and individuals with legitimate claims of difficult to adjudicate rights like fair use find themselves struggling to protect those rights. Making matters worse, some platforms have voluntarily gone well beyond even the DMCA process – for example, YouTube has a “content ID” system that proactively identifies allegedly infringing content.²⁶ Thus, we have recently seen the atrocious spectacle of abusive police officers playing Taylor Swift songs in order to prevent citizens from exposing their

²⁴ This too is not dissimilar to the governance challenges faced by states; one way to read Mancur Olson’s (1993) famous article about “stationary bandits” is as an account of the governance consequences of lengthening time horizons (much more on this in Chapter 4).

²⁵ Generally, scholars have identified that “cooperative” governance across companies, states, and the public at large is necessary for the kinds of cross-national and complex entities that platforms are (e.g., Helberger, Pierson, and Poell 2018), while “collaborative” governance in which some discretion in governing decisions is shared between the government and private entities is common domestically as well (Donahue and Zeckhauser 2011).

²⁶ For a general summary of the role of platform companies in policing intellectual property on their own initiative, see Cohen (2019, 123–25).

official misconduct – on the evident theory that if concerned citizens post videos of police misconduct on YouTube, the content ID system will take them down because of the copyrighted music playing in the background (Schiffer 2021; Cole 2021). If the abuse of platform copyright enforcement to facilitate abusive policing is too depressing, here’s a more (bleakly) amusing example: In 2009, Amazon infamously discovered a rights glitch and memory-holed George Orwell’s 1984 from users’ Kindle devices (Stone 2009).

The DMCA can perhaps stand as the nadir of government efforts to recruit private companies to govern their users. Because US copyright law is notoriously captured by media companies (Lessig 2003), it shouldn’t surprise us that an intermediary liability framework built from it would give platforms an incentive to err on the side of total overenforcement. But it could easily get worse. For example, consider the contested relationship between the notion of “terrorism” and political dissent, and the fact that US law prohibiting “material support” for terrorism is infamously overbroad, already forbidding, for example, human rights organizations from advising allegedly “terrorist” organizations even on *lawful and nonviolent* ways to achieve their political goals.²⁷ It seems like it’s only a matter of time before we see an intermediary liability framework forcing platforms to deny services to allegedly terrorist groups or their supporters. Some scholars have argued that existing material support statutes could be interpreted to subject platform companies to criminal liability for hosting the content of such groups (e.g., VanLandingham 2017).

Similarly, a number of countries have followed a model pioneered by Germany’s NetzDG law and imposed DMCA-style requirements on companies to more broadly address prohibited content.²⁸ There is evidence that some governments also engage in informal DMCA-like use of platforms as a kind of cat’s paw to demand the suppression of content they believe to be illegal without the ordinary process imposed on states (e.g., Elkin-Koren 2022). This is similar to a variety of other ways in which governments can use private actors’ control over important social affordances as a method to informally impose sanctions on individuals without complying with their own internal rules; a prominent non-internet example would be the use by some American municipalities of “nuisance property” laws to bully landlords into evicting tenants whom authorizes deem to be troublesome (Gowder 2021, 176–77). Moreover, as Citron (2018) argues, this kind of cat’s-paw regulation can expand the scope of government authority not only by freeing it from procedural constraints, but also, in the social media context, by freeing it from geographic constraints as well as substantive legal constraints – she describes

²⁷ And somehow this is considered constitutional. *Holder v. Humanitarian Law Project*, 561 U.S. 1 (2010).

²⁸ See Article 19, “Germany: The Act to Improve Enforcement of the Law in Social Networks,” August 2017, www.article19.org/wp-content/uploads/2017/12/170901-Legal-Analysis-German-NetzDG-Act.pdf; Zurth (2021) and commentary from the EFF (Rodriguez 2021).

how the European Union, by threatening companies with regulation, has begun to build the ability to create extraterritorial effect not only for its speech laws but even for its extralegal speech policies.²⁹

The DMCA model of intermediary liability – in which governments decree what the rules are to be, and then make companies enforce them on pain of punishment – is not the way to bring together government and platform regulation. Rather, I shall argue that governments might helpfully intervene in platform governance of their users by assisting and giving companies incentives to develop robust, quasi-democratic, governance *institutions* to help create and enforce platform rules. This too is a familiar strategy for governments whose citizens are disadvantaged by the governance failures of others, as exemplified by multinational and international efforts to promote the rule of law in war-torn, transitional, developing, and failed countries. In its best form (which is sadly rarely achieved in the actual world, cf. Gowder 2016, 168–76), such projects represent efforts to actually achieve the benefits of effective government elsewhere without the cost to democratic legitimacy of simply imposing such a government on one’s own. And if done right, institutions that provide for at least semi-democratic kinds of governance can also protect against the kinds of capture that the DMCA exemplifies.

Platforms Need the Help: They Are Often Unable to Govern Their Users

Platforms have shown themselves sometimes unable and (at least in part) sometimes unwilling to adequately govern the conduct of their “citizens” – that is, their users. I’ll give a couple of examples from Facebook, just because it’s the company I know the best, having done some work for them. (If this troubles you, please refer to the Appendix to this Introduction for a discussion of potential conflicts of interest raised by my past work with Facebook, and why you should still believe what I say.)

First unable: return again to the genocide in Myanmar. I discuss the problems that contributed to this disaster at length in Chapter 3, but in summary, we can fairly say that Facebook exacerbated the violence by its inaction, and that this inaction was attributable to the challenges associated with global scale, and the company’s failure – until the crisis revealed its neglect to itself and the world – to build the capacity to engage in content moderation in the language in which the genocidal incitement was conducted. In 2021, Mozilla released a report suggesting that these problems continue on other social media platforms: Content identified by users as problematic was apparently something like 60 percent more likely “in countries that do not have English as a primary language” (Mozilla Foundation 2021). In addition to Myanmar, Facebook has struggled to prevent demagogues from using its platform to abuse their citizens in other international contexts;

²⁹ Balkin (2018) has described this as “new school speech regulation.” See also Bloch-Wehba (2019) along similar lines.

another prominent example comes from Duterte’s abuse of the platform in the Philippines (Alba 2018; Etter 2017).

As for unwilling: There are credible allegations that former Republican operative and current Facebook Vice-President Joel Kaplan successfully blocked measures within Facebook that would have at least attempted to tackle its contribution to extremism, polarization, and misinformation in the United States, allegedly on the ground that such measures were biased against conservatives (but, alas, conservatives were responsible for more of the pernicious content).³⁰

Other companies, including non-social-media platform companies, have similar problems. Amazon, for example, has notorious problems with counterfeit goods, to the point that it was moved in 2019 to list counterfeiting as an investment risk factor in its annual report (Kim 2019).

The foregoing examples also illustrate that unable and unwilling are (surprising nobody) hard to distinguish: When a platform could solve a governance problem by spending a lot of money, do we say that it’s unable to do so if the solution is particularly expensive? When it bows to political pressure, do we count that as in some sense volitional? The answers to such questions can only depend on the goals motivating their asking. There is no doubt some amount of money that Amazon could spend to effectively police counterfeits, especially when sellers use Amazon’s logistics services such that the counterfeit goods pass through its own warehouses. But it might be extremely costly. For example, with a truly staggering amount of money, Amazon might have experts manually inspecting every good passing through its warehouses – but I don’t think it would be reasonable to describe Amazon’s leaders as “unwilling” to prevent counterfeit products if the only way to do so is to spend a company-ruining amount of money.

Similarly, it took Facebook years longer to invest in tools like automated hate speech classifiers in Hindi and Bengali than it did in English (Zakrzewski et al. 2021) – had it thrown money at the problem earlier, it probably could have had those tools earlier, and may have mitigated its numerous problems in preventing violent and demagogic content in India (e.g., Frenkel and Alba 2021). But companies don’t have unlimited money, and not even pre-recession Facebook could throw money at machine learning in every language on Earth. Do we call Facebook unwilling to invest earlier because it failed to prioritize languages with such huge populations on its platforms? Do we call it unable because (ex hypothesi) it would have done so with unlimited resources? Does our intuition about this change depending on whether anyone in Menlo Park knew of the dangers posed by hate speech in India? Whether anyone in Menlo Park *should have* known about those dangers, for example, because they should have had better ways to learn about them, again, given the gigantic number of people at risk?

³⁰ Wofford (2022); Birnbaum (2021); Mac and Silverman (2021); Horwitz and Seetharaman (2020); Seetharaman (2018). See Chapter 4 for more details.

We might be more inclined to label a governance failure as a failure of will when we think that there are more normatively troubling conflicts of interest playing a large role. For example, Amazon makes profits even on counterfeited goods (so long as they go undetected further down the chain); Facebook users excited by political misinformation may stay on the platform longer and look at more advertisements. If companies fail to control profitable kinds of conduct creating third-party harms, we may be inclined to demand a higher degree of corporate economic burden before we treat that failure to control as a case of inability rather than lack of motivation.

By the same token, however, we ought to recognize that, as noted above, in many situations the economic incentives of platforms may give them strong reasons to effectively regulate their users. If Facebook allows itself to be turned into 4chan, then only the sorts of people who hang out on 4chan will go there – a much smaller and more pathological user base who cannot support a megabillion dollar company. And some of our major companies have actually felt significant economic bites from governance failures: YouTube had an exodus of advertisers in 2017 due to a “brand safety” scandal after major companies found their products advertised on extremist videos (Solon 2017); Nike and Birkenstocks stopped selling their products directly on Amazon due to counterfeiting (Bain 2019). As of this writing, Elon Musk’s erratic leadership of Twitter is creating comically extreme brand safety threats for advertisers – shortly after he acquired the company, he permitted anyone to purchase a “verified” checkmark for \$8, and verified accounts shortly came into being parodying numerous major corporations. The most grimly amusing example: Someone bought a checkmark and then, under the name “Eli Lilly,” falsely (alas) declared that insulin would be given away for free. Unsurprisingly, the company canceled its Twitter advertising (Harwell 2022).

If a company is experiencing those kinds of consequences and nonetheless fails to control the behavior that leads to them, it gives us some reason to interpret those failures as rooted in inability, due, for example, to technical difficulty or to divergent incentives between the top-level leaders whose intentions animate the company’s goals and lower-level employees implementing those intentions (on which more in Chapter 4). For example, two years after many advertisers announced a boycott of Facebook in protest of its hosting and profiting from white supremacist content, it still – company representatives say inadvertently – serves advertisements against such content and in some cases algorithmically generates white supremacist pages based on user interests (Nix 2022). Similarly, in September 2022, Twitter (pre-Elon!) discovered that it had run advertisements from several major companies like Coca-Cola on accounts full of child sexual abuse material (CSAM) (Fingas 2022).³¹ In view of the fact that such content violates both companies’ stated policies, has

³¹ Incidentally, this also illustrates the extreme difficulty of automated enforcement, since there are vast databases of CSAM which companies use to detect image “fingerprints” through the PhotoDNA program – but that isn’t enough to keep it reliably off Twitter.

led to financial consequences from advertisers, and in the case of Twitter's CSAM problem is also a universally loathed major crime throughout the planet, it's pretty tempting to see the failure to eliminate it as "inability" rather than "unwillingness," at least relative to existing levels of company investment in enforcement.

Our choice of whether to label a company's governance failures as a case of inability or unwillingness, in turn, may implicate the strategies we choose as democratic citizens operating governments to remediate them. In the case of inability, we have some reason to prefer interventions on platform features like their underlying corporate governance or employee relations, or the technologies available to them (e.g., by organizing licensing schemes for such technologies). By contrast, in cases of unwillingness, we have reason to prefer more coercive interventions such as the threat of fines and antitrust action. However, I shall suggest in this book that there are some interventions that can actually avoid the inability/unwillingness dichotomy, insofar as they can remediate both problems at once, that is, to give platforms both the ability and the incentive to govern their users adequately.

It is worth identifying that there is, of course, an unavoidable amount of disagreement and contestation on what adequate governance looks like. The example of right-wing misinformation on social media could serve again: Many on the political right would deny that much of the material in question actually does misinform people. Regardless of the proportion of the debatable material that is removed, those on the left will always think that it's not enough, those on the right too much. Still, this project does not require us to resolve such questions. One of the tasks of any effort to build governing capacity includes building the capacity to come to reasonable, even if imperfect, resolutions of highly controversial cases.

The reader may object that if adequate governance can't be observed, then we can't tell whether any program to develop it is a success. To this, my response is twofold. First, we can observe improvements in uncontroversially terrible user behavior, such as despotic militaries inciting genocide and Russian spies pretending to be Black Lives Matter activists on social media, or counterfeit products on Amazon. Second, we can observe relatively uncontroversial process improvements, such as the inclusion of minorities of all stripes (racial, religious, ethnic, etc., relative to a country or a problem, for example). Given where we stand today with the immense number of social harms created by platforms, we should work to solve the easy (in an evaluative sense) cases before worrying about the hard ones.

SCHOLARS HAVE THE TOOLS TO IMPROVE PLATFORM GOVERNANCE: BORROWING FROM POLITICAL GOVERNANCE

In the tradition of academic books, part of my mission is to fill a surprising gap in the scholarly literature. There are, of course, countless scholars writing about platform governance, from disciplines such as law (one of my home fields), communications,

science and technology studies, and the like. And there are numerous scholars in political science (my other home field) and allied fields such as political economy writing about the problems associated with platforms (e.g., Tucker et al. 2018; Zhuravskaya, Petrova, and Enikolopov 2020). However, the political science literature mostly resides in the behavioral side of the discipline, that is, from scholars who empirically study how people participate in political activity.

Yet there's another side to political science, traditionally known as "institutions." Scholars in that half of the discipline, in conversation with allied disciplines like economics, sociology, and history – as well as more applied disciplines such as public policy, urban planning, and ecology – write about the effects of different organizational forms and patterns of interaction on aggregate behavior. Such scholars consider questions such as the conditions under which it might be possible to bind people (particularly, but not exclusively, top-level officials) to complying with legal rules (e.g., Hadfield and Weingast 2014; De Lara, Greif, and Jha 2008), how the independence of judges is preserved (e.g., Ferejohn 1999), the relationship between different types of property rights and the ability to manage shared resources (e.g., Ostrom 2003), and the like.³²

There is a dearth of work from the institutional tradition of political science and its allied disciplines noted above on the problem of platform governance itself. While there are a handful of articles in the vein about specific problems and specific platforms, there is no comprehensive or book-length treatment.³³ This book aims to begin the conversation on that broader basis.³⁴

³² Fascinatingly, there is a surprising affinity between some of the modern governance literature in the disciplines I have described and a more technologically oriented and science-fiction sounding discipline with essentially no direct intersection with political science, namely, "cybernetics" (Beer 2002). More or less, as far as I can discern, the cybernetics people talk to ecologists and complexity theorists, and then complexity theorists and ecologists talk to political scientists (mostly thanks to Ostrom). And the technology people sometimes talk to the cybernetics people because the name sounds like something out of science fiction. Thus, the bizarre sociology of the intelligentsia. For one story about the intersection between cybernetics and the kinds of organizational political theory that partly animate this book through the lens of anarchist(!) philosophy, see Swann (2018). Some of the ideas described below, such as on the ineffectiveness of rigidly centralized control, also feature in the cybernetic literature (Swann 2018, 433).

³³ Some examples of the most important article-length work setting up the foundations of this nascent literature from the political science and allied side include Gorwa (2019b, 2019a), Napoli (2014, 2015), Srivastava (2021), and Caplan and Boyd (2018).

³⁴ This approach is also self-consciously a product of my own unusual intellectual location. There is an active platform governance literature which I consume (although given the volume and rate at which scholarship in platform governance is produced, I've probably missed important work), but in which I have not previously been a participant. There is a sense in which I intrude on that conversation as an outsider – as a political theorist and constitutional scholar drawing on external fields to intervene in an existing literature, with all the advantages (cross-pollination of ideas, a fresh perspective) as well as the disadvantages (the risk of repeating or missing ideas extant in that community, the potential misuse of internally generated terms of art) this entails. At the same time, I am unusually familiar with the problem space from an odd angle: As described in the Appendix to this Introduction, I have worked within one of the most major platform companies on initiatives to address some of its most important

This book rests heavily on a broad cluster of theoretical ideas that has crossed many disciplines and been associated with a number of prominent scholars. The lodestar points of this cluster include, among other things, Ostrom's (2015, 2010a, 2010b) work on polycentric governance, Hayek's (1945) on the problem of knowledge, Scott (2008) on high modernism, Jacobs (1992) on urbanism, Dewey's (1927) democratic experimentalism, Ober's (2008) historical work on democracy and knowledge in classical Athens, and a mass of work in public administration and related fields that often goes under the name "New Governance" associated, for example, with Mark Bevir (2013).³⁵ This seemingly diverse set of ideas tends to converge, for different reasons, on propositions such as the following:

- Centralized top-down command-giving is often ineffective because of its difficulties with integrating knowledge from the periphery and offering legitimate rules to diverse constituencies.
- Many effective institutions of governance are grown or evolved out of the immanent behavior of people trying to solve their own problems, rather than designed or imposed.
- Novel governance strategies can be developed by permitting some play in the space between means and ends, for example, by creating local sub-institutions empowered to develop experimental or even idiosyncratic techniques to pursue shared goals in the context of dense cross-institution communication and learning.
- Rules and governing institutions frequently require revision in the light of practical experience with their operation.
- Agents and organizations engaged in the activities of governing can often be more effective when organized into complex structures including features such as overlapping and multi-scale jurisdictions and collaborative networked relationships drawing on markets and informal social interactions.

Yet despite the challenges to centralized governance and the empirical successes of alternative forms represented by this literature, the major platform companies uniformly have a centralized, top-down, authoritative governance structure for user behavior. With very few exceptions (mostly Reddit, Discord, Wikipedia,³⁶ and to a

governance challenges, and maintain a close and active engagement with workers and former workers on these problems from numerous companies through a nonprofit organization (the Integrity Institute) in which I have had a high degree of involvement. I am also a longstanding participant in the conversation on governance more generally through the central normative construct of the "rule of law," on which I have previously published two books and numerous articles. So I am, somewhat bizarrely, an outsider to the academic literature on platform governance but not to the problem of platform governance in actual implementation nor to the problem of governance in the abstract. I ask the reader to consider the arguments offered in this volume in that light, and excuse any distortions which I inadvertently impose on existing scholarship in the more narrow "platform governance" academic field.

³⁵ On the relationship between Dewey and new governance, see Simon (2010).

³⁶ While I mention Wikipedia in several places for the purposes of comparison to some of its governing entities, it does not meet the definition of a platform as used in this volume.

limited extent the Meta Oversight Board), the rules are made in a corporate hierarchy in somewhere like Menlo Park or Seattle, and enforced by a combination of machine learning algorithms and human enforcers directly answerable to the corporate chain of command and nobody else. And this is so even though the major platforms are managing globe-spanning user conduct, with deeply interconnected networks of people generating complex emergent patterns of behavior in a context of extreme diversity – quite possibly the least suitable setting for the centralized command-and-control style.

The enterprise of this book is also inspired by (although does not as directly deploy) intellectual frameworks that have long recognized that a higher-level abstraction can be used to analyze both states and firms, namely, the organization. Some of the foundational work in the political institutions research program is built on the recognition that business companies and political entities face similar organizational and governance problems, and consciously applies economic theories like Coase and Williamson's theories of the firm (e.g., Williamson 2005) to political states (e.g., Weingast and Marshall 1988; Moe 1984; North and Weingast 1989). On the other end of the social sciences, organizational theorists in sociology (W. R. Scott 2004) such as March, Olsen, Powell, and DiMaggio (e.g., March and Olsen 1984) have identified that ideas like diffusion of strategies and logics of legitimacy and appropriateness apply across organizational contexts. Early work in the new economics of organization specifically attends to governance as an enterprise that can take different structural forms – most famously understanding firms and markets as alternative ways of arranging transactions (Williamson 1996, 133).

As yet, efforts to apply these theoretical frameworks to platforms are in their early stages. The most interesting exception is Marxist economist Laurent Baronian's (2020) effort to conceptualize platforms' relationships with "users" (conceptually centered on, but not limited to, transactional platforms' relationships with worker-users, as with Uber drivers) as novel solutions to the management problem of determining the boundaries between firm and market. But I contend we can learn more by explicitly drawing from the application of these theoretical frameworks not just to firms but also to states.

The discussion in this volume will also be guided by both the strategic and the normative. The governance literature, like most contemporary political science and economics, tends to be focused on the management of the strategic incentives of participants (here, companies, users, and governments) as well as the structural features of an interactive environment that make it possible for participants to respond to those incentives (e.g., the sharing of information, and second-order incentives to conduct that sharing, and so forth). But the evaluation of the predicted outcomes of those incentives, and thus any recommendations in terms of actual policy or design outputs, necessarily depends on external normative standards.

The core normative presuppositions of this book can be described in terms of the concepts of democratization and inclusion. That is (very briefly, and with elaboration spread through the rest of the pages that follow), I suppose that (1) people ought to be able to run their own lives, collectively, through regulatory institutions that are accountable to the regulated,³⁷ and (2) we should be alert to the way in which governing arrangements can go wrong by failing to identify the people who should rightly be included. The latter category, for example, includes the danger that those in the so-called “developed” world will economically and politically dominate those in the “developing” world in ways that are objectionable both because they undermine the self-governance of the latter and because they represent the continuing unjust legacy of historical conquest and colonization. But that category also includes existing biases and exclusions within countries that company governance might replicate, such as the exclusion of racial, religious, gender, and cultural minorities.

There’s a close affinity between this book’s democratic and inclusive normative side and its pragmatic focus on modern theories of governance. I shall argue that the best path forward for democratizing platform governance and for including the legitimate claims of those who are not the powerful elites from the wealthy developed world in the decision-making processes involves the creation of institutions that can incorporate ordinary people into polycentric and densely interconnected governance processes much like those described by modern governance scholarship (but with a broader popular element than is traditional for the sort of “new governance” that mostly runs together NGOs, governments, and corporations). This is not merely coincidental: In a highly diverse ecosystem operating at an immense scale, like every major platform, the imperatives of (normative) legitimacy and the imperatives of effectiveness come together, for both demand the deep-down inclusion of a wide and ever-expanding variety of knowers and stakeholders.

A final key influence that requires specific discussion in the Introduction – for it has guided the entire book in a variety of subtle or explicit ways – is John Dewey’s major work of political theory, *The Public and Its Problems*. Dewey begins, as the title suggests, by giving an account of the domain of the public, namely those interactions between people which have an effect on third parties (what contemporary economists like to call “externalities”). The nature of the state, he argues, cannot be discovered by some theoretical derivation from first principles but rather arises out of the efforts to solve these third-party effects – and thus the nature of the state is different under different social, technological, and economic conditions. Moreover,

³⁷ This one-clause summary is self-consciously neutral with respect to whether the implication of this idea in the platform context is that democratically elected governments ought to control platforms or their users, or that people ought to have democratic control over those platforms directly – as it will turn out, the answer I offer is somewhat more complex, and will require some development.

this suggests that there is an experimental quality to the organization of states and that the goal of the study of politics is to help build the conditions under which experimentation and learning can be successful.

The Deweyian approach to the state (a kind of pragmatist Hegelianism) is perhaps best summarized in the following passage, worth quoting at length:

In no two ages or places is there the same public. Conditions make the consequences of associated action and the knowledge of them different. In addition the means by which a public can determine the government to serve its interests vary. Only formally can we say what the best state would be. In concrete fact, in actual and concrete organization and structure, there is no form of state which can be said to be the best: not at least till history is ended, and one can survey all its varied forms. The formation of states must be an experimental process. The trial process may go on with diverse degrees of blindness and accident, and at the cost of unregulated procedures of cut and try, of fumbling and groping, without insight into what men are after or clear knowledge of a good state even when it is achieved. Or it may proceed more intelligently, because guided by knowledge of the conditions which must be fulfilled. But it is still experimental. And since conditions of action and of inquiry and knowledge are always changing, the experiment must always be retried; the State must always be rediscovered. Except, once more, in formal statement of conditions to be met, we have no idea what history may still bring forth. It is not the business of political philosophy and science to determine what the state in general should or must be. What they may do is to aid in creation of methods such that experimentation may go on less blindly, less at the mercy of accident, more intelligently, so that men may learn from their errors and profit by their successes.³⁸

This perspective seems entirely sound to me when confronting novel forms of governance with a novel underlying “public” and trying to figure out what to do about it. While Dewey assumed that states would be geographically contiguous (Dewey 1927, 39–43, 212–13 – he was, after all, writing almost a century ago), I think that he would find the notion that platforms have their own publics, and thus are at least in the same family as states to be fairly congenial.³⁹ He would probably agree that the process of figuring out how to govern externalities generated by activities in these novel, somewhat state-like, entities is at heart experimental, and that what we must do is aim to build the conditions in which that experimentation can be carried out. I argue that ultimately those conditions include a kind of radical inclusiveness that recognizes that the Deweyian public for platforms is global, and that it is this global public which must be permitted to experiment. I further argue that even the scope of inclusion, that is, who is to be involved in running the institutions of platform governance itself, is itself ineluctably experimental.

³⁸ Dewey (1927, 33–34).

³⁹ He did recognize that political boundaries could diverge from the publics constituted by the effects of economic activity and change in communication technologies. See *ibid.*, 107, 114.

Because demands for inclusion are likely to come from unanticipated directions, a chief design criterion for platform governance will be to build institutions capable of responding appropriately to those demands, and thus learning from contestation even over the scope of its own stakeholders. This is all, to my mind, deeply Deweyian and Dewey's influence runs throughout this book.

WHERE WE'RE GOING

The remainder of this book is divided into three substantive parts. The first part, consisting of Chapters 1 and 2, further develops the general approach in the Introduction. Chapter 1 more carefully defines the platforms under consideration and describes their general characteristics.

Chapter 2 draws the analogy that drives the approach of this book between platforms and states – in particular, failed states, which lack the ability or the incentive to adequately govern those who use their services. It addresses several major objections to the enterprise of adopting a capacity-building approach to the project of platform governance, where our political states work to shape the incentives and abilities of platforms in order to govern their users.

Part II, consisting of Chapters 3 and 4, returns to the problems that lead this Introduction. Chapter 3 begins with the Myanmar genocide and argues that Facebook's culpability in that genocide resulted from a characteristic problem, also experienced by governments, of bringing knowledge from the periphery of the governed domain to the center. The chapter argues that democratic institutions, organized in ways attentive to the dispersal of authority and the aggregation of information across space and scale, can mitigate such knowledge problems.

Chapter 4, in turn, takes up the problem of Donald Trump – both the propaganda that led to his election and the companies' seeming inability to control his supporters while in office – leading up to, but not limited to, the January 6 coup attempt. It contextualizes these events in a broader narrative about social media “political bias,” which I interpret less as a genuine problem of bias and more as an effort by politicians to intimidate companies into under-moderating their sides. Such efforts aim to leverage the inability of the companies to exercise self-control in the face of short-term temptations and threats. Platforms, like governments, have problems of internal governance, in which personnel have incentives that diverge from the interests of the overall organization or operate under suboptimal time horizons. In the literature on governments, the tools to mitigate these problems tend to travel under the rubric of “the rule of law.” I defend the idea of dispersing power to independent institutions under the control or at least supervision of diverse groups of employees and non-employees as a key tool to create a kind of platform rule of law.

Ultimately, the two chapters of Part II point toward the same primary ideas. To wit: the social organization of governance matters; effective governance requires that

people with knowledge (about what is happening, about their needs) and distinctive interests be assembled in network structures where they have an incentive to share their knowledge – incentive conferred in substantial part by genuine power over outcomes that matter to them and the capacity and incentive to negotiate over the use of shared resources. Doing so ultimately requires the conscious building of inclusive processes in which currently under-represented stakeholders, such as those in the global South, as well as those who have the latent power to control platform companies, such as their own workers (broadly understood) are organized into groups with overlapping authority over key governance decisions. The title of this book – *The Networked Leviathan* – thus has a double meaning, referring first to the existing nature of platform companies, which occupy quasi-governmental roles due to their leveraging of network effects; but also referring to the capacity of interventions on the network structure of platform users and workers (and secondarily governments and civil society) to create a new kind of governing structure. The title is also ironic, for, *contra* Hobbes, I ultimately argue that Leviathan must share rather than hoard his power.

Finally, Part III, consisting of Chapters 5 and 6 as well as the Conclusion, turns to practical implementation of the book's overall approach. Chapter 5 examines the design and performance of the most prominent recent innovation in platform governance, the Facebook (now Meta) Oversight Board.⁴⁰ It contextualizes the examination of that board in the rule of law ideas of the preceding chapter. In the course of this analysis, it also develops some ideas about the normative function of platform legalism in building a kind of platform identity which may be valuable in resolving controversial governance issues.

Chapter 6 makes concrete proposals for the design of polycentric, decolonial, democratic governance institutions in the platform economy. Chapter 6 focuses on institutions that platforms could, in principle, build themselves; it could be understood in the first instance as being directed at senior platform executives. The end of Chapter 6 and then the Conclusion address governments, describing some ideas for ways they (particularly, but not exclusively, the United States and the European Union) could give companies the incentives necessary to implement some of these reforms, as well as other direct beneficial interventions that states could make on platform governance.

APPENDIX TO INTRODUCTION: ADDRESSING SOME ETHICAL CHALLENGES

Before turning to the substance, I must disclose and address some potential issues of research ethics. I do not write as a fully independent observer. Most importantly, I've taken money from Facebook to apply my academic knowledge on areas closely

⁴⁰ I had some involvement in the design of the board; see the Appendix to this Introduction for a full disclosure and description.

related to the topic of this book. From the summer of 2018 through the end of 2019, I was a paid consultant/contractor for Facebook: For several months in 2018, I was embedded full time in the company's Civic Integrity team advising on considerations of democratic theory and ethics relating primarily to elections; thereafter, I continued to consult on a part-time basis with that team as well as with Facebook's Product Policy Research team supporting the design of the Facebook Oversight Board. In the latter capacity, I co-wrote a report with Radha Iyengar Plumb, then Facebook's Head of Product Policy Research, which Facebook distributed in conjunction with an announcement about the Board.⁴¹

In addition to paid work, I also have other interactions with the technology industry that arguably might undermine my research objectivity. I've given informal (and unpaid) consultations to staff from other prominent technology companies who have sought my advice about their governance institutions. I've also been a participant in industry conversations, including conversations under confidentiality obligations meant to facilitate frank conversation. (This paragraph uses the nonspecific plural largely to preserve a degree of confidentiality in all of the interactions it describes.) Finally, I've been involved with research and advocacy organizations seeking to produce reforms in how social media companies do content moderation. I am a founding fellow of the Integrity Institute, a nonprofit organization composed of present and former "integrity workers" at social media companies dedicated to advising policymakers, press, and public on the challenges of platform integrity; I am also a member of the Integrity Institute's nonprofit (corporate) board as well as its Community Advisory Board.

These activities might reasonably raise several concerns in a reader of this book. First, I might be compromised by past financial gain or the anticipation of future financial gain: The arguments in this book might be unconsciously or even consciously influenced by a pro-company or just pro-Meta bias as a result of my financial interests. This is potentially particularly problematic because the company that paid me was Facebook, which has a record of getting journalists and academics into compromising financial entanglements.⁴²

Second, confidentiality obligations that I've accepted, either formally – my work at Facebook was done under an industry-standard nondisclosure agreement – or

⁴¹ Paul Gowder and Radha Iyengar Plumb, "Oversight of Deliberative Decision-making: An Analysis of Public and Private Oversight Models Worldwide," Appendix E to "Global Feedback and Input on the Facebook Oversight Board for Content Decisions," released June 27, 2019, <https://about.fb.com/wp-content/uploads/2019/06/oversight-board-consultation-report-appendix.pdf#page=138>.

⁴² For example, *New York Times* columnist David Brooks recently failed to disclose to his editors that he was producing a corporate blog post including a paean to the wonders of Facebook Groups for the company, to promote a study by researchers at NYU on how wonderful they were; while Brooks apparently was not directly paid for this work, something called the "Weave Project," which he founded, at the Aspen Institute, had received funding from Facebook, as did the NYU researchers whose report Brooks's blog post for the company was meant to introduce (Silverman and Mac 2021).

informally, such as under “Chatham House Rule” norms in various conversations in which I’ve participated, might compromise my ability to fully use knowledge that I’ve gained in these relationships to improve the scholarship represented by this book. From the other direction, my work in this book might violate those confidentiality commitments.

Third, I may have a bias as a result of the tasks I have done and the social contexts in which I’m embedded – I associate with many past and present platform company employees and directly worked on some of the things discussed in the following pages. Nobody wants to come to the conclusion that their own efforts were wasted, or the things that they and their friends have worked on are harmful. Self-justification bias is a powerful force.⁴³

To some extent, these problems go along with the territory.⁴⁴ Without having relationships with companies and others associated with the industry, I wouldn’t have the knowledge necessary to write this book sensibly – one of the most unfortunate aspects of the current debate on platform regulation is the extent to which it’s dominated by, on one side, critics of the industry with no information as to the actual constraints companies face; on the other, industry insiders with expert knowledge of those constraints but tainted motivations. I am trying to straddle that gap, but I may not be successful.

In an effort to ameliorate these problems, I’ve taken the following steps. First, of course, is this Appendix: I’ve endeavored to write a full disclosure that communicates all relevant facts, provided early on in the main text of the book rather than buried in a footnote or at the end.

Second, I have taken steps to sever my personal financial interests from those of the companies under discussion. As of the sending of this book to press, I have not done any paid work for Meta or any other technology company since the end of 2019, and I do not have any offers, ongoing discussions, or plans to do any in the future. (However, I have not committed to refusing any such work that may come along after this book is published.) Nor have I or any institution with which I’m affiliated received any research support for the production of this book from any for-profit company; the only financial support outside of Northwestern University’s ordinary resources that contributed to this book was a grant from the Knight Foundation. Similarly, I have, to my knowledge, no discrete investments in any technology company, although I do have mutual fund investments over large portions of US and

⁴³ I think this is the least important of the risks. The place where it is most likely to crop up is in a bias toward believing that the Meta Oversight Board is a good idea, but, as you will see in Chapter 5, I have no qualms about being critical of the Board’s design where warranted. Moreover, part of the reason I agreed to work on the Oversight Board was because I think it’s fundamentally a good idea – as evidenced by the fact that I published scholarship in favor of the general idea of independent judges long before I had any affiliation with Facebook (e.g., Gowder 2014b, 2016).

⁴⁴ Tarleton Gillespie (2022, 2–3) has a “methodological note” in a recent article of his which is particularly thoughtful on these kinds of issues.

international markets, including the technology sector, in my personal investment and retirement accounts, which are managed by professional fund managers without my input in choosing individual stocks.⁴⁵

Third, on the informational side, no company has reviewed or approved the manuscript for this book. Because I am covered by a nondisclosure agreement relating to all work done for Meta, I have been very careful to identify publicly reported sources for any information about Meta discussed in this book.⁴⁶

⁴⁵ However, I acknowledge that this does not completely free me from potential financial bias. My views might be biased by the hope of *future* financial opportunities from platform companies. But that's true of anyone who writes academic work about a topic in which wealthy companies are also interested; there's nothing special about what I've done in the past that changes this risk.

⁴⁶ In practice, when I have discussed more sensitive or controversial issues involving Meta this has amounted to citing many news articles drawn from the so-called "Facebook Files," documents leaked by whistleblower Frances Haugen. I do not assert the truth or falsity of anything reported in any of these sources, which should stand or fall on their own.